

Article

Estimation of Surface NO₂ Volume Mixing Ratio in Four Metropolitan Cities in Korea Using Multiple Regression Models with OMI and AIRS Data

Daewon Kim ¹, Hanlim Lee ^{1,*}, Hyunkee Hong ¹, Wonei Choi ¹, Yun Gon Lee ² and Junsung Park ¹

¹ Division of Earth Environmental System Science Major of Spatial Information Engineering, Pukyong National University, Busan 608-737, Korea; k.daewon91@gmail.com (D.K.); brunhilt77@gmail.com (H.H.); cwylh3338@gmail.com (W.C.); junsung2ek@gmail.com (J.P.)

² Department of Atmospheric Sciences, Chungnam National University, Daejeon 34134, Korea; yungonlee@gmail.com

* Correspondence: hllee@pknu.ac.kr; Tel.: +82-51-629-6688

Academic Editors: Yang Liu, Jun Wang, Omar Torres and Richard Müller

Received: 25 April 2017; Accepted: 15 June 2017; Published: 18 June 2017

Abstract: Surface NO₂ volume mixing ratio (VMR) at a specific time (13:45 Local time) (NO₂ VMR_{ST}) and monthly mean surface NO₂ VMR (NO₂ VMR_M) are estimated for the first time using three regression models with Ozone Monitoring Instrument (OMI) data in four metropolitan cities in South Korea: Seoul, Gyeonggi, Daejeon, and Gwangju. Relationships between the surface NO₂ VMR obtained from in situ measurements (NO₂ VMR_{In-situ}) and tropospheric NO₂ vertical column density obtained from OMI from 2007 to 2013 were developed using regression models that also include boundary layer height (BLH) from Atmospheric Infrared Sounder (AIRS) and surface pressure, temperature, dew point, and wind speed and direction. The performance of the regression models is evaluated via comparison with the NO₂ VMR_{In-situ} for two validation years (2006 and 2014). Of the three regression models, a multiple regression model shows the best performance in estimating NO₂ VMR_{ST} and NO₂ VMR_M. In the validation period, the average correlation coefficient (R), slope, mean bias (MB), mean absolute error (MAE), root mean square error (RMSE), and percent difference between NO₂ VMR_{In-situ} and NO₂ VMR_{ST} estimated by the multiple regression model are 0.66, 0.41, −1.36 ppbv, 6.89 ppbv, 8.98 ppbv, and 31.50%, respectively, while the average corresponding values for the other two models are 0.75, 0.41, −1.40 ppbv, 3.59 ppbv, 4.72 ppbv, and 16.59%, respectively. All three models have similar performance for NO₂ VMR_M, with average R, slope, MB, MAE, RMSE, and percent difference between NO₂ VMR_{In-situ} and NO₂ VMR_M of 0.74, 0.49, −1.90 ppbv, 3.93 ppbv, 5.05 ppbv, and 18.76%, respectively.

Keywords: surface NO₂ volume mixing ratio; NO₂; OMI; multiple regression

1. Introduction

The main anthropogenic source of nitrogen dioxide (NO₂) is fossil fuel combustion, while natural sources of NO₂ include lightning, forest fires, and soil emissions [1,2]. In particular, since NO₂ is emitted in large quantities in automobile exhaust gas, NO₂ is often used as an indicator of traffic-related air pollution in urban areas [3]. In terms of its effect on human health, long-term NO₂ exposure can lead to respiratory depression and respiratory illness [4–8]. In addition, it is a precursor of aerosol nitrate, tropospheric ozone, and the hydroxyl radical (OH), the main atmospheric oxidant [9]. It is therefore important to measure NO₂ and various methods are used, with chemiluminescence, a well-known technique for measuring surface NO₂ volume mixing ratio (VMR) [10]. In situ measurements such as

the chemiluminescence method are, in general, more accurate than remote sensing techniques, but require a large number of in situ instruments to provide the spatial distribution of the NO₂ VMR at high resolution [11]. In recent years, NO₂ vertical column density (VCD) has been measured from satellites that can monitor NO₂ at global scale over a short time scale. Space-borne sensors that have observed global distributions of NO₂ are the Global Ozone Monitoring Experiment (GOME) aboard European Remote Sensing-2 (ERS-2) (1995–2003), Scanning Imaging Absorption Spectrometer for Atmospheric Chartography/Chemistry (SCIAMACHY) aboard Environmental Satellite (Envisat) (2002–2012), the Ozone Monitoring Instrument (OMI) aboard EOS-AURA (2004–present), and GOME-2 aboard the Meteorological Operational satellite (MetOp)-A (2007–present) and MetOp-B (2012–present) [12–17]. In many countries, air quality regulation requires surface NO₂ VMR so the NO₂ VCD obtained from satellites cannot be used directly. In recent years, studies have been conducted to investigate the feasibility of estimating the surface NO₂ VMR using the NO₂ VCD obtained from satellite measurements and, in particular, the correlation between the NO₂ VCD obtained from satellite measurements and the surface NO₂ VMR.

Ordóñez et al. [18] reported the correlation between tropospheric NO₂ VCD and the NO₂ VCD measured by GOME and ground based in situ devices in Milan. Kharol et al. [3] estimated the annual average ground-level NO₂ concentrations in North America using chemical transport model (GEOS-Chem) data and OMI NO₂ columns and also reported the annual trend of the estimated ground-level NO₂ concentrations. However, no studies have attempted to estimate the surface NO₂ VMR at higher temporal resolutions such as hourly and monthly using the NO₂ VCD measured by satellites.

In this present study, we estimate for the first time the surface NO₂ VMR at a specific time (13:45 Local time (LT)) (NO₂ VMRST) and the monthly mean surface NO₂ VMR (NO₂ VMRM) using two linear regression models and a multiple regression model with the tropospheric NO₂ VCD obtained from OMI (Trop NO₂ VCDOMI) in five metropolitan cities. In addition, the performance of each regression method is evaluated by comparing the estimated surface NO₂ VMRs with those obtained from in situ measurement (NO₂ VMR_{in-situ}).

2. Study Area and Period

A large amount of anthropogenic NO_x is emitted in Northeast Asia including China, Korea and Japan [19]. Especially, the annual mean NO₂ tended to increase in Seoul from 1995 to 2009 [20]. The study areas were selected where the surface NO₂ VMR is continuously measured in Korean metropolitan cities (Figure 1). Metropolitan cities such as Busan and Incheon where the OMI pixel covers both sea and land are excluded since there are no surface NO₂ data available over the sea. Therefore, the selected areas are Seoul, Gyeonggi, Daejeon, and Gwangju. Seoul is covered by four OMI pixels and is divided into eastern and western areas (West Seoul and East Seoul). The study period is the nine years from 2006 to 2014. This is split into a seven-year training period (2007–2013) to determine the coefficients of the regression models used in this study, and two years of validation (2006 and 2014) when the surface NO₂ VMRs estimated from the resulting three regression models are evaluated by comparison with the in situ data. The three regression models used in this study are described in detail in Section 3.

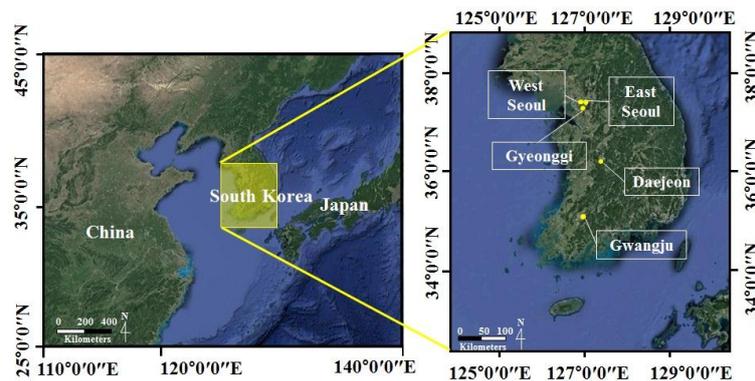


Figure 1. Study areas in South Korea.

2.1. Data

The data used in this study are Trop NO_2 VCD_{OMI} and Atmospheric Infrared Sounder (AIRS) boundary layer height (BLH_{AIRS}), atmospheric temperature ($\text{Temp}_{\text{AIRS}}$) and pressure ($\text{Press}_{\text{AIRS}}$), together with in situ measurements of NO_2 $\text{VMR}_{\text{In-situ}}$, surface temperature ($\text{Temp}_{\text{In-situ}}$), surface pressure ($\text{Press}_{\text{In-situ}}$), surface dew point ($\text{Dewpoint}_{\text{In-situ}}$), surface wind speed ($\text{WS}_{\text{In-situ}}$), and surface wind direction ($\text{WD}_{\text{In-situ}}$) (see Table 1).

Table 1. Satellite and in situ data used in this study.

	Data		Time (LT)
Satellite	Trop NO_2 VCD	OMI Level3 NO_2 Daily data (OMNO2d)	13:45
	BLH, Temperature, Pressure	AIRS/Aqua L3 Daily Support Product (AIRS + AMSU) V006 (AIRX3SPD)	13:30
In situ	Surface NO_2 VMR	Air Korea	
	Surface Temperature, Surface Pressure, Surface Dew point, Surface Wind Speed, Surface Wind Data	AWS (Automatic Weather System)	13:00 and 14:00

2.1.1. Ozone Monitoring Instrument (OMI) Data

The Trop NO_2 VCD_{OMI} data were obtained from OMI Level3 NO_2 Daily Data (OMNO2d) provided by the NASA Goddard Earth Sciences Data and Information Services Center (<http://disc.sci.gsfc.nasa.gov/Aura/data-holdings/OMI>) [17,21,22]. OMI is a nadir-viewing UV-visible (270–500 nm) spectrometer aboard the Aura platform launched in July 2004 [23]. Aura is a polar orbiting satellite with an overpass time of 13:45 LT. The spectral resolution of the OMI is about 0.5 nm and the spatial resolution is 13×24 km at nadir. Cloud-screened NO_2 data (Level-3 OMI NO_2 Cloud-Screened Total and Tropospheric Column NO_2 (V003)) are used in the present study (Cloud Fraction <30%).

2.1.2. Atmospheric Infrared Sounder (AIRS) Data

The BLH_{AIRS} , $\text{Temp}_{\text{AIRS}}$, and $\text{Press}_{\text{AIRS}}$ used in this study were obtained from the AIRS/Aqua L3 Daily Support Product (AIRS + AMSU) 1 degree \times 1 degree V006 (AIRX3SPD.00) from NASA Goddard Earth Sciences Data and Information Services Center (http://disc.sci.gsfc.nasa.gov/ui/datasets/AIRX3SPD_V006/summary?keywords=%22AIRS%22) [24–26]. The AIRS/Advanced Microwave Sounding Unit (AMSU) is a sounding suite launched in May 2002 aboard Aqua [26,27]. Aqua is a polar orbiting satellite with an overpass time of 13:30 LT and a horizontal spatial resolution of 40 km at nadir.

2.1.3. In Situ NO₂ Data

The NO₂ VMR_{In-situ} data were obtained from Air Korea (http://www.airkorea.or.kr/last_amb_hour_data). Since NO₂ VMR_{In-situ} is available hourly, the average of the values at 13:00 and 14:00 LT is used to be closer to the OMI overpass time. In a previous study [18], the in situ measurements were grouped into five different NO₂ levels: clean, slightly polluted, averagely polluted, polluted, and heavily polluted. Many stations are located close to roads and are exposed to emissions. In addition, the in situ NO₂ data from stations within GOME pixels (320 × 40 km) were averaged, since in situ measurements are only representative of a small fraction of the satellite ground scene. In the present study, the NO₂ VMR_{In-situ} obtained from in situ measurements located close to streets were excluded in this study. We used the average of three or more NO₂ VMR_{In-situ} from stations located at least 2 km from each other.

2.1.4. In Situ Meteorological Data

The Temp_{In-situ}, Press_{In-situ}, Dewpoint_{In-situ}, WS_{In-situ}, and WD_{In-situ} used in this study are Automatic Weather System (AWS) data provided by the Korea Meteorological Administration (<http://sts.kma.go.kr/jsp/home/contents/statistics/newStatisticsSearch.do?menu=SFC&MNU=MNU>). Since meteorological data are available hourly, the average of the data at 13:00 LT and 14:00 LT is used. The surface wind data, especially wind direction can be impacted by local topography and interferences.

3. Methodology

In this study, NO₂ VMR_{ST} and NO₂ VMR_M were estimated using three regression models with Trop NO₂ VCD_{OMI}. Table 2 summarizes the three models.

Table 2. Regression models used for surface NO₂ VMR estimation in this study.

	Model	Equation
M1	13:45 LT and Monthly	$\text{NO}_2 \text{ VMR}_{in\ situ} = a \text{Trop NO}_2 \text{ VCD}_{OMI}^{(a)} + b$
M2	13:45 LT and Monthly	$\text{NO}_2 \text{ VMR}_{in\ situ} = a \text{BLH NO}_2 \text{ VMR}_{OMI}^{(b)} + b$
M3	13:45 LT	Section 3, Multiple regression Equation (1)
M4	Monthly	

Notes: (a) NO₂ tropospheric vertical column density obtained from OMI; and (b) BLH NO₂ VMR_{OMI} = $\frac{\text{Trop NO}_2 \text{ VCD}_{OMI} \text{ Gas constant } R \text{ Temp}_{AIRS} \times 10^{13}}{\text{Avogadro constant } NA \text{ BLH}_{AIRS} \text{ Press}_{AIRS}}$, where the AIRS pressure and temperature are boundary layer mean values, Gas constant R = 8.314472 m³ pa K⁻¹ mol⁻¹ and Avogadro constant NA = 6.022 × 10²³ mol⁻¹.

3.1. M1

M1 is the linear regression equation where Trop NO₂ VCD_{OMI} is used as the independent variable. Figure 2 shows the linear regression between Trop NO₂ VCD_{OMI} and NO₂ VMR_{In-situ} at 13:45 LT during the training period, with R² (coefficient of determination), slope, and intercept of 0.47, 0.80 and 11.47, respectively. Figure 3 shows the linear regression between monthly mean Trop NO₂ VCD_{OMI} and monthly mean NO₂ VMR_{In-situ} during the training period, with R², slope, and intercept of 0.62, 0.77, and 10.95, respectively. The final form of the M1 equation for estimating NO₂ VMR_{ST} is shown in Table 3, and that for estimating NO₂ VMR_M in Table 4.

Tables 3 and 4 show the equations M1, M2, M3, and M4 with the regression coefficients determined from the training period.

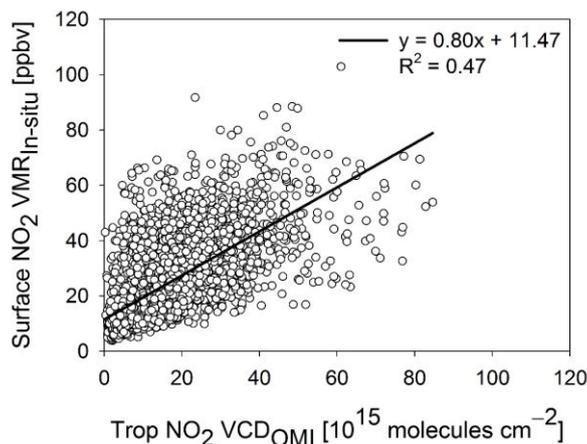


Figure 2. Scatter plot between Trop NO₂ VCD_{OMI} at 13.45 LT and NO₂ VMR_{In-situ} to determine the regression coefficient for M1 for the training period 2007–2013.

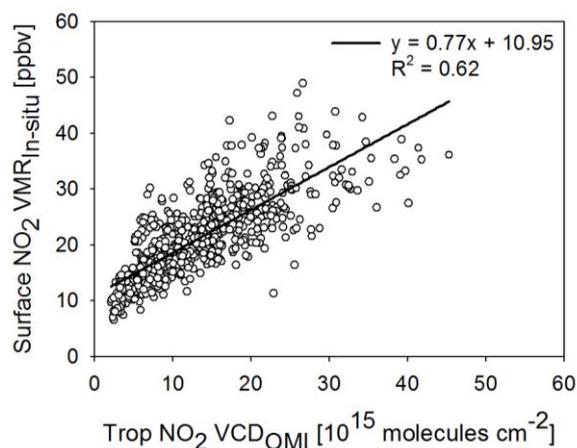


Figure 3. As Figure 2 but for the monthly mean values.

Table 3. Final form of the regression models used for estimating surface NO₂ VMR at a specific time and R² obtained from the regression between NO₂ VMR_{In-situ} and the corresponding independent variable for the training period.

	Equation	R ²
13:45 LT	M1 $NO_2\ VMR_{ST} = 1.71 \times Trop\ NO_2\ VCD_{OMI} - 0.68$	0.47
	M2 $NO_2\ VMR_{ST} = 4.19 \times BLH\ NO_2\ VMR_{OMI} + 1.57$	0.38
	M3 $NO_2\ VMR_{ST} = 0.000602 \times Trop\ NO_2\ VCD_{OMI} - 0.000107 \times Temp_{In-situ} - 0.000083 \times Dewpoint_{In-situ} + 0.000061 \times Press_{In-situ} - 0.000002 \times BLH_{AIRS} - 0.002435 \times WS_{In-situ} + 0.001190 \times WD_{In-situ} - 0.039996$	0.47

Table 4. As Table 3 but for monthly mean surface NO₂ VMR.

	Equation	R ²
Monthly mean	M1 $NO_2\ VMR_M = 1.23 \times Trop\ NO_2\ VCD_{OMI} + 4.74$	0.62
	M2 $NO_2\ VMR_M = 2.92 \times BLH\ NO_2\ VMR_{OMI} + 6.74$	0.59
	M4 $NO_2\ VMR_M = 0.657241 \times Trop\ NO_2\ VCD_{OMI} - 0.137334 \times Dewpoint_{In-situ} - 0.136096 \times Press_{In-situ} - 0.004331 \times BLH_{AIRS} - 0.770356 \times WS_{In-situ} + 2.370956 \times WD_{(west)_{In-situ}} + 157.361668$	0.63

3.2. M2

There might exist a minor fraction of the tropospheric NO_2 column in upper troposphere particularly because of lightning. However, the NO_2 amount in upper troposphere could be considered negligible in metropolitan cities, where a significant amount of NO_x is emitted. Therefore, assuming Trop NO_2 VCD_{OMI} is mostly present within the PBL, the relationship between Trop NO_2 VCD_{OMI} and the surface NO_2 VMR may change as the PBL varies. However, a minor fraction of the tropospheric NO_2 column can also be in the upper tropospheric, particularly because of lightning. This NO_2 fraction in upper tropospheric might cause either small or negligible reduction in correlations of the OMI NO_2 VCD between and surface NO_2 VMR as the upper part of the troposphere (free troposphere) contribution is assumed to be negligible [28]. To reflect the BLH in the regression equation, Trop NO_2 VCD_{OMI} is first divided by BLH_{AIRS} to calculate the NO_2 concentration in the PBL and then converted to the NO_2 mixing ratio in the PBL ($\text{BLH NO}_2 \text{VMR}_{\text{OMI}}$) using $\text{Temp}_{\text{AIRS}}$ and $\text{Press}_{\text{AIRS}}$ [29] as shown Table 2. Only a single OMI pixel contained completely within an AIRS pixel was used. Figure 4 shows the linear regression between $\text{BLH NO}_2 \text{VMR}_{\text{OMI}}$ and $\text{NO}_2 \text{VMR}_{\text{In-situ}}$ at 13:45 LT during the training period. Here R^2 , slope and intercept are 0.38, 1.58, and 14.30, respectively. Figure 5 shows the corresponding linear regression for the monthly mean data, with R^2 , slope and intercept of 0.59, 1.71, and 12.75, respectively. The final form of equation M2 to estimate $\text{NO}_2 \text{VMR}_{\text{ST}}$ is shown in Table 3, and for the monthly values in Table 4.

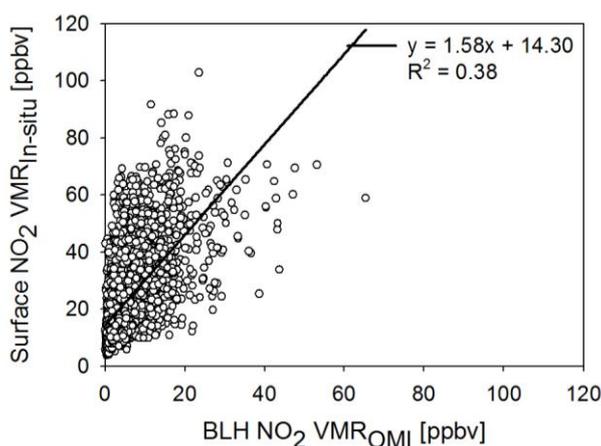


Figure 4. Scatter plot between $\text{BLH NO}_2 \text{VMR}_{\text{OMI}}$ at a specific time (13:45 LT) and $\text{NO}_2 \text{VMR}_{\text{In-situ}}$ to determine the regression coefficient for M1 for the training period 2007–2013.

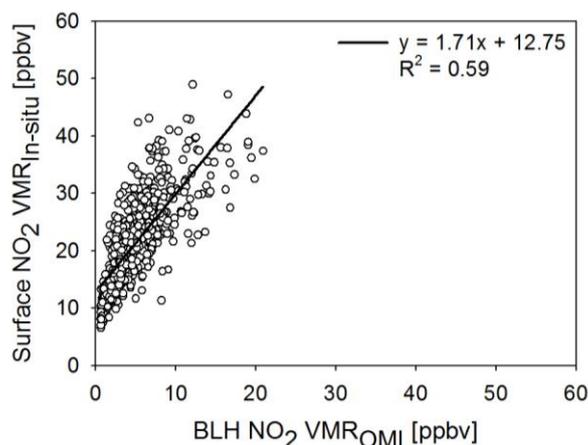


Figure 5. As Figure 4 but for the monthly mean values.

3.3. M3 and M4

M3 and M4 are multiple regression equations for estimating NO₂ VMR_{ST} and NO₂ VMR_M. Multiple regression equations consist of a dependent variable, independent variables, and their regression coefficients. In addition to Trop NO₂ VCD_{OMI} and BLH_{AIRS}, meteorological factors (surface temperature, dew point, atmospheric pressure, wind direction, and wind speed) are used as candidate independent variables for the multiple regression equation in the present study. In a previous study [30], these meteorological factors were also used as candidate independent variables to estimate surface SO₂ concentration in Shanghai, China. Temperature, pressure, boundary layer height, wind speed, and wind direction were selected as the candidates for independent variables since they are known to either directly or indirectly affect the spatial mixing of NO₂ molecules in boundary layer. Furthermore, temperature and dewpoint were selected as candidates for independent variables as they affect the boundary layer height [31].

The multiple regression equation can be defined by the following equations:

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \varepsilon \quad (1)$$

where \hat{y} and β_0 are the dependent variable (NO₂ VMR_{In-situ}) and regression coefficient, respectively; x_1, x_2, \dots, x_n are the candidate independent variables (Trop NO₂ VCD_{OMI}, Dewpoint_{In-situ}, Press_{In-situ}, Temp_{In-situ}, BLH_{AIRS}, WS_{In-situ}, and WD_{In-situ}); $\beta_1, \beta_2, \dots, \beta_n$ are the regression coefficients of the independent variables; and ε is the difference between observations (NO₂ VMR_{In-situ}) and estimated values (NO₂ VMR_{estimate}). The regression coefficients can be estimated by least square fitting:

$$\sum_{j=1}^m \varepsilon_j^2 = \sum_{j=1}^m (y_j - \hat{y}_j)^2 \quad (2)$$

where y_j is the observed value with m data points. By minimizing the sum of ε^2 , regression coefficients can be derived. These least square fitting techniques are based on the following assumptions: the linear relationship, a normal distribution and equal variance in the residuals. The least squares regression is sensitive to the presence of some points that are excessively large or small values in the training data [32]. To determine the independent variables (x_n) and regression coefficients (β_n) included in the final form of equations M3 and M4, we considered the variation inflation factor (VIF) and p -value to ensure their statistical significance. First, we examined the VIF that explains the multicollinearity of a candidate independent variable with regard to other candidate independent variables. The VIF of the j -th independent variable is expressed as:

$$\text{VIF}(x_j) = \frac{1}{1 - R_j^2} \quad (3)$$

where R_j^2 is the coefficient of determination for the regression of x_j against another independent variable (a regression that does not involve the dependent variable j). The VIF indicates how much x_j is correlated with the other candidate variables. A candidate independent variable with a very high VIF can be considered redundant and should be removed from the multiple regression equations. Candidate independent variables that do not satisfy the criterion $\text{VIF} < 10$ [33], were excluded from the independent variables. The p -value was also used to select independent variables. The highest still statistically significant p -level was shown by Sellke et al. [34] to be 5%. Among the independent variables that satisfy the VIF criterion, those that also satisfy p -value < 0.05 are selected as final independent variables in the multiple regression equations. The independent variables selected for equations M3 and M4 are shown in Table 5. The final form of equation M3 to estimate NO₂ VMR_{ST} is shown Table 3, and that for NO₂ VMR_M in Table 4.

Table 5. Final independent variables included in multiple regression equations (M3 and M4).

	Final Selected Independent Variables	<i>p</i> -Value	VIF
M3	Trop NO ₂ VCD _{OMI}	0	1.26
	Temp _{In-situ}	0.000032	7.02
	Dewpoint _{In-situ}	0.000306	7.16
	Press _{In-situ}	0.009981	3.14
	BLH _{AIRS}	1.73×10^{-15}	1.12
	WS _{In-situ}	3.86×10^{-133}	1.33
	WD _{In-situ}	1.7493×10^{-38}	1.07
M4	Trop NO ₂ VCD _{OMI}	2.4832×10^{-89}	1.64
	Dewpoint _{In-situ}	0.000421	6.47
	Press _{In-situ}	0.034582	6.65
	BLH _{AIRS}	0.000834	2.32
	WS _{In-situ}	3.86×10^{-133}	1.59
	WD _{In-situ}	1.699×10^{-7}	1.25

4. Results

4.1. Daily Estimates

Figure 6 shows the day-to-day variations of NO₂ VMR_{In-situ} and NO₂ VMR_{ST} estimated at 13:45 LT in West Seoul and East Seoul using M1, M2 and M3 in Table 3 for 2006 and 2014. A slightly larger difference in magnitude is found between NO₂ VMR_{In-situ} and NO₂ VMR_{ST} obtained with M3 compared to those between NO₂ VMR_{In-situ} and NO₂ VMR_{ST} obtained with M1 and M2. However, NO₂ obtained from M3 showed moderate agreement with NO₂ VMR_{In-situ} in the form of the day-to-day variation. Results for Daejeon, Gwangju, and Gyeonggi are included in the Supplementary Materials.

Figure 7 shows the *R*, slope, mean bias (MB), mean absolute error (MAE), root mean square error (RMSE) and percent difference between NO₂ VMR_{ST} and NO₂ VMR_{In-situ} for the validation period (2006 and 2014). The *R* obtained with M1 ranges from 0.49 to 0.71, showing better agreement than that with M2 ($0.47 < R < 0.65$). M3 showed the best correlation with NO₂ VMR_{In-situ} ($0.67 < R < 0.90$). The slopes from both M1 and M2 are close to one in East Seoul, whereas they are lower in the other cities. The MB from M1, M2, and M3 ranges from -7.74 to 5.80 ppbv. In all study areas, the MAE (5.79 ppbv $< MAE < 8.25$ ppbv) of M3 is lower than those (6.58 ppbv $< MAE < 11.41$ ppbv) of M1 and M2, which means that NO₂ VMR_{ST} estimated from M3 show moderate agreement with NO₂ VMR_{In-situ} in terms of magnitude. The RMSE from M3 is found to be lower than those from M1 and M2. The NO₂ VMR_{ST} from M3 have the lowest RMSE in all study areas (7.21 ppbv $< RMSE < 11.37$ ppbv). In addition, percent differences estimated from M3 and NO₂ VMR_{In-situ} are lower in all study areas than from M1 and M2. In estimating NO₂ VMR_{ST}, M3, which is a multiple regression method with various independent variables as inputs, generally showed good statistical performance except for MB.

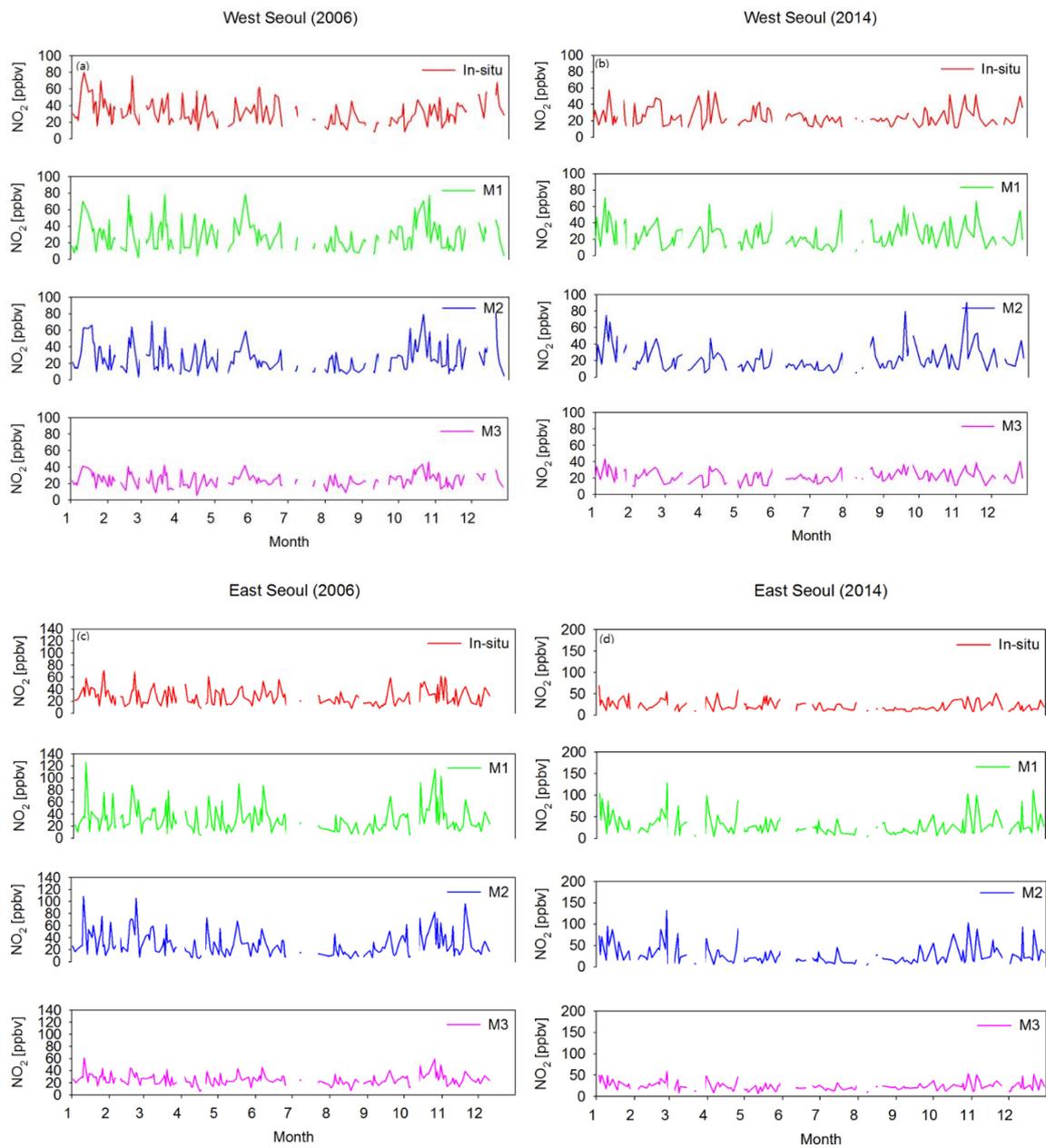


Figure 6. Time series of $\text{NO}_2 \text{ VMR}_{\text{In-situ}}$ and $\text{NO}_2 \text{ VMR}_{\text{ST}}$ at 13:45 LT estimated by M1, M2 and M3 in East Seoul and West Seoul for: 2006 (a,c); and 2014 (b,d).

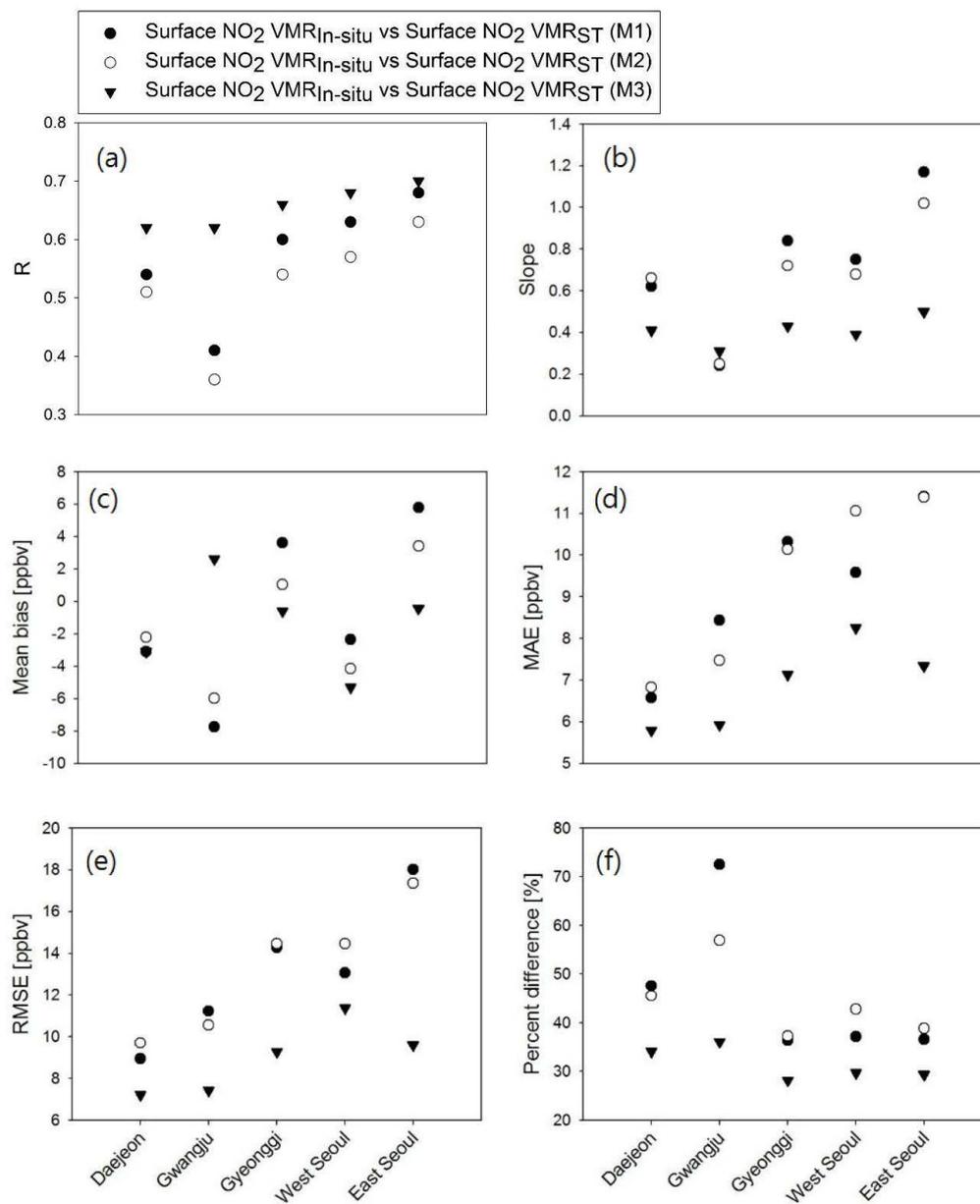


Figure 7. (a) R ; (b) slope; (c) MB; (d) MAE; (e) RMSE; and (f) percent difference between NO₂ VMR_{ST} against NO₂ VMR_{In-situ} in 2006 and 2014 for M1, M2, and M3.

4.2. Monthly Estimates

Figure 8 shows the temporal variation of monthly mean NO₂ VMR_{In-situ} and NO₂ VMR_M estimated using M1, M2 and M4 of Table 4 in West Seoul and East Seoul using monthly mean independent variables during the validation period (see the detailed input data in Section 2.1). Figure 8 shows good agreement in terms of the temporal pattern between the estimated NO₂ VMR_M and monthly mean NO₂ VMR_{In-situ}. However, we found a large difference between NO₂ VMR_{In-situ} and NO₂ VMR_M in periods when there was a jump in NO₂ VMR_{In-situ} between successive months. For example, no models calculated NO₂ VMR_M that were similar to NO₂ VMR_{In-situ} in December 2006, which is very different from that in November 2006. NO₂ VMR_{In-situ} (NO₂ VMR_M from M1, M2, and M4) in November and December in 2006 are 19.32 ppbv (15.94, 17.96, and 17.62 ppbv) and 30.30 ppbv (15.94, 17.96, and 17.62 ppbv) in Daejeon, 15.26 ppbv (12.29, 13.87, and 18.09 ppbv) and 32.55 ppbv (12.73, 14.57, and 18.46 ppbv) in Gwangju, 29.31 ppbv (25.86, 25.97, and 22.35 ppbv) and

40.64 ppbv (29.91, 29.15, and 26.85 ppbv) in Gyeonggi, and 31.25 ppbv (22.80, 24.55, and 23.64 ppbv) and 45.93 ppbv (28.65, 28.92, and 26.49 ppbv) in West Seoul. Especially in West Seoul, there are several periods when NO_2 $\text{VMR}_{\text{In-situ}}$ changes rapidly compared with the previous month. The NO_2 VMR_M obtained from the three models at these times are in poor agreement with the pattern of monthly NO_2 $\text{VMR}_{\text{In-situ}}$. As described in Section 2, despite the use of NO_2 $\text{VMR}_{\text{In-situ}}$ located away from the streets, the in situ measurement sites in West Seoul are located closer to the streets than the in situ measurement sites in Daejeon and Gwangju. This may explain why there are more periods when NO_2 $\text{VMR}_{\text{In-situ}}$ changes rapidly from one month to the next. It is difficult to estimate the rapid change of NO_2 VMR near the NO_2 source using regression models that reflect the relationship between the in situ measurements and the OMI sensor covering both source and non-source areas in a single pixel.

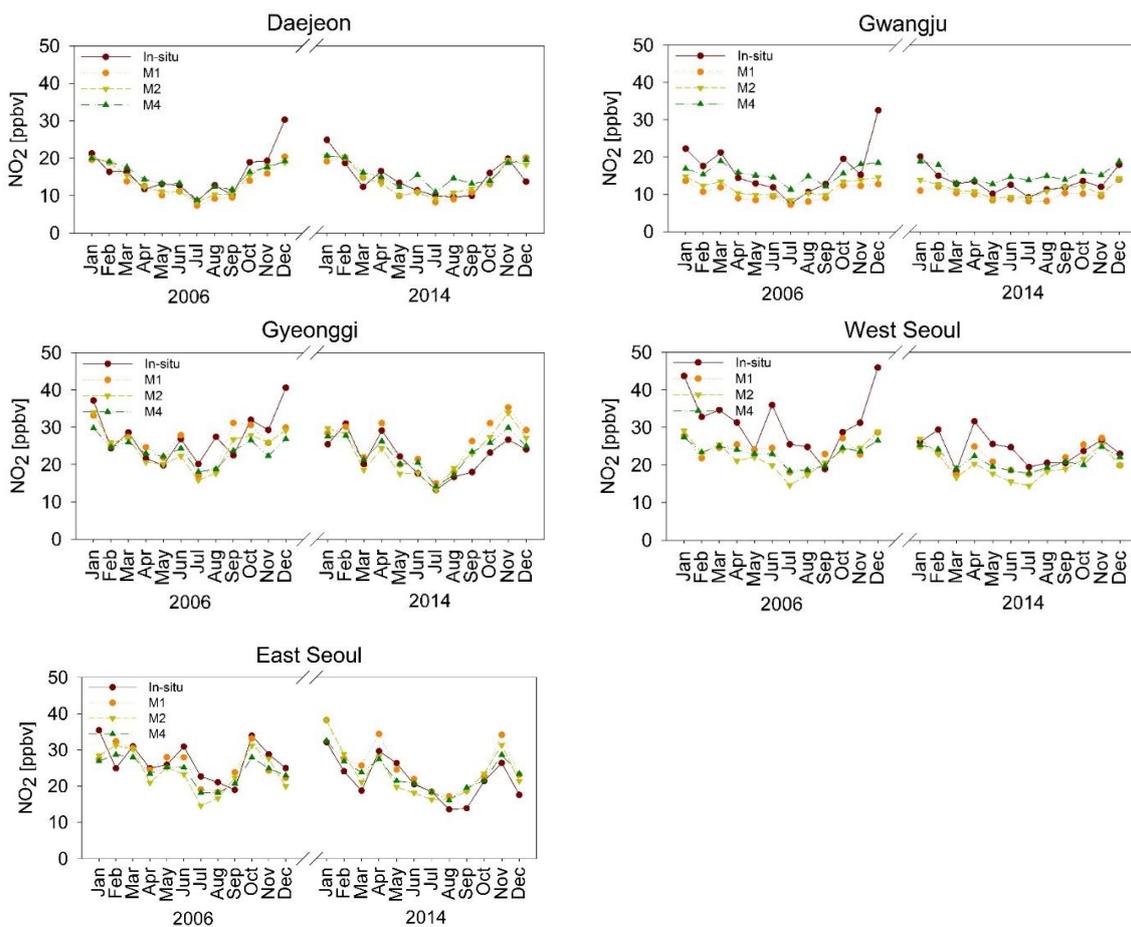


Figure 8. Time series of NO_2 $\text{VMR}_{\text{In-situ}}$ and NO_2 VMR_M estimated by M1, M2, and M4 for 2006 and 2014.

Figure 9 shows the R , slope, MB, MAE, RMSE and percent difference between NO_2 VMR_M and monthly mean NO_2 $\text{VMR}_{\text{In-situ}}$ in 2006 and 2014. In general, NO_2 VMR_M agreed better with NO_2 $\text{VMR}_{\text{In-situ}}$ than did the NO_2 VMR_{ST} . The value of R from M1, M2 and M4 and monthly mean NO_2 $\text{VMR}_{\text{In-situ}}$ ranged from 0.68 to 0.82 in all areas. MB was close to 0 in most study areas. MAE was less than 5 ppbv in Daejeon, Gwangju, Gyeonggi, and East Seoul where there is good agreement between NO_2 VMR_M from M1, M2, and M4 and monthly mean NO_2 $\text{VMR}_{\text{In-situ}}$, whereas MAEs in West Seoul ranged from 5.66 to 6.79. RMSEs between NO_2 $\text{VMR}_{\text{In-situ}}$ and NO_2 VMR_M from M1, M2, and M3 are found to be lower than 7 ppbv in the study areas except for West Seoul. In addition, the three models showed percent differences of less than 30% except for the value estimated from M1 in Gwangju.

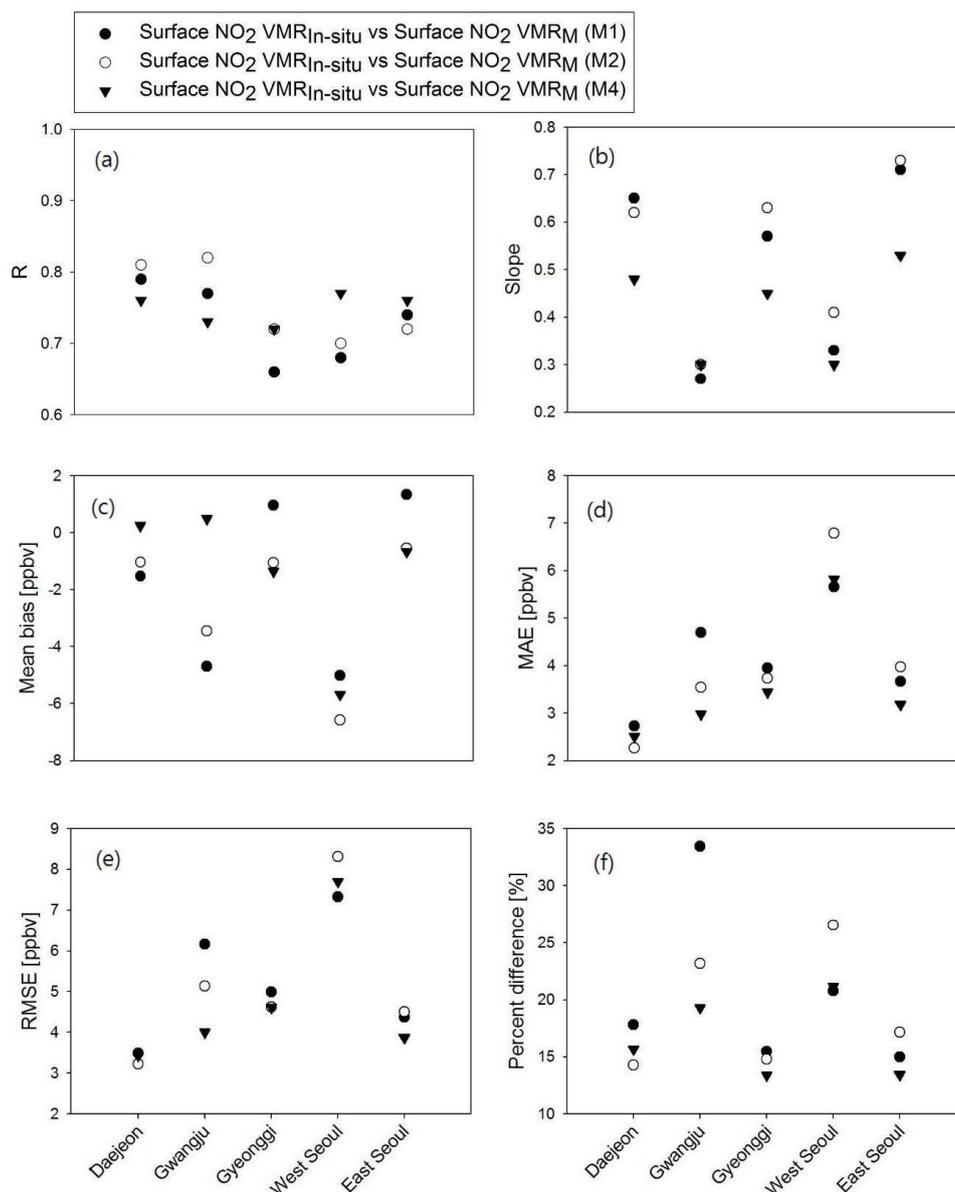


Figure 9. (a) R ; (b) slope; (c) MB; (d) MAE; (e) RMSE; and (f) percent difference between NO₂ VMR_M and monthly mean NO₂ VMR_{In-situ} in 2006 and 2014.

5. Discussion

In a previous study [18], tropospheric NO₂ VCDs obtained from GOME were compared with tropospheric NO₂ VCDs calculated using NO₂ concentrations obtained from both in situ measurements and the Model of Ozone and Related Tracers 2 (MOZART-2). There are also several previous studies estimating surface NO₂ VMR using satellite data [3,35]. Among them, Kharol et al. [3] estimated the annual variation of ground-level NO₂ concentrations using both GEOS-Chem data and OMI data. However, in the present study, NO₂ VMR_{ST} and NO₂ VMR_M were estimated for the first time at higher temporal resolution using three regression models with Trop NO₂ VCD_{OMI} as input.

5.1. Estimation of Surface NO₂ VMRs at a Specific Time (13:45 LT)

- Among the three regression models, the multiple regression model M3 performed best in estimating NO₂ VMR_{ST}. The linear regression model (M2), in which BLH is used as an independent variable in addition to Trop NO₂ VCD_{OMI}, has comparable performance to that of

the model (M1) which uses Trop NO_2 VCD_{OMI} as the only independent variable. The BLH varies with latitude [36], but the latitudinal variation of BLH is not well represented since the spatial resolution of the AIRS used in this study is coarser than the spatial resolution of OMI. It might also be associated with the BLH_{AIRS} data quality. We expect better results using BLH data obtained from LIDAR.

- The average difference was found to be 46.04% between NO_2 $\text{VMR}_{\text{In-situ}}$ and NO_2 VMR_{ST} obtained from M1, 44.29% between NO_2 $\text{VMR}_{\text{In-situ}}$ and NO_2 VMR_{ST} obtained from M2, and 31.50% between NO_2 $\text{VMR}_{\text{In-situ}}$ and NO_2 VMR_{ST} obtained from M3 in all cities, while there was moderate agreement in the temporal pattern of NO_2 variation between NO_2 $\text{VMR}_{\text{In-situ}}$ and NO_2 VMR_{ST} obtained from M1, M2, and M3 (Figure 6).
- In terms of statistical evaluation with respect to the in situ data, M3 showed the best performance in general.
- The results produced by M2 are not improved compared to those by M1 which may imply that surface NO_2 VMR is dominantly affected by tropospheric NO_2 column while the BLH effect could be negligible in areas of the present study. It might also be associated with the AIRS BLH data quality.

5.2. Estimation of Monthly Mean Surface NO_2 VMRs of a Specific Time (13:45 LT)

- We found good agreement in the temporal pattern between the estimated NO_2 VMR_M and monthly mean NO_2 $\text{VMR}_{\text{In-situ}}$ (Figure 8). However, there was a large difference between NO_2 $\text{VMR}_{\text{In-situ}}$ and NO_2 VMR_M in the period when there was a clear change in NO_2 VMR_M between one month and the next. Despite the use of NO_2 $\text{VMR}_{\text{In-situ}}$ located away from streets, the in situ measurement sites in West Seoul are located closer to streets than the in situ measurement sites in Daejeon and Gwangju. This may explain why there are more periods when NO_2 $\text{VMR}_{\text{In-situ}}$ changes rapidly in successive months. It is difficult to estimate the rapid change of NO_2 VMR near NO_2 sources with regression models that reflect the relationship between the in situ measurements and the OMI sensor covering both source and non-source areas in a single pixel.
- In terms of statistical evaluation, the three regression models (M1, M2, and M4) were found to be similar (Figure 9).
- NO_2 VMR_M shows better agreement with the NO_2 $\text{VMR}_{\text{In-situ}}$ than does NO_2 VMR_{ST} . The reason for the better performance in the monthly mean estimation could be attributed to reduced errors in the monthly mean OMI data [37] as well as fewer occasions with sudden monthly changes in NO_2 $\text{VMR}_{\text{In-situ}}$ than rapid day-to-day changes in NO_2 $\text{VMR}_{\text{In-situ}}$.

This present study provides the results in the condition of 2 km distance between the in situ NO_2 measurement location and NO_x point source. For a future study, performances of the models need to be investigated depending on the distance between the in situ NO_2 data and point sources. We expect that the regression methods used to estimate the surface NO_2 VMR using Trop NO_2 VCD_{OMI} will be useful in providing information on surface NO_2 VMR in metropolitan cities on a monthly timescale. In future research, the estimation of surface NO_2 VMR may be attempted at higher time resolution with geostationary satellite sensors (e.g., geostationary environmental monitoring spectrometer (GEMS), tropospheric emissions: monitoring of pollution (TEMPO), and Sentinel-4). In further work, improvements are needed in the input data or the model formulation before the surface NO_2 can be estimated on a daily basis.

6. Conclusions

In this study, monthly and specific time estimates of NO_2 VMR were obtained for the first time using three regression models in four metropolitan cities for two years, 2006 and 2014. The multiple regression model (M3) was found to perform best in estimating NO_2 VMR_{ST} in all cities. For surface NO_2 estimates at the specific time (13:45 LT), M3 generally gives better R, MAE, RMSE, and percent difference than the other two models (M1 and M2). A comparison between monthly surface NO_2

VMR estimates and those at the specific time showed that agreement with NO₂ VMR_{In-situ} was better for monthly estimates. In estimating NO₂ VMR_M, three regression models (M1, M2, and M4) showed similar performance. In estimating daily and monthly surface NO₂ VMR variations, when the surface NO₂ VMR changes rapidly, the difference between surface NO₂ VMR estimated from all models and NO₂ VMR_{In-situ} is found to be large. In future studies, using higher spatial resolution satellites is expected to improve the relationship with in situ measurements. In addition, the use of other independent variables that may co-vary with rapid changes of surface NO₂ VMR should be investigated.

Supplementary Materials: The Supplementary Materials are available online at <http://www.mdpi.com/2072-4292/9/6/627/s1>.

Acknowledgments: This subject is supported by Korea Ministry of Environment (MOE) as “K-COSEM Research Program”. This work was financially supported by the BK21 plus Project of the Graduate School of Earth Environmental Hazard System.

Author Contributions: Ozone Monitoring Instrument (OMI) and Atmospheric Infrared Sounder (AIRS) data collection and analysis were done by Wonei Choi. In situ data collection and analysis were done by Yun Gon Lee and Junsung Park. Estimation of surface NO₂ mixing ratios using three regression models was conducted by Daewon Kim, Hanlim Lee and Hyunkee Hong.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. IPCC. Climate Change 2007: The Physical Science Basis. In *Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*; Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K.B., Tignor, M., Miller, H.L., Eds.; Cambridge University Press: Cambridge, UK; New York, NY, USA, 2007; p. 2007.
2. Van der A, R.J.; Eskes, H.J.; Boersma, K.F.; Van Noije, T.P.C.; Van Roozendael, M.; De Smedt, I.; Peters, D.H.M.U.; Meijer, E.W. Trends, seasonal variability and dominant NO_x source derived from a ten year record of NO₂ measured from space. *J. Geophys. Res. Atmos.* **2008**, *113*. [[CrossRef](#)]
3. Kharol, S.K.; Martin, R.V.; Philip, S.; Boys, B.; Lamsal, L.N.; Jerrett, M.; Brauer, M.; Crouse, D.L.; McLinden, C.; Burnett, R.T.; et al. Assessment of the magnitude and recent trends in satellite-derived ground-level nitrogen dioxide over North America. *Atmos. Environ.* **2015**, *118*, 236–245.
4. Ackermann-Liebrich, U.; Leuenberger, P.; Schwartz, J.; Schindler, C.; Monn, C.; Bolognini, G.; Bongard, J.P.; Brändli, O.; Domenighetti, G.; Elsasser, S.; et al. Lung function and long term exposure to air pollutants in Switzerland. Study on Air Pollution and Lung Diseases in Adults (SAPALDIA) Team. *Am. J. Respir. Crit. Care Med.* **1997**, *155*, 122–129. [[PubMed](#)]
5. Schindler, C.; Ackermann-Liebrich, U.; Leuenberger, P.; Monn, C.; Rapp, R.; Bolognini, G.; Bongard, J.P.; Brändli, O.; Domenighetti, G.; Karrer, W.; et al. Associations between Lung Function and Estimated Average Exposure to NO₂ in Eight Areas of Switzerland. *Epidemiology* **1998**, *9*, 405–411. [[CrossRef](#)] [[PubMed](#)]
6. James Gauderman, W.; McConnell, R.O.B.; Gilliland, F.; London, S.; Thomas, D.; Avol, E.; Vora, H.; Berhane, K.; Rappaport, E.B.; Lurmann, F.; et al. Association between air pollution and lung function growth in southern California children. *Am. J. Respir. Crit. Care Med.* **2000**, *162*, 1383–1390. [[CrossRef](#)] [[PubMed](#)]
7. Panella, M.; Tommasini, V.; Binotti, M.; Palin, L.; Bona, G. Monitoring nitrogen dioxide and its effects on asthmatic patients: Two different strategies compared. *Environ. Monit. Assess.* **2000**, *63*, 447–458.
8. Smith, B.J.; Nitschke, M.; Pilotto, L.S.; Ruffin, R.E.; Pisaniello, D.L.; Willson, K.J. Health effects of daily indoor nitrogen dioxide exposure in people with asthma. *Eur. Respir. J.* **2000**, *16*, 879–885. [[PubMed](#)]
9. Boersma, K.F.; Jacob, D.J.; Trainic, M.; Rudich, Y.; DeSmedt, I.; Dirksen, R.; Eskes, H.J. Validation of urban NO₂ concentrations and their diurnal and seasonal variations observed from the SCIAMACHY and OMI sensors using in situ surface measurements in Israeli cities. *Atmos. Chem. Phys.* **2009**, *9*, 3867–3879. [[CrossRef](#)]
10. Demerjian, K.L. A review of national monitoring networks in North America. *Atmos. Environ.* **2000**, *34*, 1861–1884. [[CrossRef](#)]
11. Bechle, M.J.; Millet, D.B.; Marshall, J.D. Remote sensing of exposure to NO₂: Satellite versus ground-based measurement in a large urban area. *Atmos. Environ.* **2013**, *69*, 345–353.

12. Leue, C.; Wenig, M.; Wagner, T.; Klimm, O.; Platt, U.; Jähne, B. Quantitative analysis of NO_x emissions from Global Ozone Monitoring Experiment satellite image sequences. *J. Geophys. Res. Atmos.* **2001**, *106*, 5493–5505. [[CrossRef](#)]
13. Richter, A.; Burrows, J.P. Tropospheric NO₂ from GOME measurements. *Adv. Space Res.* **2002**, *29*, 1673–1683. [[CrossRef](#)]
14. Martin, R.V.; Chance, K.; Jacob, D.J.; Kurosu, T.P.; Spurr, R.J.; Bucsela, E.; Gleason, J.F.; Palmer, P.I.; Bey, I.; Fiore, A.M.; et al. An improved retrieval of tropospheric nitrogen dioxide from GOME. *J. Geophys. Res. Atmos.* **2002**, *107*. [[CrossRef](#)]
15. Boersma, K.F.; Eskes, H.J.; Brinksma, E.J. Error analysis for tropospheric NO₂ retrieval from space. *J. Geophys. Res. Atmos.* **2002**, *109*. [[CrossRef](#)]
16. Boersma, K.F.; Eskes, H.J.; Veefkind, J.P.; Brinksma, E.J.; Van Der A, R.J.; Sneep, M.; Van Der Oord, G.H.J.; Levelt, P.F.; Stammes, P.; Gleason, J.F.; et al. Near-real time retrieval of tropospheric NO₂ from OMI. *Atmos. Chem. Phys. Discuss.* **2006**, *6*, 12301–12345. [[CrossRef](#)]
17. Bucsela, E.J.; Celarier, E.A.; Wenig, M.O.; Gleason, J.F.; Veefkind, J.P.; Boersma, K.F.; Brinksma, E.J. Algorithm for NO₂ vertical column retrieval from the Ozone Monitoring Instrument. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 1245–1258. [[CrossRef](#)]
18. Ordóñez, C.; Richter, A.; Steinbacher, M.; Zellweger, C.; Nüß, H.; Burrows, J.P.; Prévôt, A.S.H. Comparison of 7 years of satellite-borne and ground-based tropospheric NO₂ measurements around Milan, Italy. *J. Geophys. Res. Atmos.* **2006**, *111*. [[CrossRef](#)]
19. Kim, N.K.; Kim, Y.P.; Morino, Y.; Kurokawa, J.I.; Ohara, T. Verification of NO_x emission inventory over South Korea using sectoral activity data and satellite observation of NO₂ vertical column densities. *Atmos. Environ.* **2013**, *77*, 496–508. [[CrossRef](#)]
20. Studies–Seoul, M.A.P. NIER. Available online: https://espo.nasa.gov/home/sites/default/files/documents/MAPS-Seoul_White%20Paper_26%20Feb%202015_Final.pdf (accessed on 17 June 2017).
21. Levelt, P.F.; Noordhoek, R. *OMI Algorithm Theoretical Basis Document Volume I: OMI Instrument, Level 0–1b Processor, Calibration & Operations*; ATBD-OMI-01; Version 1.1; NASA: Washington, DC, USA, 2002. Available online: http://www.sciamachy-validation.org/omi/documents/data/OMI_ATBD_Volume_1_V1d1.pdf (accessed on 17 June 2017).
22. Celarier, E.A.; Brinksma, E.J.; Gleason, J.F.; Veefkind, J.P.; Cede, A.; Herman, J.R.; Ionov, D.; Goutail, F.; Pommereau, J.P.; Lambert, J.C.; et al. Validation of Ozone Monitoring Instrument nitrogen dioxide columns. *J. Geophys. Res. Atmos.* **2008**, *113*. [[CrossRef](#)]
23. Levelt, P.F.; Hilsenrath, E.; Leppelmeier, G.W.; van den Oord, G.H.; Bhartia, P.K.; Tamminen, J.; de Haan, J.F.; Veefkind, J.P. Science objectives of the ozone monitoring instrument. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 1199–1208. [[CrossRef](#)]
24. Aumann, H.; Gaiser, S.; Ting, D.; Manning, E. *AIRS Algorithm Theoretical Basis Document Level 1B Part 1: Infrared Spectrometer*; EOS Project Science Office, NASA Goddard Space Flight Center: Greenbelt, MD, USA, 1999.
25. Fetzer, E.; McMillin, L.M.; Tobin, D.; Aumann, H.H.; Gunson, M.R.; McMillan, W.W.; Hagan, D.E.; Hofstadter, M.D.; Yoe, J.; Whiteman, D.N.; et al. Airs/amsu/hsb validation. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 418–431. [[CrossRef](#)]
26. Aumann, H.H.; Chahine, M.T.; Gautier, C.; Goldberg, M.D.; Kalnay, E.; McMillin, L.M.; Revercomb, H.; Rosenkranz, P.W.; Smith, W.L.; Staelin, D.H.; et al. AIRS/AMSU/HSB on the Aqua mission: Design, science objectives, data products, and processing systems. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 253–264. [[CrossRef](#)]
27. Chahine, M.T.; Pagano, T.S.; Aumann, H.H.; Atlas, R.; Barnet, C.; Blaisdell, J.; Chen, L.; Divakarla, M.; Fetzer, E.J.; Goldberg, M.; et al. AIRS: Improving Weather Forecasting and Providing New Data on Greenhouse Gases. *Bull. Am. Meteorol. Soc.* **2006**, *87*, 911–926. [[CrossRef](#)]
28. Lee, C.J.; Brook, J.R.; Evans, G.J.; Martin, R.V.; Mihele, C. Novel application of satellite and in-situ measurements to map surface-level NO₂ in the Great Lakes region. *Atmos. Chem. Phys.* **2011**, *11*, 11761–11775. [[CrossRef](#)]
29. Knepp, T.; Pippin, M.; Crawford, J.; Chen, G.; Szykman, J.; Long, R.; Cowen, L.; Cede, A.; Abuhassan, N.; Herman, J.; et al. Estimating surface NO₂ and SO₂ mixing ratios from fast-response total column observations and potential application to geostationary missions. *J. Atmos. Chem.* **2015**, *72*, 261–286. [[CrossRef](#)] [[PubMed](#)]

30. Xue, D.; Yin, J. Meteorological influence on predicting surface SO₂ concentration from satellite remote sensing in Shanghai, China. *Environ. Monit. Assess.* **2014**, *186*, 2895–2906. [[CrossRef](#)] [[PubMed](#)]
31. Seidel, D.J.; Ao, C.O.; Li, K. Estimating climatological planetary boundary layer heights from radiosonde observations: Comparison of methods and uncertainty analysis. *J. Geophys. Res. Atmos.* **2010**, *115*. [[CrossRef](#)]
32. Cui, W.; Yan, X. Adaptive weighted least square support vector machine regression integrated with outlier detection and its application in QSAR. *Chemom. Intell. Lab. Syst.* **2009**, *98*, 130–135. [[CrossRef](#)]
33. Kutner, M.H.; Nachtsheim, C.; Neter, J. *Applied linear Regression Models*; McGraw-Hill/Irwin: New York, NY, USA, 2004.
34. Sellke, T.; Bayarri, M.J.; Berger, J.O. Calibration of ρ values for testing precise null hypotheses. *Am. Stat.* **2001**, *55*, 62–71. [[CrossRef](#)]
35. Lamsal, L.N.; Martin, R.V.; Van Donkelaar, A.; Steinbacher, M.; Celarier, E.A.; Bucsela, E.; Dunlea, E.J.; Pinto, J.P. Ground-level nitrogen dioxide concentrations inferred from the satellite-borne Ozone Monitoring Instrument. *J. Geophys. Res. Atmos.* **2008**, *113*. [[CrossRef](#)]
36. Zeng, X.; Brunke, M.A.; Zhou, M.; Fairall, C.; Bond, N.A.; Lenschow, D.H. Marine atmospheric boundary layer height over the eastern Pacific: Data analysis and model evaluation. *J. Clim.* **2004**, *17*, 4159–4170. [[CrossRef](#)]
37. Team, O.M.I. *Ozone Monitoring Instrument (OMI) Data User's Guide*; NASA: Washington, DC, USA, 2009; Volume 7, pp. 6–10.



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).