# Supplement to "Evaluation of version 3.0B of the BEHR $OMI NO_2$ product"

Joshua L. Laughner, Qindan Zhu, and Ronald C. Cohen

## 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.

Campaign	Product	Extend Slope	led with GEOS Intercept	S-Chem $R^2$	Exten Slope	ded by extrapc Intercept	lation $R^2$
DISCOVER-MD	BEHR v3.0B (D) BEHR v3.0B (M) BEHR v2.1C SP v3.0	N/A 0.888 1.92 0.774	$\begin{array}{c} {\rm N/A} \\ -8.96 \times 10^{14} \\ -2.69 \times 10^{15} \\ -5.35 \times 10^{14} \end{array}$	N/A 0.149 0.0747 0.0806	N/A 0.998 2.27 0.884	$\begin{array}{c} {\rm N/A} \\ 1.28 \times 10^{14} \\ -7.24 \times 10^{14} \\ 3.67 \times 10^{14} \end{array}$	N/A 0.0813 0.0921 0.0469
DISCOVER-CA	BEHR v3.0B (D) BEHR v3.0B (M) BEHR v2.1C SP v3.0	$\begin{array}{c} 0.627 \\ 0.665 \\ 0.944 \\ 0.912 \end{array}$	$\begin{array}{c} 1.12 \times 10^{15} \\ 9.91 \times 10^{14} \\ 8.29 \times 10^{14} \\ 9.66 \times 10^{14} \end{array}$	$\begin{array}{c} 0.567 \\ 0.521 \\ 0.554 \\ 0.291 \end{array}$	$\begin{array}{c} 0.716\\ 0.775\\ 1.14\\ 1.14\\ 0.513\end{array}$	$\begin{array}{c} 9.6 \times 10^{14} \\ 7.98 \times 10^{14} \\ 5.25 \times 10^{14} \\ 9.93 \times 10^{14} \end{array}$	$\begin{array}{c} 0.587 \\ 0.554 \\ 0.625 \\ 0.311 \end{array}$
DISCOVER-TX	BEHR v3.0B (D) BEHR v3.0B (M) BEHR v2.1C SP v3.0	$\begin{array}{c} 0.483 \\ 0.407 \\ 0.867 \\ 0.299 \end{array}$	$\begin{array}{c} 2.66 \times 10^{14} \\ 3.9 \times 10^{14} \\ 1.03 \times 10^{15} \\ 7.15 \times 10^{14} \end{array}$	$\begin{array}{c} 0.295\\ 0.321\\ 0.386\\ 0.352\\ 0.352\end{array}$	$\begin{array}{c} 0.512 \\ 0.425 \\ 0.836 \\ 0.317 \end{array}$	$\begin{array}{c} 4.54 \times 10^{14} \\ 5.54 \times 10^{14} \\ 1.58 \times 10^{15} \\ 7.9 \times 10^{14} \end{array}$	$\begin{array}{c} 0.275 \\ 0.238 \\ 0.359 \\ 0.355 \end{array}$
DISCOVER-CO $(V > 0)$	BEHR v3.0B (D) BEHR v3.0B (M) BEHR v2.1C SP v3.0	$1.03 \\ 0.915 \\ 1.84 \\ 0.705$	$\begin{array}{c} -5.82 \times 10^{14} \\ -5.27 \times 10^{14} \\ -1.32 \times 10^{15} \\ -4.04 \times 10^{14} \end{array}$	$\begin{array}{c} 0.0279\\ 0.0244\\ 0.018\\ 0.0553\end{array}$	$1.01 \\ 0.957 \\ 1.84 \\ 0.65$	$\begin{array}{c} 1.26 \times 10^{14} \\ 8.24 \times 10^{13} \\ -1.65 \times 10^{14} \\ 1.69 \times 10^{14} \end{array}$	$\begin{array}{c} 0.0234 \\ 0.062 \\ 0.0186 \\ 0.0423 \end{array}$
SENEX	BEHR v3.0B (D) BEHR v3.0B (M) BEHR v2.1C SP v3.0	$\begin{array}{c} 2.25\\ 0.953\\ 1.43\\ 1.07\end{array}$	$\begin{array}{c} -2.59 \times 10^{15} \\ -4.6 \times 10^{14} \\ -2.25 \times 10^{14} \\ -5.84 \times 10^{14} \end{array}$	$\begin{array}{c} 0.154 \\ 0.265 \\ 0.164 \\ 0.274 \end{array}$	$\begin{array}{c} 1.72 \\ 0.879 \\ 1.48 \\ 0.835 \end{array}$	$\begin{array}{c} -1.2 \times 10^{15} \\ -8.89 \times 10^{13} \\ -2.44 \times 10^{14} \\ 2.35 \times 10^{13} \end{array}$	$\begin{array}{c} 0.0549 \\ 0.153 \\ 0.111 \\ 0.107 \end{array}$
SEAC4RS	BEHR v3.0B (D) BEHR v3.0B (M) BEHR v2.1C SP v3.0	$\begin{array}{c} 0.888 \\ 1.22 \\ 2.63 \\ 1.01 \end{array}$	$\begin{array}{c} -1.51 \times 10^{14} \\ -4.56 \times 10^{14} \\ -1.55 \times 10^{15} \\ -3.97 \times 10^{14} \end{array}$	$\begin{array}{c} 0.0122 \\ 0.228 \\ 0.597 \\ 0.236 \end{array}$	$\begin{array}{c} 0.695 \\ 0.971 \\ 2.53 \\ 0.823 \end{array}$	$\begin{array}{c} 5.89 \times 10^{14} \\ 5.81 \times 10^{14} \\ -7.82 \times 10^{13} \\ 4.31 \times 10^{14} \end{array}$	$\begin{array}{c} 0.0473 \\ 0.376 \\ 0.422 \\ 0.154 \end{array}$

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 S1: Scatter plots comparing  $(\mathbf{a}, \mathbf{c}, \mathbf{e})$  NASA Standard Product and  $(\mathbf{b}, \mathbf{d}, \mathbf{f})$  the BEHR product VCDs to  $(\mathbf{a}, \mathbf{b})$  aircraft profiles extended with GEOS-Chem,  $(\mathbf{c}, \mathbf{d})$  aircraft profiles extended by extrapolation,  $(\mathbf{e}, \mathbf{f})$  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.



Figure S2: Scatter plots comparing  $(\mathbf{a}, \mathbf{c}, \mathbf{e})$  NASA Standard Product and  $(\mathbf{b}, \mathbf{d}, \mathbf{f})$  the BEHR product VCDs to  $(\mathbf{a}, \mathbf{b})$  aircraft profiles extended with GEOS-Chem,  $(\mathbf{c}, \mathbf{d})$  aircraft profiles extended by extrapolation,  $(\mathbf{e}, \mathbf{f})$  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  $(\mathbf{a}, \mathbf{c}, \mathbf{e})$  NASA Standard Product and  $(\mathbf{b}, \mathbf{d}, \mathbf{f})$  the BEHR product VCDs to  $(\mathbf{a}, \mathbf{b})$  aircraft profiles extended with GEOS-Chem,  $(\mathbf{c}, \mathbf{d})$  aircraft profiles extended by extrapolation,  $(\mathbf{e}, \mathbf{f})$  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  $(\mathbf{a,c,e})$  NASA Standard Product and  $(\mathbf{b,d,f})$  the BEHR product VCDs to  $(\mathbf{a,b})$  aircraft profiles extended with GEOS-Chem,  $(\mathbf{c,d})$  aircraft profiles extended by extrapolation,  $(\mathbf{e,f})$  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) BEHR v3.0B (M) BEHR v2.1C SP v3.0	N/A 0.443 0.766 0.355	$ \begin{array}{c} {\rm N/A} \\ 2.55 \times 10^{15} \\ 3.2 \times 10^{15} \\ 3.02 \times 10^{15} \end{array} $	N/A 0.139 0.128 0.125
DISCOVER-CA	BEHR v3.0B (D) BEHR v3.0B (M) BEHR v2.1C SP v3.0	$\begin{array}{c} 0.813 \\ 0.781 \\ 0.774 \\ 0.599 \end{array}$	$\begin{array}{c} -9.99\times 10^{14} \\ -8.66\times 10^{14} \\ -1.22\times 10^{14} \\ -1.01\times 10^{14} \end{array}$	$\begin{array}{c} 0.175 \\ 0.16 \\ 0.178 \\ 0.169 \end{array}$
DISCOVER-TX	BEHR v3.0B (D) BEHR v3.0B (M) BEHR v2.1C SP v3.0	$     1.08 \\     0.868 \\     1.43 \\     0.688 $	$\begin{array}{c} -7.57\times10^{14}\\ 2.05\times10^{14}\\ 6.54\times10^{14}\\ 9.54\times10^{14} \end{array}$	$\begin{array}{c} 0.152 \\ 0.173 \\ 0.102 \\ 0.136 \end{array}$
DISCOVER-CO $(V > 0)$	BEHR v3.0B (D) BEHR v3.0B (M) BEHR v2.1C SP v3.0	$\begin{array}{c} 0.655 \\ 0.628 \\ 0.658 \\ 0.468 \end{array}$	$\begin{array}{c} 7.71 \times 10^{14} \\ 8.26 \times 10^{14} \\ 1.5 \times 10^{15} \\ 1.38 \times 10^{15} \end{array}$	$\begin{array}{c} 0.211 \\ 0.243 \\ 0.224 \\ 0.213 \end{array}$

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

Campaign P							
	roduct	Slope	Intercept	$\mathbb{R}2$	Slope	Intercept	$\mathbb{R}2$
B	3EHR v3.0B (D)	N/A	N/A	N/A	N/A	N/A	N/A
DISCOVER-MD B	3EHR v 3.0B (M)	0.804	$-4.16 \times 10^{14}$	0.407	0.637	$8.46 \times 10^{14}$	0.361
B	SEHR v2.1C	1.27	$-2.92 \times 10^{14}$	0.292	0.874	$2.23  imes 10^{15}$	0.304
S	P v3.0	0.788	$-4.65\times10^{14}$	0.383	0.496	$1.72  imes 10^{15}$	0.286
B	3EHR v3.0B (D)	0.493	$1.3 \times 10^{15}$	0.525	0.677	$1.27 \times 10^{14}$	0.321
	3EHR v 3.0B (M)	0.509	$1.18  imes 10^{15}$	0.514	0.661	$1.5  imes 10^{14}$	0.317
	SEHR v2.1C	0.568	$1.46 \times 10^{15}$	0.469	0.682	$6.54 \times 10^{14}$	0.321
S	P v3.0	0.407	$1.31 \times 10^{15}$	0.408	0.539	$4.11 \times 10^{14}$	0.282
B	BEHR v3.0B (D)	0.692	$1.35 \times 10^{14}$	0.157	1	$-7.39 \times 10^{14}$	0.208
	3EHR v 3.0B (M)	0.598	$3.22 \times 10^{14}$	0.217	0.871	$-2.26 \times 10^{14}$	0.221
	SEHR v2.1C	1.06	$1.07 imes 10^{15}$	0.19	1.33	$7.16  imes 10^{14}$	0.173
S	P v3.0	0.527	$5.25 \times 10^{14}$	0.12	0.736	$2.45 \times 10^{14}$	0.136
B	3EHR v3.0B (D)	0.663	$6.38 \times 10^{14}$	0.394	0.664	$6.92 \times 10^{14}$	0.267
	BEHR v3.0B (M)	0.696	$2.68 \times 10^{14}$	0.347	0.628	$7.56 \times 10^{14}$	0.277
	3EHR v2.1C	0.741	$1.09  imes 10^{15}$	0.422	0.679	$1.33  imes 10^{15}$	0.326
S	P v3.0	0.531	$4.32 \times 10^{14}$	0.312	0.498	$1.05  imes 10^{15}$	0.224

Matched slopes use only Pandora data approximately coincident with aircraft profiles to get similar sampling; all uses all valid Table S3: Slopes and  $1\sigma$  uncertainties of BEHR vs. combined aircraft (extended with GEOS-Chem profiles) and Pandora VCDs. Pandora data. Outliers and negative VCDs are removed before computing slopes.

# S2 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. S7. 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. S8). The simulated flash density in the convection core is lower than observation despite the total flash counts are still comparable.



Figure S7: 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 S8: 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.

### S3 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:

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

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

- I: the top-of-atmosphere intensity at wavelength  $\lambda$  for the given viewing zenith angle (VZA,  $\theta$ ), solar zenith angle (SZA,  $\theta_0$ ), relative azimuth angle (RAA,  $\phi$ ), surface pressure ( $P_s$ ), and BRDF function ( $R_{\text{BRDF}}$ )
- $I_0$ : 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  $\pi$  (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):

$$\begin{bmatrix} (I - I_0)_{R=0.05} & 1\\ (I - I_0)_{R=0.1} & 1 \end{bmatrix} \begin{bmatrix} S_b\\ T \end{bmatrix} = \begin{bmatrix} \left(\frac{I - I_0}{R}\right)_{R=0.05}\\ \left(\frac{I - I_0}{R}\right)_{R=0.1} \end{bmatrix}$$
(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 S9 shows the results for 85 sites (a combination of urban, power plant, and rural sites) with 189 geometry combinations for each site. Figure S9a shows only a 3% variation from a 1:1 line in the regression, and Fig. S9b 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 ± 4% (1 $\sigma$ ) decrease in the NO<sub>2</sub> 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 (~ 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  $1.011 \pm 0.001$  to  $1.0395 \pm 0.0005$ ), indicating that the average uncertainty does not vary significantly with season. This is explored in more detail in Sect. S4 of the main paper.



Figure S9: 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.

#### S4 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<sub>2</sub> 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:

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

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

Quantity	Perturbation	Reasoning
Surface reflectivity	$\pm 17\%$	Quadrature sum of 14% LER
		error and 10% from Schaaf et
		al. (2010)
Surface pressure	$\pm 1.5\%$	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 $\pm 0.05$	Acarreta et al. $(2004)$ with cor-
		relation 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<sub>2</sub> 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 S10 shows the summed uncertainty for the four seasons as well as the individual contributions to the uncertainty. In all seasons, the a priori NO<sub>2</sub> 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<sub>2</sub> 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 ~ 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 S10: (**a**–**d**) Total percent uncertainty in tropospheric NO<sub>2</sub> 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<sub>2</sub> columns to it and the terrain height ( $\% \Delta V_{NO_2}/\% \Delta p_{trop}$ ) are generally similar (~ 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 ~ 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<sub>x</sub> lifetime means that more pixels will have high levels of surface NO<sub>2</sub>, 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° × 3.75° 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.
- Laughner, J.L., Q. Zhu, and R.C. Cohen (2018). "The Berkeley High Resolution Tropospheric NO<sub>2</sub> Product". In: *Earth System Science Data*.
- 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.
- 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.