Estimation of Surface NO\textsubscript{2} Volume Mixing Ratio in Four Metropolitan Cities in Korea Using Multiple Regression Models with OMI and AIRS Data

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

\textsuperscript{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.); cwyh3338@gmail.com (W.C.); junsung2ek@gmail.com (J.P.)

\textsuperscript{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 NO\textsubscript{2} volume mixing ratio (VMR) at a specific time (13:45 Local time) (NO\textsubscript{2} VMR\textsubscript{ST}) and monthly mean surface NO\textsubscript{2} VMR (NO\textsubscript{2} VMR\textsubscript{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 NO\textsubscript{2} VMR obtained from in situ measurements (NO\textsubscript{2} VMR\textsubscript{in-situ}) and tropospheric NO\textsubscript{2} 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 NO\textsubscript{2} VMR\textsubscript{in-situ} for two validation years (2006 and 2014). Of the three regression models, a multiple regression model shows the best performance in estimating NO\textsubscript{2} VMR\textsubscript{ST} and NO\textsubscript{2} VMR\textsubscript{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 NO\textsubscript{2} VMR\textsubscript{in-situ} and NO\textsubscript{2} VMR\textsubscript{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 NO\textsubscript{2} VMR\textsubscript{M}, with average R, slope, MB, MAE, RMSE, and percent difference between NO\textsubscript{2} VMR\textsubscript{in-situ} and NO\textsubscript{2} VMR\textsubscript{M} of 0.74, 0.49, \(-1.90\) ppbv, \(3.93\) ppbv, \(5.05\) ppbv, and \(18.76\)%, respectively.

Keywords: surface NO\textsubscript{2} volume mixing ratio; NO\textsubscript{2}; OMI; multiple regression

1. Introduction

The main anthropogenic source of nitrogen dioxide (NO\textsubscript{2}) is fossil fuel combustion, while natural sources of NO\textsubscript{2} include lightning, forest fires, and soil emissions [1,2]. In particular, since NO\textsubscript{2} is emitted in large quantities in automobile exhaust gas, NO\textsubscript{2} 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 NO\textsubscript{2} 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 NO\textsubscript{2} and various methods are used, with chemiluminescence, a well-known technique for measuring surface NO\textsubscript{2} 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 NO$_2$ VMR at high resolution [11]. In recent years, NO$_2$ vertical column density (VCD) has been measured from satellites that can monitor NO$_2$ at global scale over a short time scale. Space-borne sensors that have observed global distributions of NO$_2$ 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 NO$_2$ VMR so the NO$_2$ VCD obtained from satellites cannot be used directly. In recent years, studies have been conducted to investigate the feasibility of estimating the surface NO$_2$ VMR using the NO$_2$ VCD obtained from satellite measurements and, in particular, the correlation between the NO$_2$ VCD obtained from satellite measurements and the surface NO$_2$ VMR.

Ordóñez et al. [18] reported the correlation between tropospheric NO$_2$ VCD and the NO$_2$ VCD measured by GOME and ground based in situ devices in Milan. Kharol et al. [3] estimated the annual average ground-level NO$_2$ concentrations in North America using chemical transport model (GEOS-Chem) data and OMI NO$_2$ columns and also reported the annual trend of the estimated ground-level NO$_2$ concentrations. However, no studies have attempted to estimate the surface NO$_2$ VMR at higher temporal resolutions such as hourly and monthly using the NO$_2$ VCD measured by satellites.

In this present study, we estimate for the first time the surface NO$_2$ VMR at a specific time (13:45 Local time (LT)) (NO$_2$ VMR$_{ST}$) and the monthly mean surface NO$_2$ VMR (NO$_2$ VMR$_{MM}$) using two linear regression models and a multiple regression model with the tropospheric NO$_2$ VCD obtained from OMI (Trop NO$_2$ VCD$_{OMI}$) in five metropolitan cities. In addition, the performance of each regression method is evaluated by comparing the estimated surface NO$_2$ VMRs with those obtained from in situ measurement (NO$_2$ 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 NO$_2$ tended to increase in Seoul from 1995 to 2009 [20]. The study areas were selected where the surface NO$_2$ 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 NO$_2$ 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 NO$_2$ 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.
2.1. Data

The data used in this study are Trop NO\textsubscript{2} VCD\textsubscript{OMI} and Atmospheric Infrared Sounder (AIRS) boundary layer height (BLH\textsubscript{AIRS}), atmospheric temperature (Temp\textsubscript{AIRS}) and pressure (Press\textsubscript{AIRS}), together with in situ measurements of NO\textsubscript{2} VMR\textsubscript{In-situ}, surface temperature (Temp\textsubscript{In-situ}), surface pressure (Press\textsubscript{In-situ}), surface dew point (Dewpoint\textsubscript{In-situ}), surface wind speed (WS\textsubscript{In-situ}), and surface wind direction (WD\textsubscript{In-situ}) (see Table 1).

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

<table>
<thead>
<tr>
<th>Data</th>
<th>Time (LT)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Satellite</td>
<td></td>
</tr>
<tr>
<td>Trop NO\textsubscript{2} VCD</td>
<td></td>
</tr>
<tr>
<td>OMI Level3 NO\textsubscript{2} Daily data (OMNO2d)</td>
<td>13:45</td>
</tr>
<tr>
<td>BLH, Temperature, Pressure</td>
<td></td>
</tr>
<tr>
<td>AIRS/Aqua L3 Daily Support Product (AIRS + AMSU) V006 (AIRX3PD)</td>
<td>13:30</td>
</tr>
<tr>
<td>In situ</td>
<td></td>
</tr>
<tr>
<td>Surface NO\textsubscript{2} VMR</td>
<td></td>
</tr>
<tr>
<td>Air Korea</td>
<td></td>
</tr>
<tr>
<td>Surface Temperature, Surface Pressure, Surface Dew point, Surface Wind Speed, Surface Wind Data</td>
<td>13:00 and 14:00</td>
</tr>
</tbody>
</table>

2.1.1. Ozone Monitoring Instrument (OMI) Data

The Trop NO\textsubscript{2} VCD\textsubscript{OMI} data were obtained from OMI Level3 NO\textsubscript{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\textsubscript{2} data (Level-3 OMI NO\textsubscript{2} Cloud-Screened Total and Tropospheric Column NO\textsubscript{2} (V003)) are used in the present study (Cloud Fraction <30%).

2.1.2. Atmospheric Infrared Sounder (AIRS) Data

The BLH\textsubscript{AIRS}, Temp\textsubscript{AIRS}, and Press\textsubscript{AIRS} used in this study were obtained from the AIRS/Aqua L3 Daily Support Product (AIRS + AMSU) 1 degree × 1 degree V006 (AIRX3PD.00) from NASA Goddard Earth Sciences Data and Information Services Center (http://disc.sci.gsfc.nasa.gov/uii/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 NO$_2$ Data

The NO$_2$ VMR$_{\text{In-situ}}$ data were obtained from Air Korea (http://www.airkorea.or.kr/last_amb_hour_data). Since NO$_2$ VMR$_{\text{In-situ}}$ 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 NO$_2$ 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 NO$_2$ 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 NO$_2$ VMR$_{\text{In-situ}}$ obtained from in situ measurements located close to streets were excluded in this study. We used the average of three or more NO$_2$ VMR$_{\text{In-situ}}$ from stations located at least 2 km from each other.

2.1.4. In Situ Meteorological Data

The Temp$_{\text{In-situ}}$, Press$_{\text{In-situ}}$, Dewpoint$_{\text{In-situ}}$, WS$_{\text{In-situ}}$, and WD$_{\text{In-situ}}$ 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&FMNU=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, NO$_2$ VMR$_{\text{ST}}$ and NO$_2$ VMR$_{\text{M}}$ were estimated using three regression models with Trop NO$_2$ VCD$_{\text{OMI}}$. Table 2 summarizes the three models.

Table 2. Regression models used for surface NO$_2$ VMR estimation in this study.

<table>
<thead>
<tr>
<th>Model</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>NO$<em>2$ VMR$</em>{\text{In situ}}$ = $a$ Trop NO$<em>2$ VCD$</em>{\text{OMI}}$ + $b$</td>
</tr>
<tr>
<td>M2</td>
<td>NO$<em>2$ VMR$</em>{\text{In situ}}$ = aBLH NO$<em>2$ VMR$</em>{\text{OMI}}$ + b</td>
</tr>
<tr>
<td>M3</td>
<td>Section 3, Multiple regression Equation (1)</td>
</tr>
<tr>
<td>M4</td>
<td>Monthly</td>
</tr>
</tbody>
</table>

Notes: (a) NO$_2$ tropospheric vertical column density obtained from OMI; and (b) BLH NO$_2$ VMR$_{\text{OMI}}$ = Trop NO$_2$ VCD$_{\text{OMI}}$ Gas constant R * Temp$^{\text{AIRS}}$ * Press$^{\text{AIRS}}$ / Avogadro constant NA, where the AIRS pressure and temperature are boundary layer mean values, Gas constant R = 8.314472 m$^3$ pa K$^{-1}$ mol$^{-1}$ and Avogadro constant NA = 6.022 × 10$^{23}$ mol$^{-1}$.

3.1. M1

M1 is the linear regression equation where Trop NO$_2$ VCD$_{\text{OMI}}$ is used as the independent variable. Figure 2 shows the linear regression between Trop NO$_2$ VCD$_{\text{OMI}}$ and NO$_2$ VMR$_{\text{In-situ}}$ at 13:45 LT during the training period, with R$^2$ (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 NO$_2$ VCD$_{\text{OMI}}$ and monthly mean NO$_2$ VMR$_{\text{In-situ}}$ during the training period, with R$^2$, slope, and intercept of 0.62, 0.77, and 10.95, respectively. The final form of the M1 equation for estimating NO$_2$ VMR$_{\text{ST}}$ is shown in Table 3, and that for estimating NO$_2$ VMR$_{\text{M}}$ in Table 4.

Tables 3 and 4 show the equations M1, M2, M3, and M4 with the regression coefficients determined from the training period.
Table 3. Final form of the regression models used for estimating surface NO2 VMR at a specific time and $R^2$ obtained from the regression between NO2 VMR$_{\text{In-situ}}$ and the corresponding independent variable for the training period.

<table>
<thead>
<tr>
<th>Equation</th>
<th>R$^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>13:45 LT</td>
<td></td>
</tr>
<tr>
<td>M1 $NO_2$ VMR$<em>{ST}$ = 1.71 × Trop $NO_2$ VCD$</em>{OMI}$ - 0.68</td>
<td>0.47</td>
</tr>
<tr>
<td>M2 $NO_2$ VMR$<em>{ST}$ = 4.19 × BLH $NO_2$ VCD$</em>{OMI}$ + 1.57</td>
<td>0.38</td>
</tr>
<tr>
<td>M3 $NO_2$ VMR$<em>{ST}$ = 0.000602 × Trop $NO_2$ VCD$</em>{OMI}$ - 0.000107 × Temp$<em>{\text{In-situ}}$ - 0.000083 × Dewpoint$</em>{\text{In-situ}}$ + 0.000061 × Press$<em>{\text{In-situ}}$ - 0.000002 × BLH$</em>{\text{AIRS}}$ - 0.002435 × WS$<em>{\text{In-situ}}$ + 0.001190 × WD$</em>{\text{In-situ}}$ - 0.039996</td>
<td>0.47</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Equation</th>
<th>R$^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monthly mean</td>
<td></td>
</tr>
<tr>
<td>M1 $NO_2$ VMR$<em>{M}$ = 1.23 × Trop $NO_2$ VCD$</em>{OMI}$ + 4.74</td>
<td>0.62</td>
</tr>
<tr>
<td>M2 $NO_2$ VMR$<em>{M}$ = 2.92 × BLH $NO_2$ VCD$</em>{OMI}$ + 6.74</td>
<td>0.59</td>
</tr>
<tr>
<td>M3 $NO_2$ VMR$<em>{M}$ = 0.657241 × Trop $NO_2$ VCD$</em>{OMI}$ - 0.137334 × Dewpoint$<em>{\text{In-situ}}$ - 0.136096 × Press$</em>{\text{In-situ}}$ - 0.004331 × BLH$<em>{\text{AIRS}}$ - 0.770356 × WS$</em>{\text{In-situ}}$ + 2.370956 × WD (west)$_{\text{In-situ}}$ + 157.361668</td>
<td>0.63</td>
</tr>
</tbody>
</table>
3.2. M2

There might exist a minor fraction of the tropospheric NO₂ column in upper troposphere particularly because of lightning. However, the NO₂ amount in upper troposphere could be considered negligible in metropolitan cities, where a significant amount of NOX is emitted. Therefore, assuming Trop NO₂ VCD_{OMI} is mostly present within the PBL, the relationship between Trop NO₂ VCD_{OMI} and the surface NO₂ VMR may change as the PBL varies. However, a minor fraction of the tropospheric NO₂ column can also be in the upper tropospheric, particularly because of lightning. This NO₂ fraction in upper tropospheric might cause either small or negligible reduction in correlations of the OMI NO₂ VCD between and surface NO₂ 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 NO₂ VCD_{OMI} is first divided by BLH_{AIRS} to calculate the NO₂ concentration in the PBL and then converted to the NO₂ mixing ratio in the PBL (BLH NO₂ VMR_{OMI}) using Temp_{AIRS} and Press_{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 BLH NO₂ VMR_{OMI} and NO₂ VMR_{in-situ} at 13:45 LT during the training period. Here R², 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 R², slope and intercept of 0.59, 1.71, and 12.75, respectively. The final form of equation M2 to estimate NO₂ VMR_{ST} is shown in Table 3, and for the monthly values in Table 4.

![Figure 4](image-url)  
**Figure 4.** Scatter plot between BLH NO₂ VMR_{OMI} at a specific time (13:45 LT) and NO₂ VMR_{in-situ} to determine the regression coefficient for M1 for the training period 2007–2013.

![Figure 5](image-url)  
**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 NO$_2$ VMR$_{ST}$ and NO$_2$ VMR$_M$. Multiple regression equations consist of a dependent variable, independent variables, and their regression coefficients. In addition to Trop NO$_2$ 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 SO$_2$ 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 NO$_2$ 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:

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n + \epsilon$$

(1)

where $\hat{y}$ and $\beta_0$ are the dependent variable (NO$_2$ VMR$_{In-situ}$) and regression coefficient, respectively; $x_1, x_2, \ldots, x_n$ are the candidate independent variables (Trop NO$_2$ VCD$_{OMI}$, Dewpoint$_{In-situ}$, Press$_{In-situ}$, Temp$_{In-situ}$, BLH$_{AIRS}$, WS$_{In-situ}$, and WD$_{In-situ}$); $\beta_1, \beta_2, \ldots, \beta_n$ are the regression coefficients of the independent variables; and $\epsilon$ is the difference between observations (NO$_2$ VMR$_{In-situ}$) and estimated values (NO$_2$ VMR$_{estimate}$). The regression coefficients can be estimated by least square fitting:

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

(2)

where $y_j$ is the observed value with $m$ data points. By minimizing the sum of $\epsilon^2$, 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 ($\beta_n$) 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:

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

(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 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 NO$_2$ VMR$_{ST}$ is shown Table 3, and that for NO$_2$ VMR$_M$ in Table 4.
Table 5. Final independent variables included in multiple regression equations (M3 and M4).

<table>
<thead>
<tr>
<th>Final Selected Independent Variables</th>
<th>p-Value</th>
<th>VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trop NO$<em>2$ VCD$</em>{OMI}$</td>
<td>0</td>
<td>1.26</td>
</tr>
<tr>
<td>Temp$_{In-situ}$</td>
<td>0.000032</td>
<td>7.02</td>
</tr>
<tr>
<td>Dewpoint$_{In-situ}$</td>
<td>0.000306</td>
<td>7.16</td>
</tr>
<tr>
<td>Press$_{In-situ}$</td>
<td>0.009981</td>
<td>3.14</td>
</tr>
<tr>
<td>BLH$_{AIRS}$</td>
<td>$1.73 \times 10^{-15}$</td>
<td>1.12</td>
</tr>
<tr>
<td>WS$_{In-situ}$</td>
<td>$3.86 \times 10^{-133}$</td>
<td>1.33</td>
</tr>
<tr>
<td>WD$_{In-situ}$</td>
<td>$1.7493 \times 10^{-38}$</td>
<td>1.07</td>
</tr>
</tbody>
</table>

M3

<table>
<thead>
<tr>
<th>Final Selected Independent Variables</th>
<th>p-Value</th>
<th>VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trop NO$<em>2$ VCD$</em>{OMI}$</td>
<td>$2.4832 \times 10^{-89}$</td>
<td>1.64</td>
</tr>
<tr>
<td>Dewpoint$_{In-situ}$</td>
<td>0.000421</td>
<td>6.47</td>
</tr>
<tr>
<td>Press$_{In-situ}$</td>
<td>0.034582</td>
<td>6.65</td>
</tr>
<tr>
<td>BLH$_{AIRS}$</td>
<td>0.000834</td>
<td>2.32</td>
</tr>
<tr>
<td>WS$_{In-situ}$</td>
<td>$3.86 \times 10^{-133}$</td>
<td>1.59</td>
</tr>
<tr>
<td>WD$_{In-situ}$</td>
<td>$1.699 \times 10^{-7}$</td>
<td>1.25</td>
</tr>
</tbody>
</table>

M4

4. Results

4.1. Daily Estimates

Figure 6 shows the day-to-day variations of NO$_2$ VMR$_{In-situ}$ and NO$_2$ 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 NO$_2$ VMR$_{In-situ}$ and NO$_2$ VMR$_{ST}$ obtained with M3 compared to those between NO$_2$ VMR$_{In-situ}$ and NO$_2$ VMR$_{ST}$ obtained with M1 and M2. However, NO$_2$ obtained from M3 showed moderate agreement with NO$_2$ 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 NO$_2$ VMR$_{ST}$ and NO$_2$ 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 NO$_2$ 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 NO$_2$ VMR$_{ST}$ estimated from M3 show moderate agreement with NO$_2$ VMR$_{In-situ}$ in terms of magnitude. The RMSE from M3 is found to be lower than those from M1 and M2. The NO$_2$ 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 NO$_2$ VMR$_{In-situ}$ are lower in all study areas than from M1 and M2. In estimating NO$_2$ 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 NO$_2$ VMR$_{\text{In-situ}}$ and NO$_2$ 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).
MAE < 8.25 ppbv) of M3 is lower than those (6.58 ppbv < MAE < 11.41 ppbv) of M1 and M2, which means that NO$_2$ VMR$_{ST}$ estimated from M3 show moderate agreement with NO$_2$ VMR$_{In-situ}$ in terms of magnitude. The RMSE from M3 is found to be lower than those from M1 and M2. The NO$_2$ 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 NO$_2$ VMR$_{In-situ}$ are lower in all study areas than from M1 and M2.

4.2. Monthly Estimates

Figure 8 shows the temporal variation of monthly mean NO$_2$ VMR$_{In-situ}$ and NO$_2$ 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 NO$_2$ VMR$_M$ and monthly mean NO$_2$ VMR$_{In-situ}$. However, we found a large difference between NO$_2$ VMR$_{In-situ}$ and NO$_2$ VMR$_M$ in periods when there was a jump in NO$_2$ VMR$_{In-situ}$ between successive months. For example, no models calculated NO$_2$ VMR$_M$ that were similar to NO$_2$ VMR$_{In-situ}$ in December 2006, which is very different from that in November 2006. NO$_2$ VMR$_{In-situ}$ (NO$_2$ 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$_2$ VMR$_{in-situ}$ 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$_2$ VMR$_{in-situ}$. As described in Section 2, despite the use of NO$_2$ VMR$_{in-situ}$ 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$_2$ VMR$_{in-situ}$ 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$_2$ VMR$_{in-situ}$ and NO$_2$ VMR$_M$ estimated by M1, M2, and M4 for 2006 and 2014.](image)

Figure 9 shows the $R$, slope, MB, MAE, RMSE and percent difference between NO$_2$ VMR$_M$ and monthly mean NO$_2$ VMR$_{in-situ}$ in 2006 and 2014. In general, NO$_2$ VMR$_M$ agreed better with NO$_2$ VMR$_{in-situ}$ than did the NO$_2$ VMR$_ST$. The value of $R$ from M1, M2 and M4 and monthly mean NO$_2$ VMR$_{in-situ}$ 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$_2$ VMR$_{in-situ}$, whereas MAEs in West Seoul ranged from 5.66 to 6.79. RMSEs between NO$_2$ VMR$_{in-situ}$ 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.
Remote Sens. 2017, 9, 627  11 of 15

Figure 9. (a) $R$; (b) slope; (c) MB; (d) MAE; (e) RMSE; and (f) percent difference between NO$_2$ VMR$_M$ and monthly mean NO$_2$ VMR$_{In-situ}$ in 2006 and 2014.

5. Discussion

In a previous study [18], tropospheric NO$_2$ VCDs obtained from GOME were compared with tropospheric NO$_2$ VCDs calculated using NO$_2$ 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 NO$_2$ VMR using satellite data [3,35]. Among them, Kharol et al. [3] estimated the annual variation of ground-level NO$_2$ concentrations using both GEOS-Chem data and OMI data. However, in the present study, NO$_2$ VMR$_{ST}$ and NO$_2$ VMR$_M$ were estimated for the first time at higher temporal resolution using three regression models with Trop NO$_2$ VCD$_{OMI}$ as input.

5.1. Estimation of Surface NO$_2$ VMRs at a Specific Time (13:45 LT)

- Among the three regression models, the multiple regression model M3 performed best in estimating NO$_2$ VMR$_{ST}$. The linear regression model (M2), in which BLH is used as an independent variable in addition to Trop NO$_2$ 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 associate 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$_2$ VMR$_{In-situ}$ and NO$_2$ VMR$_{ST}$ obtained from M1, 44.29% between NO$_2$ VMR$_{In-situ}$ and NO$_2$ VMR$_{ST}$ obtained from M2, and 31.50% between NO$_2$ VMR$_{In-situ}$ 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$_2$ VMR$_{In-situ}$ 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 associate 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$_2$ VMR$_{In-situ}$ (Figure 8). However, there was a large difference between NO$_2$ VMR$_{In-situ}$ 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$_2$ VMR$_{In-situ}$ 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$_2$ VMR$_{In-situ}$ 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$_2$ VMR$_{In-situ}$ 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$_2$ VMR$_{In-situ}$ than rapid day-to-day changes in NO$_2$ VMR$_{In-situ}$.

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 cites 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 NO\textsubscript{2} VMR\textsubscript{In-situ} was better for monthly estimates. In estimating NO\textsubscript{2} VMR\textsubscript{M}, three regression models (M1, M2, and M4) showed similar performance. In estimating daily and monthly surface NO\textsubscript{2} VMR variations, when the surface NO\textsubscript{2} VMR changes rapidly, the difference between surface NO\textsubscript{2} VMR estimated from all models and NO\textsubscript{2} VMR\textsubscript{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 NO\textsubscript{2} VMR should be investigated.

**Supplementary Materials:** The Supplementary Materials are available online at [http://www.mdpi.com/2072-4292/9/6/627/s1](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 NO\textsubscript{2} 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**

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\textsubscript{x} source derived from a ten year record of NO\textsubscript{2} measured from space. *J. Geophys. Res. Atmos.* 2008, 113. [CrossRef]
9. Boersma, K.F.; Jacob, D.J.; Trainic, M.; Rudich, Y.; De Smedt, I.; Dirksen, R.; Eskes, H.J. Validation of urban NO\textsubscript{2} 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]


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 NO$_2$ vertical column densities. *Atmos. Environ.* **2013**, *77*, 496–508. [CrossRef]


© 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/).