Supplement of
Validation of TROPOMI and WRF-Chem NO_2 across seasons using SWING+ and surface observations over Bucharest

Antoine Pasternak[1], Jean-François Müller[1], Catalina Poraicu[1], Alexis Merlaud[1], Frederik Tack[1], and Trissevgeni Stavrakou[1]

[1]Atmospheric Composition Department, Royal Belgian Institute for Space Aeronomy (BIRA-IASB), Avenue Circulaire 3, 1180 Brussels, Belgium

Correspondence: Antoine Pasternak (antoine.pasternak@aeronomie.be)

Supplement 1 Model evaluation against airborne column measurements

Figures S1 to S15 present temporal series and maps of SWING+ measurements and WRF-Chem modeled values for the 15 flights not shown in the main manuscript.

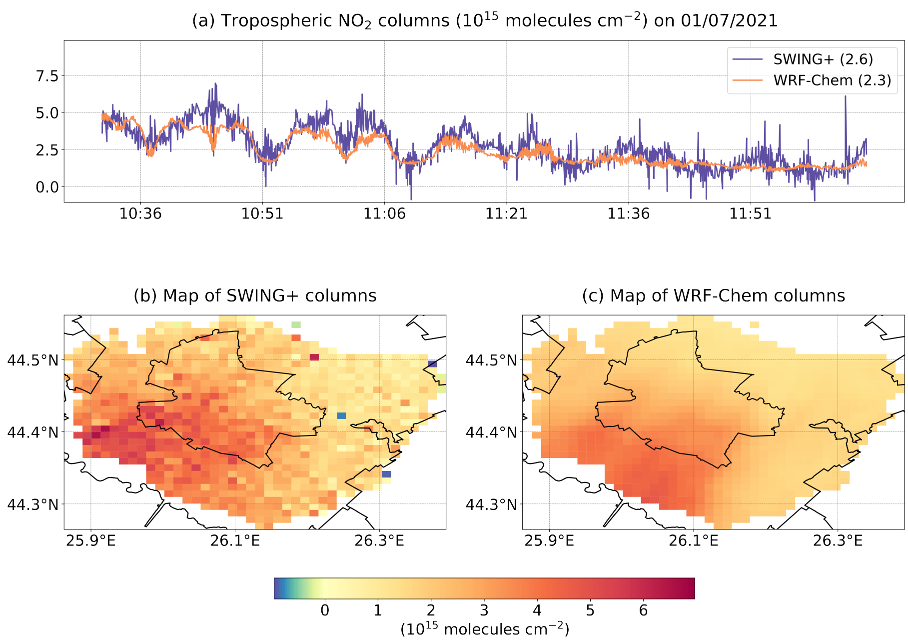

Figure S1. Tropospheric NO_2 columns on Thursday 1 July, 2021, presented as a temporal series of SWING+ and WRF-Chem values plotted against local time, with mean values in parentheses in (a), and corresponding maps in (b) and (c).

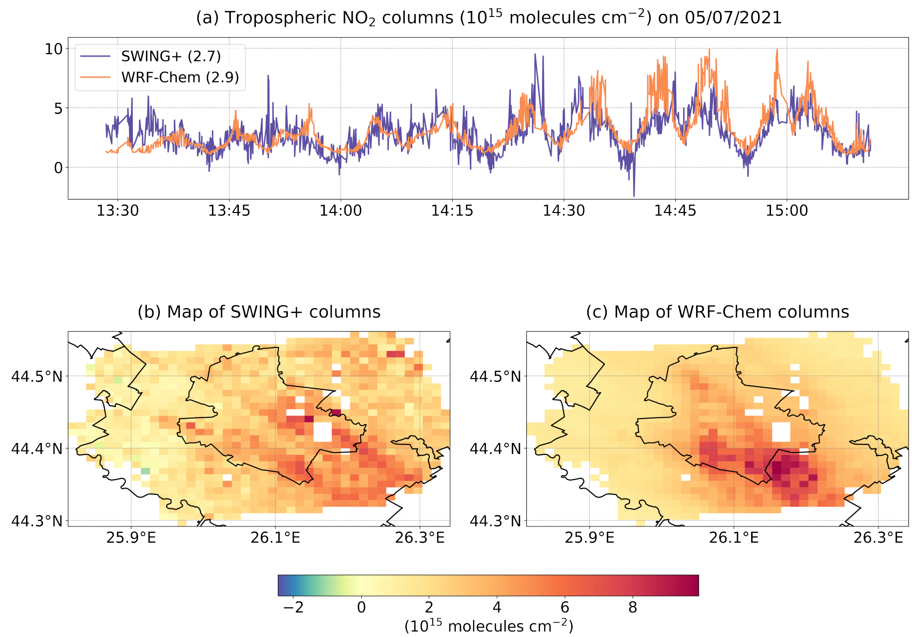

Figure S2. Same as Fig. S1, but for Monday 5 July, 2021.

Supplement 2 TROPOMI validation

In the main manuscript, a linear regression LR_1 is used to evaluate SWING+ measurements Ω_S against the corresponding WRF-Chem modeled values $\Omega_{W,S}$:

$$LR_1(\Omega_{W,S}) = \alpha_0 + \alpha_1 \Omega_{W,S}\,, \tag{S1}$$

where α_0 and α_1 are scalar coefficients determined separately for each selected flight. The non-parametric Theil–Sen (TS) method is presented and used in the main analysis. We now compare it with parametric ordinary least squares (OLS) and weighted least squares (WLS) regression methods (both implemented via the Python module statsmodels). The latter explicitly accounts for the random error $\sigma_{S,\,rand}$ associated with each SWING+ measurement. Two evaluation metrics, described in Table S1, are used for comparison. The coefficient of determination (R^2) is not included due to its subtle interpretation in the context of non-parametric methods. The results, presented by flight date and for all dates combined, are shown in Table S2.

As shown in the last row of Table S2, the MAD and RMSE obtained using TS are comparable to those from OLS and WLS. TS ranks better than WLS but slightly worse than OLS. Based on these two metrics alone, the differences among the methods are small and do not support a clear preference. However, TS is a robust estimator that handles outliers effectively, which is an important consideration given the potential for poor model performance for specific data subsets. For this reason, we favor TS over the parametric methods.

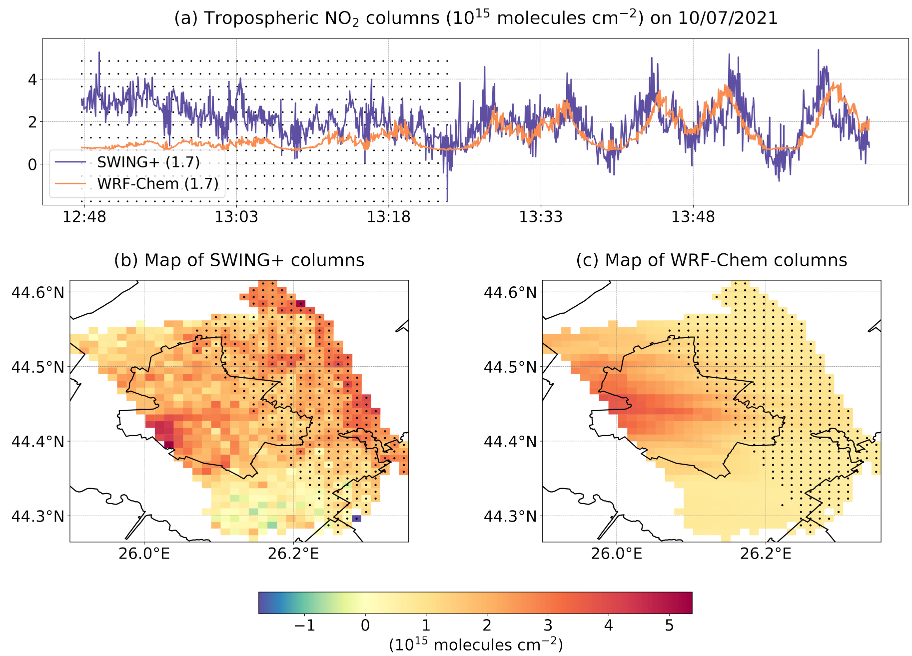

Figure S3. Same as Fig. S1, but for Saturday 10 July, 2021. Dotted values, acquired from 12:47 to 13:24 LT, are excluded from the analysis for reasons explained in the main text.

Table S1. Statistical metrics used to evaluate the linear regression method LR. The formulas are written for N observed values O_i and the corresponding modeled data M_i, with $i = 1, \ldots, N$. The weights depend on the individual random error of each measurement, $\sigma_{\text{rand},i}$, in the case of WLS and ODR regression methods. The definition of RMSE differs from that used in the main manuscript and is adapted for the evaluation of a regression method.

Metric	Formula
Mean absolute deviation	$\text{MAD} = \frac{1}{N}\sum_{i=1}^{N} w_i \lvert O_i - \text{LR}(M_i)\rvert$
Root mean square error	$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} w_i (O_i - \text{LR}(M_i))^2}$
Weight for OLS and TS	$w_i = 1$
Weight for WLS and ODR	$w_i = \dfrac{1/\sigma^2_{\text{rand},i}}{\sum_{j=1}^{N} 1/\sigma^2_{\text{rand},j}}$

After selecting the TS method for the linear regression LR_1, we proceed to evaluate LR_2, based on the TROPOMI measured values Ω_T and the corresponding modeled values $\Omega^{\text{bc}}_{\text{W,T}}$:

$$\text{LR}_2(\Omega^{\text{bc}}_{\text{W, T}}) = \beta_0 + \beta_1 \Omega^{\text{bc}}_{\text{W, T}}. \quad \text{(S2)}$$

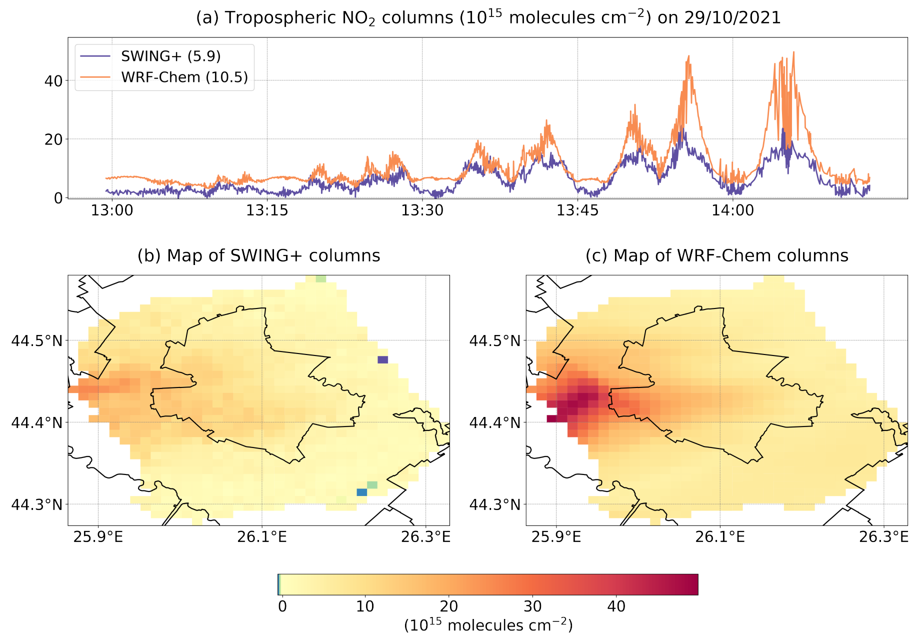

Figure S4. Same as Fig. S1, but for Friday 29 October, 2021.

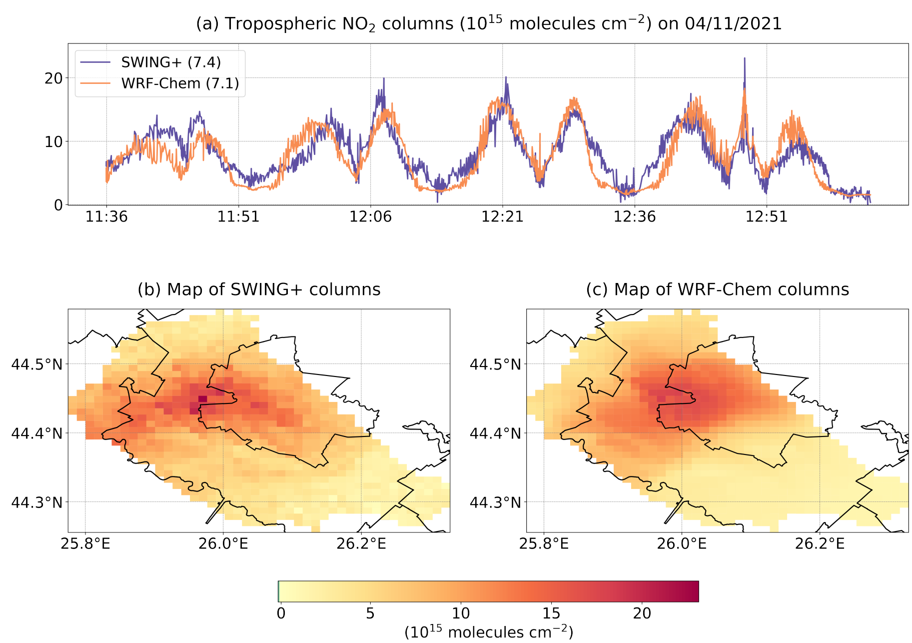

Figure S5. Same as Fig. S1, but for Thursday 4 November, 2021.

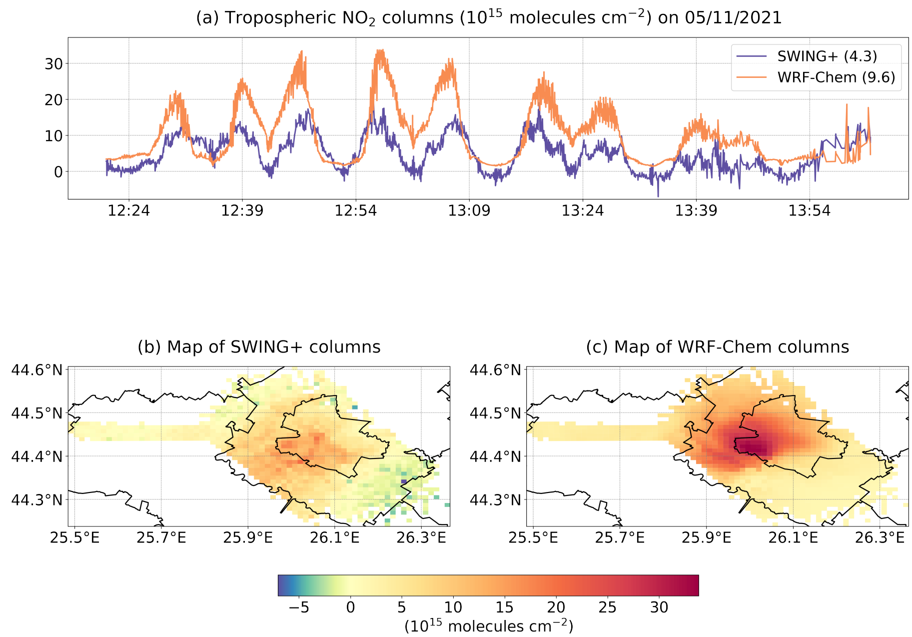

Figure S6. Same as Fig. S1, but for Friday 5 November, 2021.

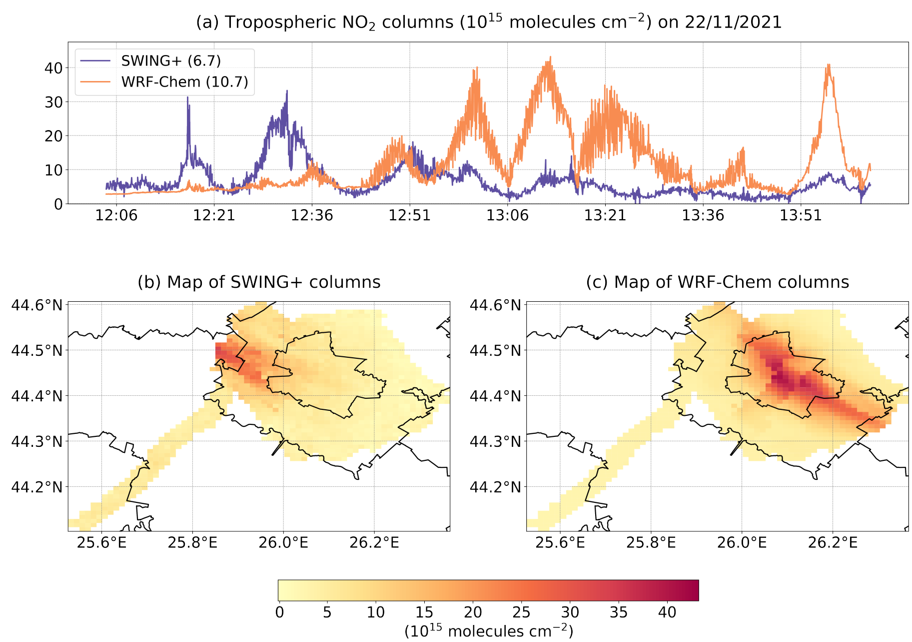

Figure S7. Same as Fig. S1, but for Monday 22 November, 2021.

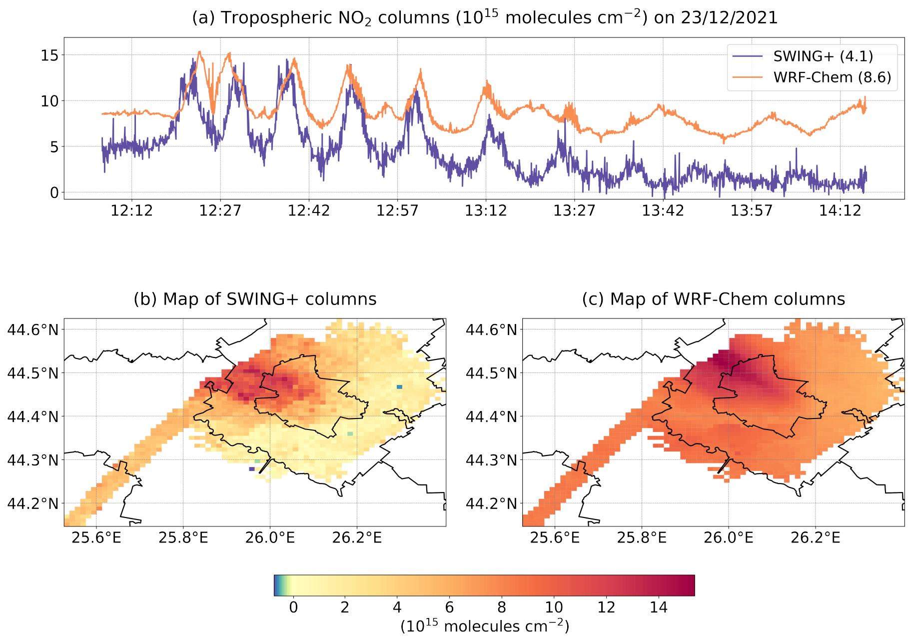

Figure S8. Same as Fig. S1, but for Thursday 23 December, 2021.

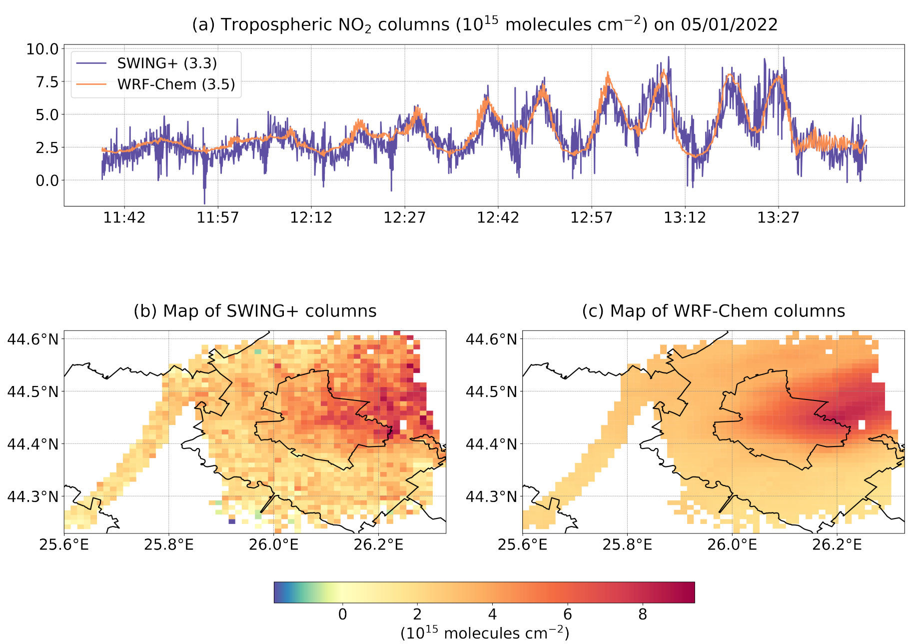

Figure S9. Same as Fig. S1, but for Wednesday 5 January, 2022.

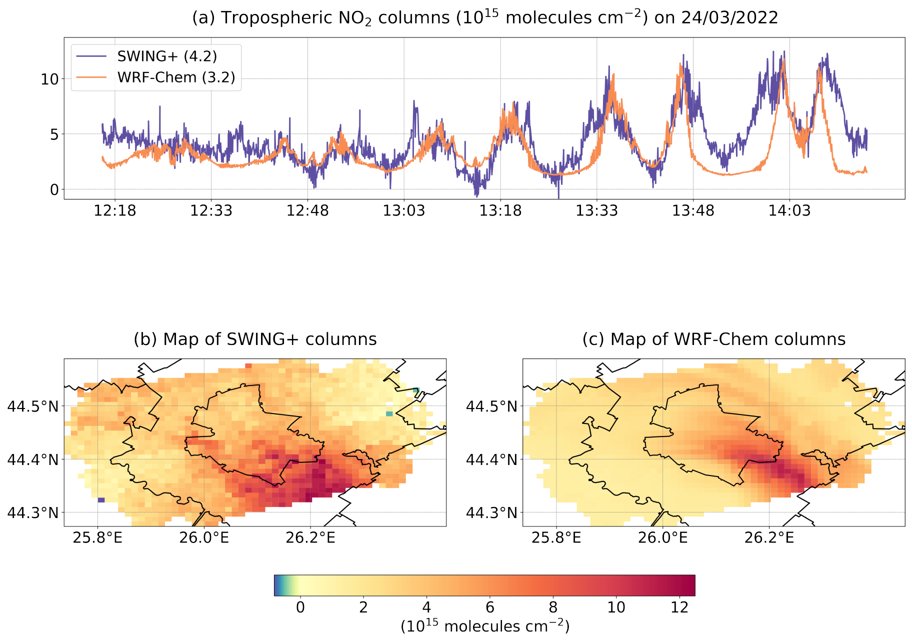

Figure S10. Same as Fig. S1, but for Thursday 24 March, 2022.

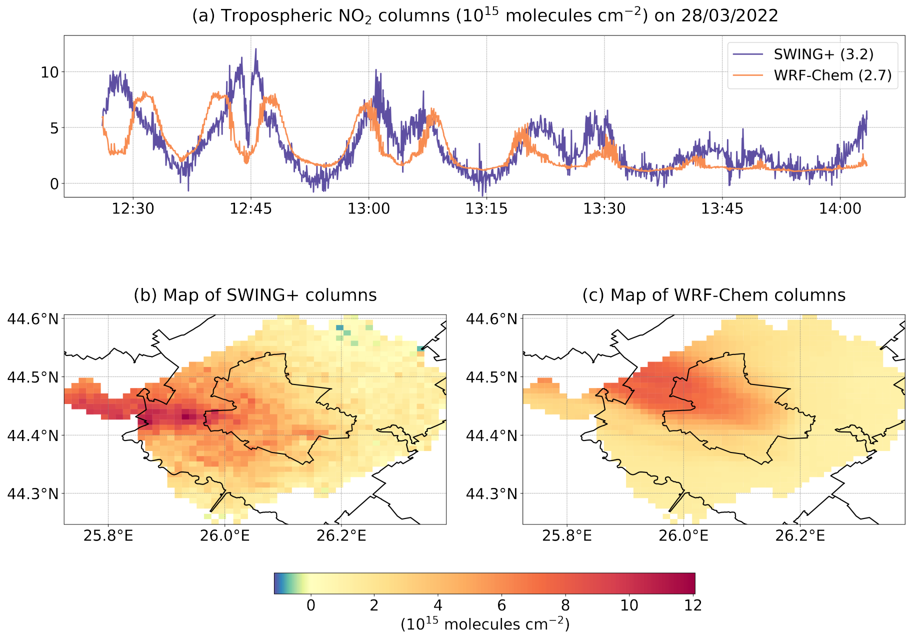

Figure S11. Same as Fig. S1, but for Monday 28 March, 2022.

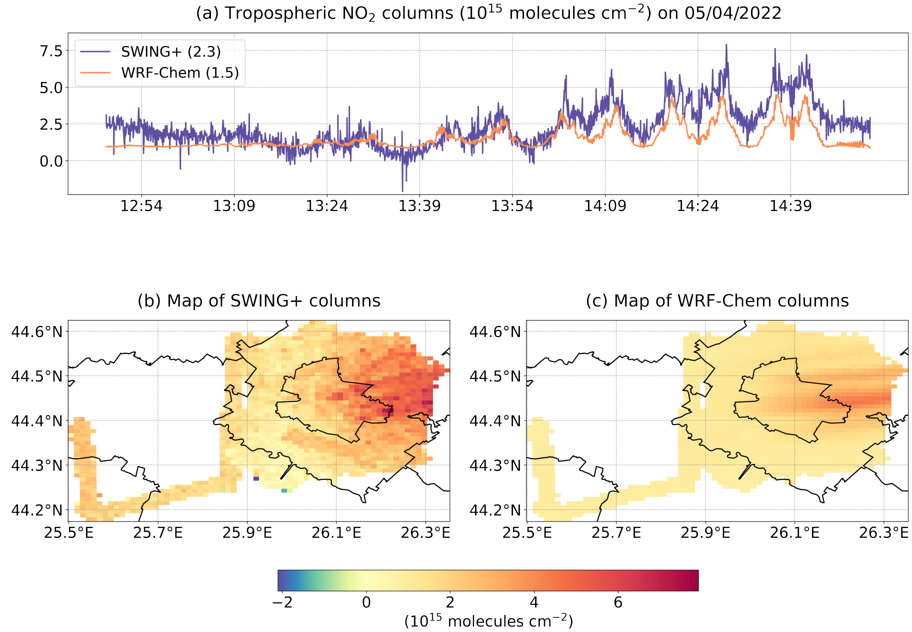

Figure S12. Same as Fig. S1, but for Tuesday 5 April, 2022.

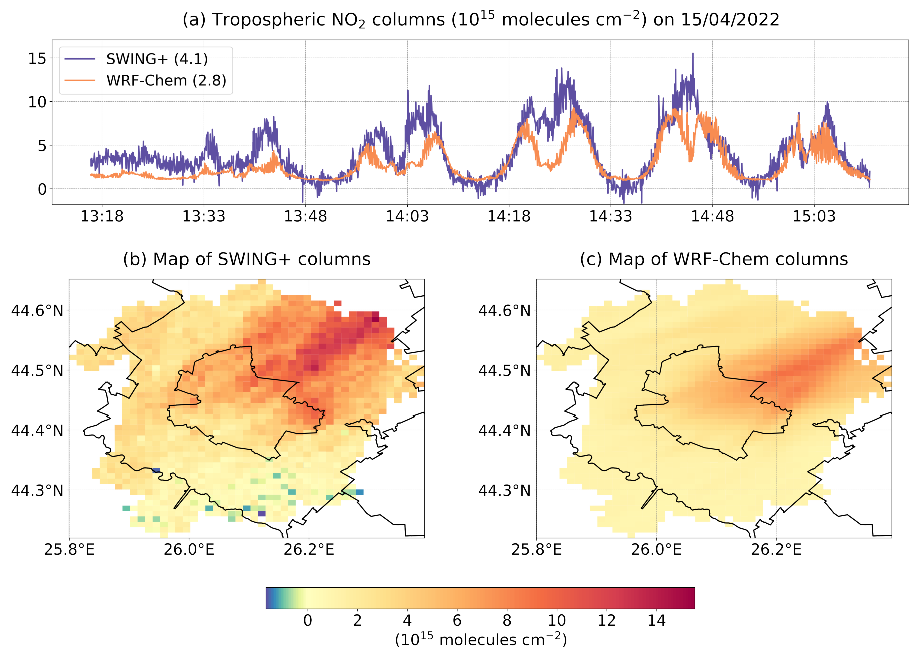

Figure S13. Same as Fig. S1, but for Friday 15 April, 2022.

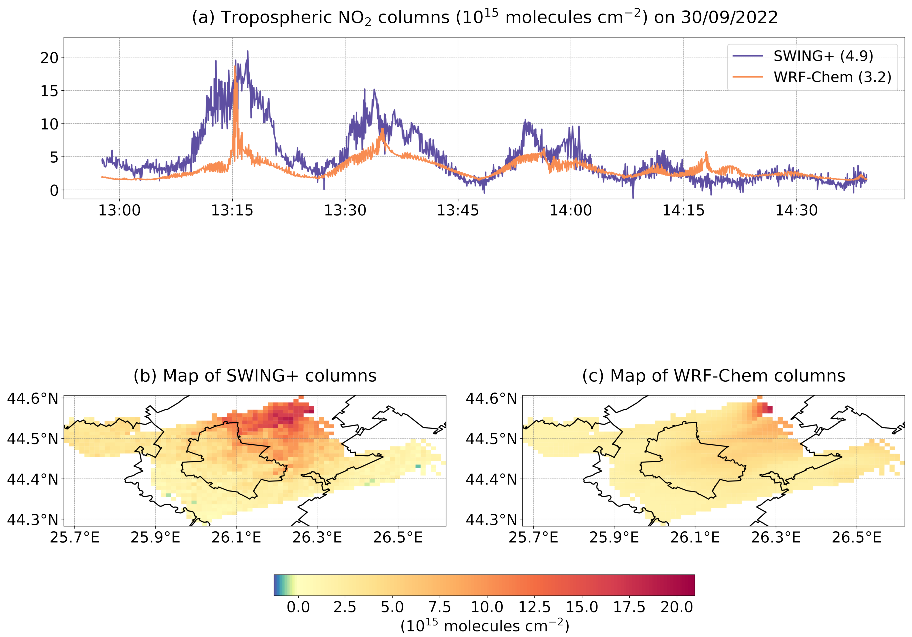

Figure S14. Same as Fig. S1, but for Friday 30 September, 2022.

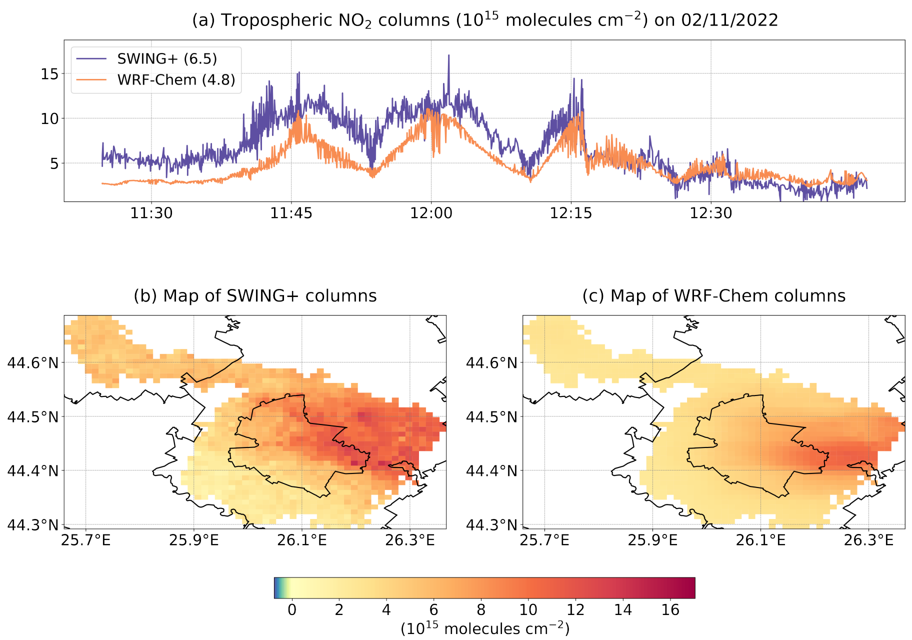

Figure S15. Same as Fig. S1, but for Wednesday 2 November, 2022.

In the main manuscript, we present and adopt the orthogonal distance regression with weights (ODR), rather than applying TS a second time. As shown in Fig. S16, both methods yield similar results in terms of slope, differing only slightly in their estimation of the intercept. We interpret this as evidence that outliers have little influence on the resulting regression line produced by the parametric ODR method. Moreover, ODR yields lower MAD and RMSE values, further supporting its selection. Note that the errors in the weights used for ODR, as defined in Table S1, are defined by the quadrature sum of uncertainties for each measurement: $\sigma_{\mathrm{T}}^2 + \beta_1^2 \sigma_{\mathrm{LR}_1}^2$, where σ_{T} is TROPOMI precision and σ_{LR_1} is the random error on $\Omega_{\mathrm{W,T}}^{\mathrm{bc}}$ resulting from the uncertainty of the linear regression method LR_1.

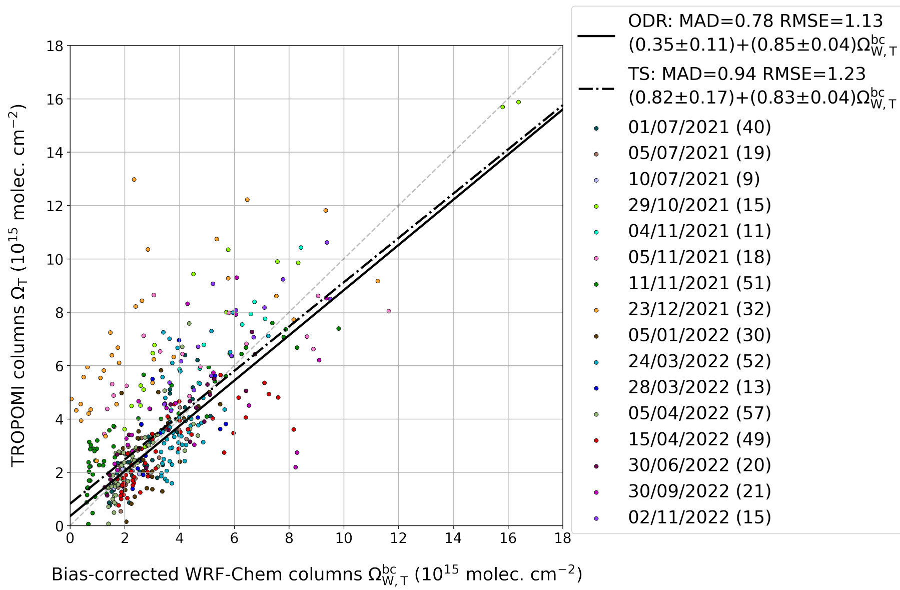

Figure S16. Comparison of ODR and TS linear regression methods applied to the scatter plot of 452 TROPOMI and bias-corrected WRF-Chem column values for all flight days (except 22/11/2021). MAD, RMSE, and the regression lines are expressed in units of 10^{15} molec. cm^{-2}. For each date, the number of columns is displayed in parentheses.

Table S2. Comparison of ordinary least squares (OLS), weighted least squares (WLS), and Theil–Sen (TS) linear regression methods. The slope α_1 is dimensionless, while all other quantities are expressed in 10^{15} molec. cm^{-2}.

Date	Sample	OLS				WLS				TS			
	size	α_0	α_1	MAD	RMSE	α_0	α_1	MAD	RMSE	α_0	α_1	MAD	RMSE
01/07/2021	40	0.06 ± 0.22	1.11 ± 0.09	0.43	0.53	-0.02 ± 0.20	1.16 ± 0.08	0.39	0.46	0.06 ± 0.31	1.18 ± 0.13	0.46	0.56
05/07/2021	19	0.80 ± 0.37	0.61 ± 0.10	0.57	0.70	0.63 ± 0.32	0.63 ± 0.09	0.46	0.60	0.90 ± 0.47	0.63 ± 0.11	0.61	0.72
10/07/2021	9	0.56 ± 0.58	0.76 ± 0.28	0.47	0.55	0.49 ± 0.50	0.80 ± 0.26	0.39	0.49	0.78 ± 1.19	0.70 ± 0.59	0.51	0.57
29/10/2021	15	0.24 ± 0.70	0.54 ± 0.05	1.17	1.45	0.32 ± 0.72	0.53 ± 0.05	1.19	1.49	-0.42 ± 1.77	0.64 ± 0.17	1.23	1.76
04/11/2021	11	2.08 ± 0.93	0.73 ± 0.11	1.16	1.27	2.51 ± 0.98	0.72 ± 0.11	1.07	1.26	1.91 ± 1.49	0.76 ± 0.12	1.17	1.28
05/11/2021	18	0.35 ± 1.10	0.41 ± 0.07	1.94	2.61	1.08 ± 1.10	0.39 ± 0.07	2.05	2.60	0.35 ± 1.53	0.44 ± 0.08	2.07	2.64
11/11/2021	51	-0.75 ± 0.15	0.92 ± 0.03	0.51	0.66	-0.73 ± 0.16	0.93 ± 0.03	0.54	0.66	-0.73 ± 0.27	0.95 ± 0.06	0.54	0.68
23/12/2021	32	-7.01 ± 1.16	1.32 ± 0.13	1.18	1.50	-8.18 ± 1.28	1.42 ± 0.14	1.33	1.57	-6.99 ± 1.40	1.32 ± 0.17	1.18	1.50
05/01/2022	30	0.13 ± 0.23	0.89 ± 0.06	0.31	0.39	0.05 ± 0.20	0.90 ± 0.05	0.27	0.35	0.13 ± 0.28	0.89 ± 0.08	0.31	0.39
24/03/2022	52	2.41 ± 0.51	0.63 ± 0.14	1.27	1.72	2.11 ± 0.47	0.67 ± 0.12	1.15	1.54	2.61 ± 0.45	0.53 ± 0.14	1.25	1.74
28/03/2022	13	0.87 ± 0.88	0.88 ± 0.26	1.26	1.45	1.17 ± 0.89	0.77 ± 0.26	1.21	1.41	1.07 ± 1.10	0.89 ± 0.41	1.30	1.47
05/04/2022	57	-0.02 ± 0.21	1.52 ± 0.12	0.54	0.63	-0.25 ± 0.21	1.55 ± 0.12	0.49	0.61	0.24 ± 0.33	1.44 ± 0.18	0.56	0.65
15/04/2022	49	0.57 ± 0.40	1.20 ± 0.12	1.22	1.55	0.48 ± 0.38	1.18 ± 0.11	1.05	1.35	0.40 ± 0.58	1.45 ± 0.22	1.31	1.67
30/06/2022	20	0.16 ± 0.36	1.18 ± 0.09	0.54	0.71	0.17 ± 0.33	1.20 ± 0.08	0.49	0.66	-0.09 ± 0.53	1.24 ± 0.10	0.52	0.72
30/09/2022	21	-2.43 ± 1.16	2.33 ± 0.33	1.57	2.04	-2.94 ± 1.09	2.35 ± 0.31	1.17	1.74	-1.46 ± 1.18	1.91 ± 0.42	1.45	2.16
02/11/2022	15	-0.75 ± 1.43	1.46 ± 0.27	1.44	1.61	-1.57 ± 1.19	1.51 ± 0.23	1.14	1.39	-0.91 ± 1.94	1.53 ± 0.40	1.50	1.64
All dates	452			0.91	1.29			0.95	1.37			0.93	1.34