

First-Time Estimation of HCHO Column over Asia using Multiple Regression with OMI and MODIS Data

W. N. Choi¹, H. K. Hong¹, Y. W. Lee¹, J. Y. Ryu², J. S. Park¹, and H. L. Lee^{1*}

¹*Division of Earth Environmental System Science Major of Spatial Information Engineering, Pukyong National University, Busan 608-737, Republic of Korea*

²*Department of Urban Environmental Engineering, Kyungnam University, Changwon 631-701, Republic of Korea*

Received October 1, 2015; revised June 18, 2016; accepted September 18, 2016; published May 7, 2019

ABSTRACT. A multiple regression method (MRM) is used for the first time with Ozone Monitoring Instrument (OMI) and Moderate Resolution Imaging Spectroradiometer (MODIS) data to estimate formaldehyde (HCHO) column over Asia in two categorized areas: (1) anthropogenic areas in East Asia (east-central China, Beijing, Seoul, and Tokyo) and (2) biogenic and pyrogenic areas (Indochina and South Borneo). In order to determine the multiple regression equations (MREs) for each study area, monthly mean and daily values of Nitrogen dioxide (NO₂) level, land surface temperature (LST), enhanced vegetation index (EVI), and fire radiative power (FRP) from 2005 to 2006 were used as independent variable candidates and monthly mean and daily HCHO measured by OMI (HCHO_{OMI}) for the same period was used as the dependent variable. The MREs estimated HCHO (HCHO_{MRM}) with inputs of independent variable data for from 2005 to 2006 whereas they estimated HCHO (HCHO_{VAL}) with inputs of independent variable data in 2007. The agreement between HCHO_{OMI} and HCHO_{VAL} is comparable with that between HCHO_{OMI} and HCHO_{MRM} in both monthly and daily scale. The average correlation coefficient, slope, mean bias, mean absolute error, root mean square error, and percent difference between monthly HCHO_{OMI} and monthly HCHO_{VAL} (between daily HCHO_{OMI} and daily HCHO_{MRM}) are 0.89 (0.58), 0.77 (0.34), -0.82×10^{15} molecules cm⁻² (-0.91×10^{15} molecules cm⁻²), 1.28×10^{15} molecules cm⁻² (3.01×10^{15} molecules cm⁻²), 1.75×10^{15} molecules cm⁻² (4.32×10^{15} molecules cm⁻²), 10.2% (24.1%). MRM can be a useful tool to provide HCHO in certain area in Asia.

Keywords: statistical modeling, multiple regression, OMI, MODIS, HCHO column, trace gas

1. Introduction

Formaldehyde (HCHO) is known to be toxic, with adverse effects on the human body that include, for example, damage to oral fibroblasts and epithelial cells (Nilsson et al., 1998). Biomass burning (Andreae and Merlet, 2001) and fossil fuel combustion (Olivier et al., 2003) are responsible for primary production of HCHO in the atmosphere, but only contribute about 1% of global HCHO (Stavrakou et al., 2009a). Most HCHO is generated by secondary production from photo-oxidation of both Volatile Organic Compounds (VOCs) and Non Methane Volatile Organic Compounds (NMVOCs), which originate from human activity, vegetation, and biomass burning (Palmer et al., 2003, 2006; Steinbacher et al., 2005; Fu et al., 2007; Barkley et al., 2009; Stavrakou et al., 2009a, 2009b). VOCs are one of the causes of photochemical smog while NMVOCs are a precursor of tropospheric ozone formation. The lifetime of HCHO is only a few hours, driven by photolysis mainly at wavelengths less than 400 nm and chemical reaction with hydroxyl radicals (Arlander et al., 1995).

Regional and seasonal variations of HCHO can be large due to its diverse emission sources and variability in its formation. In order to understand the temporal and spatial characteristics of HCHO on regional and global scales, satellite sensors for HCHO measurements have been developed over the last two decades by taking advantage of the extensive spatial coverage of the satellite measurements. Satellite sensors that have measured HCHO include the Global Ozone Monitoring Experiment (GOME) on the European Remote Sensing-2 satellite (ERS-2) launched in April 1995, the SCanning Imaging Absorption SpectroMeter for Atmospheric CHartographY (SCIAMACHY) on the European Environment Satellite (ENVISAT) launched in March 2002, the Atmospheric Chemistry Experiment-Fourier Transform Spectrometer (ACE-FTS) on Science Satellite (SCISAT-1) launched in August 2003, the Ozone Monitoring Instrument (OMI) on Aura launched in June 2004, GOME-2 on the Meteorological Operational satellites launched in January 2007 (MetOp-A) and September 2012 (MetOp-B), the Ozone Mapping and Profiler Suite (OMPS) on Suomi National Polar-orbiting Partnership (Suomi NPP) launched in October 2011.

Up to the present, measurements and analysis of global distributions of HCHO and its precursors have been actively conducted using hyperspectral Ultraviolet (UV) sensors on board of such satellite platforms. De Smedt et al. (2008, 2012) re-

* Corresponding author. Tel.: +82 51 629 6688; fax: +82 51 629 6653.

E-mail address: hllee@pknu.ac.kr (H. L. Lee).

ported the global-scale temporal and spatial characteristics of HCHO based on data from GOME, SCIAMACHY, OMI, and GOME-2 observations. Observations of HCHO over the West Pacific from SCIAMACHY and GOME-2 were validated using ship-based Multi axis differential optical absorption spectroscopy (MAX-DOAS) observations (Peters et al., 2012). The spatial and temporal variability of OMI-derived HCHO was compared with GOME, SCIAMACHY and the Goddard Earth Observing System-Chem Model (GEOS-Chem) (Bey et al., 2001) data over the US and Asia (Baek et al., 2014). Recently, spatial and seasonal characteristics of HCHO have been estimated using a Chemical Transport Model (CTM). Boeke et al. (2011) estimated HCHO using the GEOS-Chem model and validated the results by comparison with OMI data.

The multiple regression method (MRM) has been widely used for estimating concentrations of atmospheric constituents such as particulate matter (PM) and trace gases. Kim et al. (2012) used MRM to estimate primary organic carbon (POC) and secondary organic carbon (SOC) concentrations. Gupta et al. (2009) estimated PM_{2.5} in the planetary boundary layer using MRM with aerosol optical thickness measured from the Moderate Resolution Imaging Spectroradiometer (MODIS) as an independent variable. MRM often provides statistically robust predictions, although the accuracy of the estimation may vary with the target species. MRM has been also utilized to estimate diurnal variations in ambient ozone concentration (Abdul-Wahab et al., 2005). To date, however, there have been no attempts to estimate trace gases using MRM except for ozone.

Sometimes, HCHO cannot be retrieved from satellite observations because of poor measurement conditions (e.g., poor signal to noise ratio (S/N) at UV in winter) (Chance, 2002; Steck et al., 2008). The noise of the sensor itself is constant, however, S/N become poorer when radiation signal get weak especially in winter. The spectral fitting procedure (e.g., DOAS) takes place to retrieve the HCHO slant column density (SCD),

which is later converted into vertical column density using air mass factor (AMF). When the spectral fitting is done to retrieve the SCDs with the observation Level 1b data, if a magnitude of the unexplained residual signal's optical depth is larger than that of the HCHO, the noise included in the observation data is too large to retrieve the HCHO SCDs, which can be regarded as the poor observation data.

Therefore, as part of the efforts to provide HCHO column data, we, for the first time, used MRM with OMI and MODIS observational data to estimate HCHO vertical column density (HCHO VCD) over Asia where various HCHO sources are present. The present study also aims to demonstrate the feasibility of utilizing MRM for HCHO estimation, based on the agreement between HCHO VCD estimated by MRM (HCHO_{MRM}) and that measured by OMI (HCHO_{OMI}) in monthly scale. Additionally, MRM and simple interpolation method are used for the estimation of HCHO column density in daily basis.

2. Data and Methods

2.1. Multiple Regression Method

In the present study, MRM is used to estimate the spatial distribution of HCHO VCD in Asia. MRM can be carried out with least square fitting of a multiple regression equation (MRE) that consists of a dependent variable, independent variables, and their regression coefficients. In order to find the candidates for independent variables to be used in the MRE, we considered the candidates for independent variables that are known to have a high correlation with HCHO. Figure 1 shows the spatial distribution of correlation coefficient (R) of monthly mean values of HCHO_{OMI} against those of (a) Nitrogen dioxide (NO₂), (b) Fire Radiative Power (FRP), (c) Land Surface Temperature (LST), and (d) Enhanced Vegetation Index (EVI) for pixels where $|R| > 0.4$. In Figure 1, the HCHO_{OMI} data are obtained from the OMI measurements (OMI Formaldehyde Level2G Global

Table 1. Information about Data Products and Considered Flags Used in this Present Study

Sensor	Algorithm	Product name	Spatial resolution	Filtered flags and conditions	Accuracy
OMI	version 003	HCHO (OMHCHOG)	$0.25^\circ \times 0.25^\circ$	-Missing -Suspect -Bad -Cloud Fraction > 0.2 -Solar zenith angle > 70°	50-105% (Chance, 2002)
		NO ₂ (OMNO2d)	$0.25^\circ \times 0.25^\circ$	Cloud-free and atmospherically corrected -Solar zenith angle > 85° -Root Mean Squared Error Of Fit > 0.0003 -Terrain Reflectivity > 30% -Cloud Fraction > 0.3 (cloud-screened total column NO ₂)	5% (unpolluted case) 20% (polluted case) (Chance, 2002)
MODIS	version 005	LST (MOD11C3, MOD11C1)	$0.05^\circ \times 0.05^\circ$	Cloud-free and atmospherically corrected	1K (Wan, 1999)
		EVI (MOD13C2, MOD13C1) FRP (MYD14CMH, MYD14C8H)	$0.5^\circ \times 0.5^\circ$	Cloud-free and atmospherically corrected	1.5% (Miura et al., 2000) 5% (Freenborn et al., 2014)

binned data; OMHCHOG). The HCHO_{OMI} used in this present study are those with cloud free (cloud fraction < 0.2) and those flagged as “0” which indicates good quality level. The NO_2 VCD data ($\text{NO}_{2\text{OMI}}$) are also obtained from OMI measurements (OMI Level-3 Global Total and Tropospheric NO_2 VCD Data Product; OMNO2d). The HCHO_{OMI} and $\text{NO}_{2\text{OMI}}$ data are obtained from the NASA Goddard Earth Sciences Data and Information Service Center (<http://disc.sci.gsfc.nasa.gov/Aura/data-holdings/OMI>). The NO_2 retrievals of OMI have relatively reliable accuracy even for the poor measurement condition (e.g., large solar zenith angle, viewing zenith angle, and cloud fraction (less than 0.3)) since it uses visible wavelengths where the signal is stronger than that in UV wavelengths, so that the MRM method may be helpful to compensate the operational OMI HCHO retrievals. Also, the monthly FRP ($\text{FRP}_{\text{MODIS}}$), LST ($\text{LST}_{\text{MODIS}}$), and EVI data ($\text{EVI}_{\text{MODIS}}$) data are obtained from MODIS products (MYD14CHM, MOD11C3, and MOD13C2, respectively) available from Reverb (<http://reverb.echo.nasa.gov/reverb>). For the estimation of daily HCHO , MYD14C8H, MOD11C1, and MOD13C1 are used for $\text{FRP}_{\text{MODIS}}$, $\text{LST}_{\text{MODIS}}$, and $\text{EVI}_{\text{MODIS}}$, respectively. Since the temporal resolution of MYD14C8H and MOD13C1 are 8 days and 16 days, it may have an effect on the performance of daily HCHO estimation using MRM. The independent variable data that are used in this present study have been utilized in a lot of researches for several study areas in various purposes (Zhang et al., 2004; Giglio et al., 2006; Sims et al., 2008; Irie et al., 2012; Russell et al., 2012). Table 1 summarizes sensor, algorithm version, product name, spatial resolution, filtered flags and conditions, and accuracy for OMI and MODIS products.

NO_2 was selected as one of the candidates for independent variables to account for both primary (Olivier et al., 2003) and secondary (Steinbacher et al., 2005) HCHO formation related to anthropogenic fossil fuel combustion, and to biomass burning (Andreae and Merlet, 2001; Palmer et al., 2003, 2006; Gonzi et al., 2011). Furthermore, high level of nitrogen oxides ($\text{NO}_x = \text{NO} + \text{NO}_2$) can lead to active oxidation of isoprene which is one of the precursor of the HCHO in physicochemical point of view (Trainer et al., 1987). In Figure 1, there are diverse distributions of correlations between HCHO_{OMI} and the independent variable candidates due to various HCHO sources and their contribution to HCHO level on each area. In Figure 1a, there are high positive correlations between HCHO_{OMI} and $\text{NO}_{2\text{OMI}}$ in the Indochina peninsula, northeast Sumatra, and south Borneo in Southeast Asia. However, there are high negative correlations between HCHO_{OMI} and $\text{NO}_{2\text{OMI}}$ values over the east coast of China, South Korea, and the south coast of Japan due to their opposite seasonal cycles (Figure 1a). FRP, which represents the radiative heat emitted from fires such as forest fires, was chosen as a candidate for independent variables, since HCHO is produced from biomass burning. According to the previous study (Barkley et al., 2009), biomass burning is a significant source of HCHO which implies that the stronger biomass burning can lead to the more HCHO produced. Thus, FRP can be a reasonable independent variable candidate to explain HCHO in biomass burning areas. There are high positive correlations between HCHO_{OMI} and $\text{FRP}_{\text{MODIS}}$ in Sou-

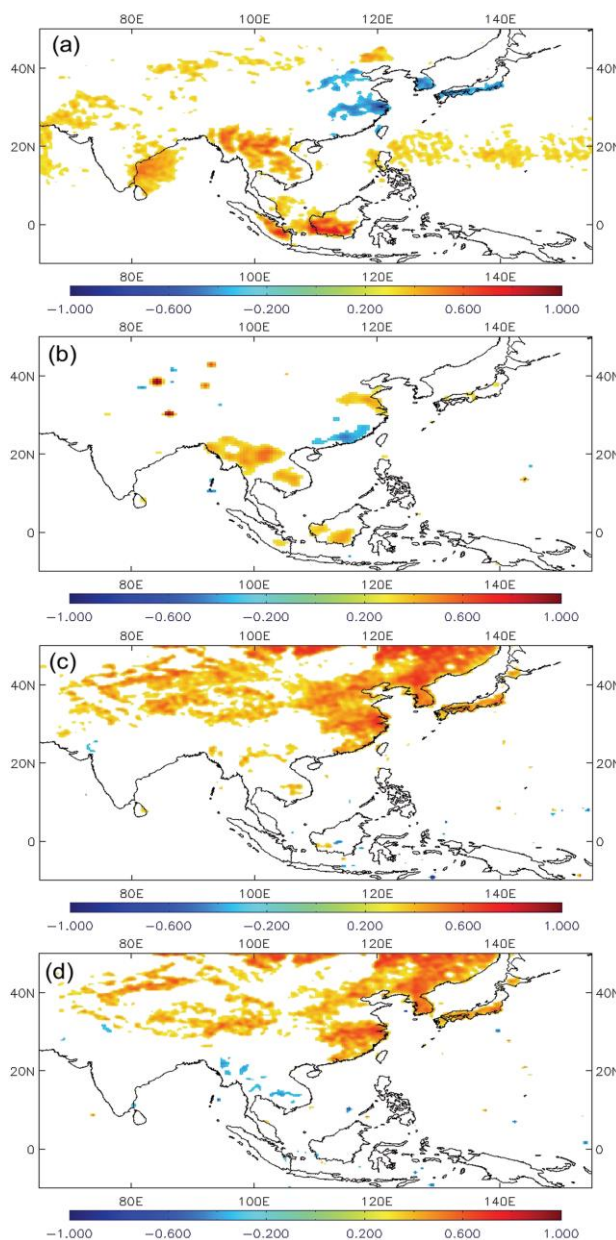


Figure 1. Distribution of correlation coefficient (R) between HCHO_{OMI} and (a) $\text{NO}_{2\text{OMI}}$, (b) $\text{FRP}_{\text{MODIS}}$, (c) $\text{LST}_{\text{MODIS}}$, and (d) $\text{EVI}_{\text{MODIS}}$ from January 2005 to December 2007 in Asia.

theast Asia and south Borneo in Figure 1b. However, high negative correlations are found in South China, due to the opposite seasonal cycles of HCHO_{OMI} and $\text{FRP}_{\text{MODIS}}$ in this area: HCHO_{OMI} and $\text{FRP}_{\text{MODIS}}$ have their maximum values in summer and winter, respectively. Although winter FRP in South China was likely to be enhanced by active wood fuel combustion, the enhanced FRP could not lead to increase in HCHO level due to negligible HCHO emission from heating with wood fuel (Andreae and Merlet, 2001; Cuiping et al., 2004). In order to account for biogenic HCHO , $\text{LST}_{\text{MODIS}}$ and $\text{EVI}_{\text{MODIS}}$ are selected as candidates for independent variables in the MRE since HCHO precursors such as isoprene are significantly in-

fluenced by type of vegetation, foliar density, leaf age, photosynthetic photon flux density (PPFD), and temperature (Guenther et al., 2000). Leaf area index was also initially considered as one of the independent variable candidates. The emission of biogenic VOCs such as isoprene, emitted from biosynthesis of vegetation are dependent with temperature (Laothawornkitkul et al., 2009) so that LST_{MODIS} and EVI_{MODIS} can be physico-chemically suitable independent variable to estimate $HCHO_{OMI}$ column density. Leaf area index, however, is eventually excluded from the independent variable candidates due to its high value of variation inflation factor (VIF) against EVI . The details of VIF are explained later in this Section 2.1. Additional explanation of correlations between independent variable candidates and $HCHO_{OMI}$ can be found later in Section 2.3. In general, positive correlations between $HCHO_{OMI}$ and LST_{MODIS} and EVI_{MODIS} in Northeast Asia tend to be higher than those in Southeast Asia (Figure 1c and 1d). In Northeast Asia, we found that the spatial distribution of R between $HCHO_{OMI}$ and LST_{MODIS} (Figure 1c) is very similar to that between $HCHO_{OMI}$ and EVI_{MODIS} except in the central east industrial region in China (Figure 1d). However, we found no significant correlations between $HCHO_{OMI}$ and LST_{MODIS} and EVI_{MODIS} in Southeast Asia (Figure 1c and 1d). These insignificant correlations may be driven by the small amplitude of seasonal LST_{MODIS} and EVI_{MODIS} variations due to strong but steady solar radiation throughout the year.

The multiple regression equation can be defined as 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 dependent variable ($HCHO_{MRM}$) and constant coefficient, x_1, x_2, \dots, x_n are the independent variables (NO_{2OMI} , LST_{MODIS} , EVI_{MODIS} , and FRP_{MODIS}), $\beta_1, \beta_2, \dots, \beta_n$ are the regression coefficients of the independent variables, and ε is the difference between observations ($HCHO_{OMI}$) and estimated values ($HCHO_{MRM}$). The regression coefficients can be estimated by the least square fitting (Equation 2):

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

where y_j is observed value with m numbers of data points. By minimizing the sum of ε^2 , regression coefficients can be derived. To determine MREs which ensure statistical significances, two criteria are considered: VIF and p -value. First, we examined the VIF that explains the multicollinearity of an independent variable candidate with regard to other independent variable candidates. The VIF of the j -th independent variable is expressed as:

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

where R_j^2 is the coefficient of determination for the regression of x_j against the other (a regression that does not involve the dependent variable y). The VIF indicates how much x_j is

correlated with the variables. A candidate for independent variables with a very high VIF can be considered redundant and should be removed from the MRE. The candidates for independent variables that do not satisfy the criterion $VIF > 10$ (Kutner et al. 2004), were excluded from the independent variables. We also used p -value to select independent variables. The highest, still statistically significant p -level was shown by Sellke et al. (2001) to be 5%. Among the independent variables that satisfy the VIF criterion, those that also satisfy the p -value less than 0.05 (p -value < 0.05) are selected as final independent variables in the MRE. The independent variables and regression coefficients determined by least square fitting for each area are shown in Table 2.

In order to determine the final form of the MREs in monthly and daily basis, monthly mean and daily values of NO_{2OMI} , FRP_{MODIS} , LST_{MODIS} , and EVI_{MODIS} from 2005 and 2006 were used as independent variable candidates and monthly mean and daily $HCHO_{OMI}$ for the same period was used as the dependent variable. In order to reflect the various $HCHO$ source types possibly present in each area selected in Figure 2, the data for the independent variables are averaged within each area.

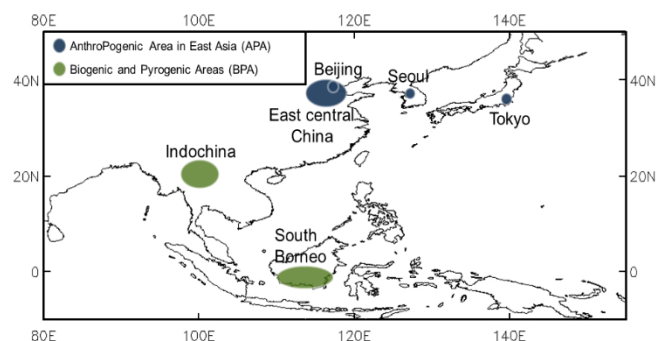


Figure 2. The areas where the MRM was used for $HCHO$ estimation.

2.2. Regions of Interest

We selected the areas where MRM is used to estimate $HCHO$ according to two criteria: first, in order to include the areas where $HCHO$ estimation using MRM may be feasible, we selected areas where there are high correlations ($|R| > 0.4$) between $HCHO_{OMI}$ and one or more independent variables in Figure 1. Among the areas that satisfy the primary criterion, only the areas where the monthly value of $HCHO_{OMI}$ is within the top 10% ($> 1.28 \times 10^{16}$ molecules cm^{-2}) were selected since areas with high $HCHO$ are more likely to have explicit seasonal cycles. Depending on the correlation coefficient from the linear regressions, the area types are subjectively classified as:

Biogenic and pyrogenic area: the regions where there are positive correlations between $HCHO_{OMI}$ against NO_{2OMI} and FRP_{MODIS} ($R \geq 0.4$).

Anthropogenic area: The regions where there are positive correlations between $HCHO_{OMI}$ against LST_{MODIS} and EVI_{MODIS} ($R \geq 0.4$) but negative correlations between $HCHO_{OMI}$ and NO_{2OMI} ($R \leq -0.4$).

The areas that satisfy these two criteria in Southeast Asia are Indochina (18° - 22° N, 98° - 102° E) and South Borneo (4° S - 0° N, 108° - 116° E) where biomass burning is reported to take place regularly (Fu et al., 2007). In East Asia, the areas that satisfy these criteria are the three megacities; Beijing (39.5° - 40.5° N, 115.5° - 117.5° E), Seoul (37° - 38° N, 126.5° - 127.5° E), and Tokyo (35.5° - 36.5° N, 139° - 140° E), along with industrial areas along the east coast of China (36° - 40° N, 113° - 119.75° E). Finally, these selected areas (Figure 2) were classified into two categories: (1) Biogenic and Pyrogenic Areas (BPA; green circles; Indochina, South Borneo) and (2) Anthropogenic Areas in East Asia (APA; blue circles; east-central China, Beijing, Seoul, Tokyo). South China is excluded in the regions of interest due to its complexity of HCHO sources including biogenic, anthropogenic, and biomass burning in the post-harvest season.

2.3. Seasonal Characteristics of the Variables Used in the MRE

2.3.1. Biogenic and Pyrogenic Areas (BPA)

Figure 3 shows time series of HCHO_{OMI} , $\text{NO}_{2\text{OMI}}$, $\text{FRP}_{\text{MODIS}}$, $\text{LST}_{\text{MODIS}}$, and $\text{EVI}_{\text{MODIS}}$ from January 2005 to December 2007 in the BPA of Indochina and South Borneo, which are marked with green circles in Figure 2. In South Borneo, there is good agreement between the seasonal cycles of HCHO_{OMI} and the three independent variables ($\text{NO}_{2\text{OMI}}$, $\text{FRP}_{\text{MODIS}}$, and $\text{LST}_{\text{MODIS}}$) in Figure 3. There are high $\text{FRP}_{\text{MODIS}}$ values from August to November, which could be associated with an increase in $\text{FRP}_{\text{MODIS}}$ due to biomass burning events. Most fires occurred in the period were related to slash and burn activities (Langner et al., 2007). In 2006, the monthly average HCHO_{OMI} is 53% higher than in other years, which is in good agreement with the 56% increase in monthly average $\text{FRP}_{\text{MODIS}}$ during the biomass burning period in 2006 relative to 2005 and 2007. Furthermore, during the biomass burning period in 2006, $\text{NO}_{2\text{OMI}}$ increases by about 26% and $\text{LST}_{\text{MODIS}}$ reaches its maximum. In Indochina, the seasonal cycle of HCHO_{OMI} is in good agreement with those of the independent variables during the biomass burning events that took place annually from January to May.

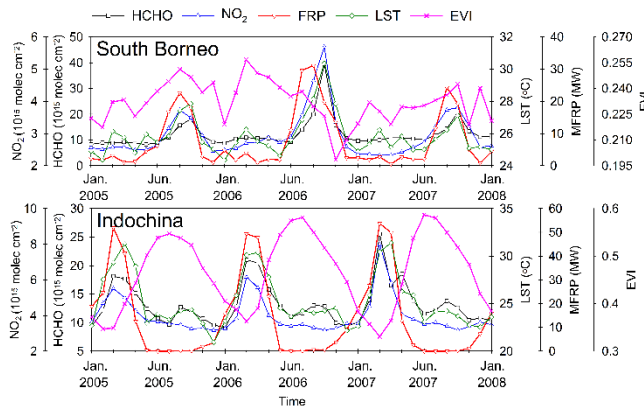


Figure 3. Time series of monthly HCHO_{OMI} , $\text{NO}_{2\text{OMI}}$, $\text{FRP}_{\text{MODIS}}$, $\text{LST}_{\text{MODIS}}$, and $\text{EVI}_{\text{MODIS}}$ (see text for an explanation of these variables) in the BPA (South Borneo and Indochina).

2.3.2. Anthropogenic Areas in East Asia (APA)

Figure 4 shows time series of HCHO_{OMI} , $\text{NO}_{2\text{OMI}}$, $\text{LST}_{\text{MODIS}}$, and $\text{EVI}_{\text{MODIS}}$ from January 2005 to December 2007 in the APA marked with blue circles in Figure 2 (east-central China, Beijing, Seoul, and Tokyo). We have excluded $\text{FRP}_{\text{MODIS}}$ from Figure 4 as no fires were detected in those areas during our study period. The values of HCHO_{OMI} tend to increase in summer, which could be attributed to increased secondary HCHO from fossil fuel combustion under enhanced solar radiation conditions in summer (Ahmad et al., 2011). According to a previous study (Pang et al., 2009), such an increase in HCHO_{OMI} could be partly explained by enhanced emission from biogenic HCHO even in the APA in summer when both $\text{LST}_{\text{MODIS}}$ and $\text{EVI}_{\text{MODIS}}$ tend to increase as shown in Figure 4. NO_2 tends to reach its minimum in summer when the solar radiation is strongest. Thus, $\text{NO}_{2\text{OMI}}$ tend to be negatively correlated with HCHO_{OMI} , $\text{LST}_{\text{MODIS}}$, and $\text{EVI}_{\text{MODIS}}$ as shown in Figure 4.

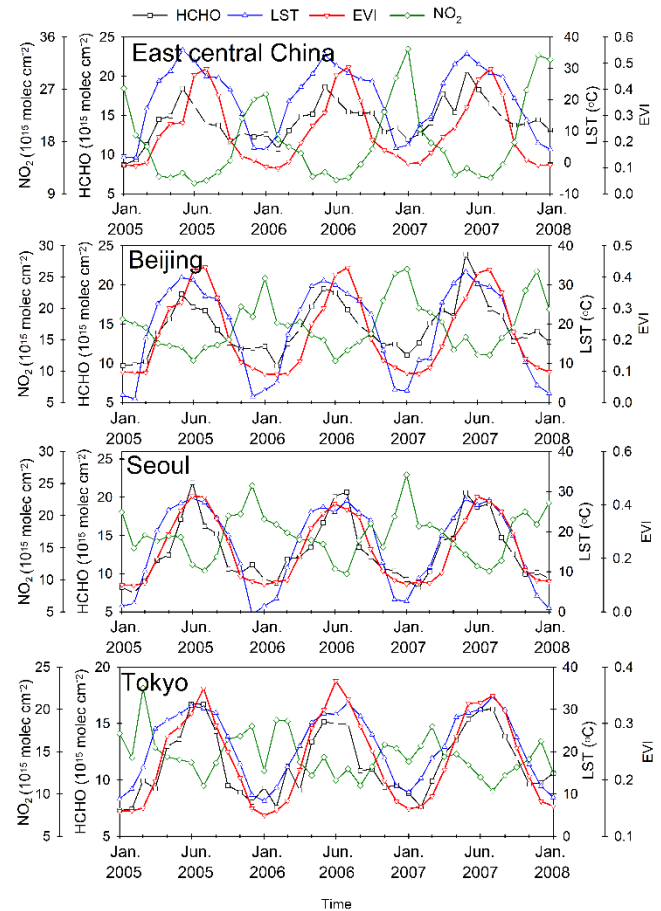


Figure 4. Time series of monthly HCHO_{OMI} , $\text{NO}_{2\text{OMI}}$, $\text{LST}_{\text{MODIS}}$, and $\text{EVI}_{\text{MODIS}}$ (see text for an explanation of these variables) in the APA.

3. Results

3.1. Estimation of Monthly HCHO Column Density

3.1.1. Determination of the MREs

The regression coefficients are obtained by least square

fitting (Timm, 2002). Table 2 presents the MREs with regression coefficients determined for each region of interest. The variables, that satisfy the two criteria $VIF > 10$ (Kutner et al., 2004) and $p\text{-value} < 0.05$ (Sellke et al., 2001), are included in the MREs in Table 2. We used two different terms for NO_2 in Table 2 to represent the two different major sources, biomass burning and fossil fuel combustion. In the biogenic and pyrogenic dominant areas with very low population density where there is negligible fossil fuel combustion activity, NO_2 is categorized as BB_{NO_2} (NO_2 derived from biomass burning). Since biomass burning was not detected in the APA during the study period, as inferred from the unavailability of $\text{FRP}_{\text{MODIS}}$, NO_2 in the APA is categorized as FFC_{NO_2} (NO_2 derived from fossil fuel combustion). We found better agreement between monthly HCHO_{MRM} and monthly HCHO_{OMI} in South Borneo than in Indochina. In BPA, both BB_{NO_2} and $\text{LST}_{\text{MODIS}}$ are included in the MREs for both South Borneo and Indochina. In the MRE for South Borneo, $\text{FRP}_{\text{MODIS}}$ is included while $\text{EVI}_{\text{MODIS}}$ is excluded due to a high $p\text{-value}$ (0.753). In the MRE for Indochina, $\text{EVI}_{\text{MODIS}}$ is included whereas $\text{FRP}_{\text{MODIS}}$ is excluded from the MRE for Indochina due to high VIF (14.8). The correlation coefficients (R) obtained from linear regression between monthly HCHO_{OMI} and monthly HCHO_{MRM} are 0.99 in South Borneo and 0.94 in Indochina, respectively.

In APA, $\text{FRP}_{\text{MODIS}}$ is excluded from independent variable candidates for MREs due to its unavailability in these areas. Only a few number of $\text{FRP}_{\text{MODIS}}$ data are available in east-central China, Beijing, Seoul, and Tokyo. Both FFC_{NO_2} and $\text{LST}_{\text{MODIS}}$ are included in the MREs for Beijing and east-central China. For Beijing, $\text{EVI}_{\text{MODIS}}$ is included in the MRE. However, $\text{EVI}_{\text{MODIS}}$ is excluded from the MRE for east-central China due to a high $p\text{-value}$ (0.382). For Seoul and Tokyo, only $\text{EVI}_{\text{MODIS}}$ is included in the MREs. $\text{LST}_{\text{MODIS}}$ and FFC_{NO_2} is excluded from the MREs for Seoul and Tokyo due to the high $p\text{-values}$. The correlation coefficients obtained from linear regression between monthly HCHO_{OMI} and monthly HCHO_{MRM} are 0.89,

0.97, 0.91, and 0.90 in east-central China, Beijing, Seoul, and Tokyo, respectively.

3.1.2. An Evaluation of the Performance of the MREs

The regression coefficients of the MREs in Table 2 were determined based on monthly mean data for dependent and independent variables for 24 months period from 2005 and 2006. For the assessment of the performance of the MREs in Table 2, monthly mean HCHO_{OMI} data in 2005 and 2006 were compared with HCHO estimated (monthly HCHO_{MRM}) using the MREs in Table 2 with the inputs of monthly mean independent variable data for 24 months period from 2005 and 2006. In addition, we tried to carry out an unbiased assessment of the performance of the MREs in Table 2 via estimating HCHO (monthly HCHO_{VAL}) using the MREs in Table 2 with the inputs of monthly mean independent variable data for 12 months period in 2007. Monthly HCHO_{VAL} was compared with monthly HCHO_{OMI} in 2007.

Figure 5 b-c shows spatial distributions of two years mean HCHO_{OMI} for 24 months period from 2005 to 2006 and two years mean values of HCHO_{MRM} for the same period in South Borneo, whereas Figure 5d-e shows annual mean HCHO_{OMI} for 12 months period in 2007, and annual mean values of HCHO_{VAL} for the same period in South Borneo. Figures 6 and 7 are the same as Figure 5 but for Indochina and east-central China, respectively. In general, spatial distributions of two years mean HCHO_{MRM} and annual mean HCHO_{VAL} in Figures 5, 6, and 7 show good agreements with HCHO_{OMI} . Spatial distributions of two years mean HCHO_{MRM} and annual mean HCHO_{VAL} are not shown here for Beijing, Seoul, and Tokyo due to the small number of pixels which compose each of those megacity areas. However, in order to quantify the performance of the MREs for all of the study areas including Beijing, Seoul, and Tokyo, the statistical values between monthly HCHO_{OMI} and HCHO_{MRM} is compared with that between monthly HCHO_{OMI} and monthly HCHO_{VAL} .

Table 2. Final Form of the MRE and Correlation Coefficient (R) Obtained from the Linear Regression between Monthly HCHO_{MRM} and HCHO_{OMI} for 24 Months Period from 2005 and 2006

Type	Area	VIF*	p-value**	Multiple regression equation [$\times 10^{15}$ molecules cm^{-2}]***	Comparison with monthly HCHO_{OMI} R ****
BPA	South Borneo ($4^\circ \text{ S} - 0^\circ \text{ N}$, $108^\circ - 116^\circ \text{ E}$)	—	EVI (0.753)	$6.21 \times \text{BB}_{\text{NO}_2} + 0.01 \times 1.27^{\text{LST}} - 0.202 \times \text{FRP} - 11.7$	0.99
	Indochina ($18^\circ - 22^\circ \text{ N}$, $98^\circ - 102^\circ \text{ E}$)	FRP (14.8)	—	$2.51 \times \text{BB}_{\text{NO}_2} + 0.58 \times \text{LST} + 11.4 \times \text{EVI} - 16.8$	0.94
APA	East-central China ($36^\circ - 40^\circ \text{ N}$, $113^\circ - 119.75^\circ \text{ E}$)	—	EVI (0.382)	$0.284 \times \text{FFC}_{\text{NO}_2} + 0.326 \times \text{LST} + 2.34$	0.89
	Beijing ($39.5^\circ - 40.5^\circ \text{ N}$, $115.5^\circ - 117.5^\circ \text{ E}$)	—	—	$0.298 \times \text{FFC}_{\text{NO}_2} + 0.247 \times 1.12^{\text{LST}} + 7.14 \times \text{EVI} + 3.90$	0.97
	Seoul ($37^\circ - 38^\circ \text{ N}$, $126.5^\circ - 127.5^\circ \text{ E}$)	—	LST (0.748) FFC_{NO_2} (0.245)	$7.50 \times 9.11^{\text{EVI}} - 0.269$	0.91
	Tokyo ($35.5^\circ - 36.5^\circ \text{ N}$, $139^\circ - 140^\circ \text{ E}$)	—	FFC_{NO_2} (0.275) LST (0.327)	$5.34 \times 20.7^{\text{EVI}} + 0.031$	0.90

* The independent variable candidates that do not satisfy the VIF criterion ($VIF < 10$) are indicated.

** The independent variable candidates that do not satisfy the $p\text{-value}$ criterion ($p\text{-value} < 0.05$) are indicated.

*** FRP is excluded from independent variable candidates for MREs in APA.

**** Correlation coefficient between monthly HCHO_{OMI} and HCHO_{MRM} .

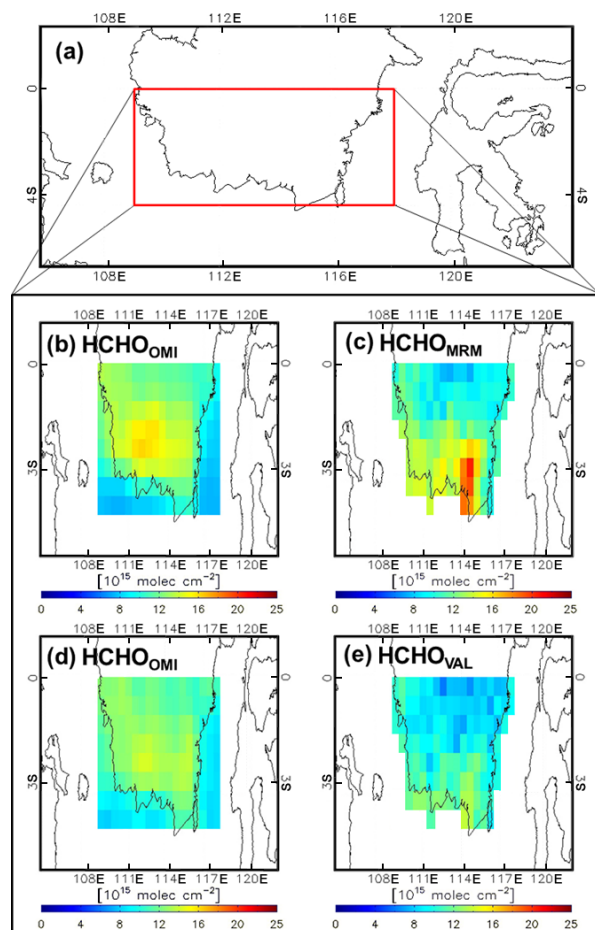


Figure 5. (a) Map of South Borneo, spatial distributions of (b) two years mean HCHO_{OMI} for 24 months period from 2005 to 2006, (c) two years mean values of HCHO estimated (HCHO_{MRM}) using the MREs in Table 1 with the inputs of monthly mean independent variable data for 24 months period from 2005 and 2006, and (d) annual mean HCHO_{OMI} for 12 months period in 2007, and (e) annual mean values of HCHO estimated (HCHO_{VAL}) using the MREs in Table 1 with the inputs of monthly mean independent variable data for 12 months period in 2007.

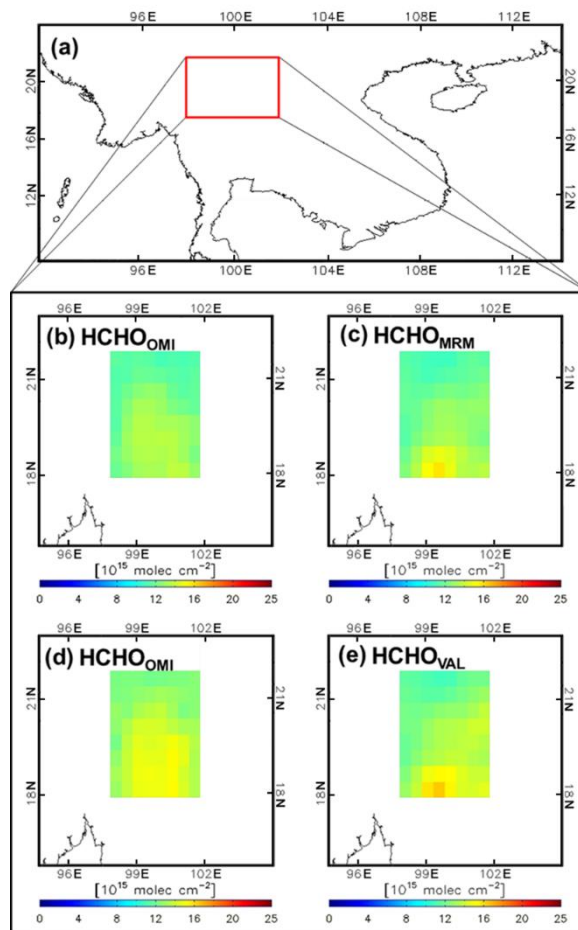


Figure 6. Same as Figure 5 but for Indochina.

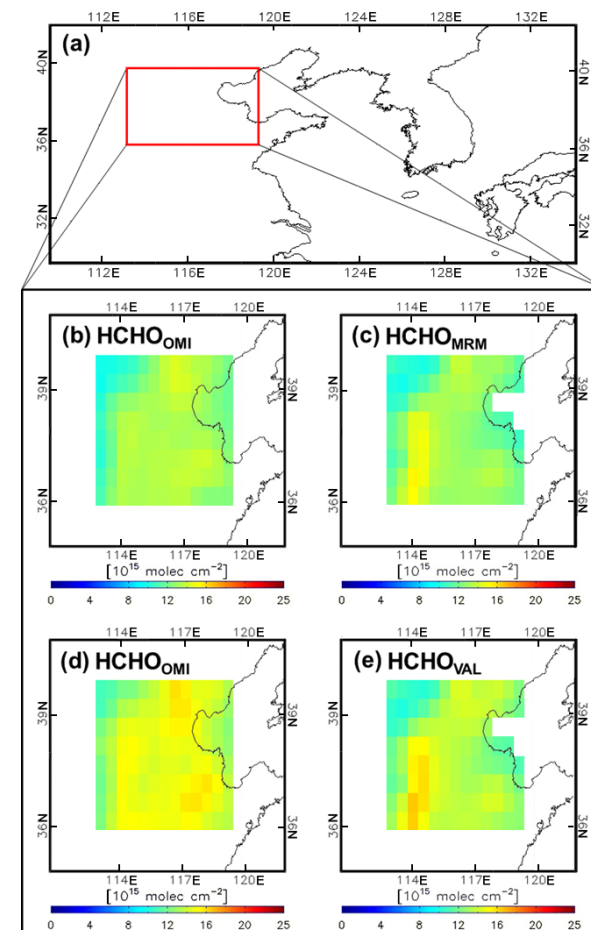


Figure 7. Same as Figure 5 but for China.

Figure 8 shows the (a) R (correlation coefficient), (b) slope, (c) mean bias, (d) MAE (mean absolute error), (e) RMSE (root mean square error) and (f) percent difference between monthly mean monthly HCHO_{OMI} in 2005 and 2006 against monthly HCHO_{MRM} , and between monthly HCHO_{OMI} in 2007 against monthly HCHO_{VAL} . The correlation coefficient (R) and the slopes are found to be up to 0.99 and 0.98, respectively, showing very good agreements between monthly HCHO_{OMI} and monthly HCHO_{MRM} . The mean bias values are found to be close to zero in Figure 5c, implying negligible bias of monthly HCHO_{MRM} against monthly HCHO_{OMI} . The MAE values ranges from 0.58 to 1.26×10^{15} molecules cm^{-2} in Figure 5d. Figures 5 e and 5f present RMSE and percent difference obtained from the linear regressions between monthly HCHO_{OMI} and monthly HCHO_{MRM} . The RMSE and percent difference are found to range from 0.75 to 1.64×10^{15} molecules cm^{-2} and from 7.1 to 9.9%, respectively, showing good performances of MRM. The statistical values between monthly HCHO_{OMI} in 2005 and 2006 against monthly HCHO_{MRM} are found to be either similar or slightly different from those between monthly HCHO_{OMI} in 2007 against monthly HCHO_{VAL} . The correlation coefficient (R) and the slopes are found to be up to 0.96 and 0.91, respectively, showing good agreements between monthly HCHO_{OMI} and monthly HCHO_{VAL} . In terms of east-central China, we found R (0.81) and slope (0.50) between monthly HCHO_{OMI} in 2007 against monthly HCHO_{VAL} lower than those between monthly HCHO_{OMI} in 2005 and 2006 against monthly HCHO_{MRM} . However, we found almost negligible differences in other statistical values (mean bias, MAE, RMSE, percent difference) for all of the study areas. The mean bias values are found to be either close to or smaller than zero in Figure 5c, implying negligible bias of monthly HCHO_{VAL} against monthly HCHO_{OMI} . The MAE values range from 0.77 to 1.54×10^{15} molecules cm^{-2} in Figure 5d. Figures 5 e and 5f present RMSE and percent difference obtained from the linear regressions between monthly HCHO_{OMI} and monthly HCHO_{VAL} . The RMSE and percent difference range from 0.96 to 2.15×10^{15} molecules cm^{-2} and 7.4 to 14.1%, respectively, showing similar performances of monthly HCHO_{VAL} to those of monthly HCHO_{MRM} . For Tokyo, most of the statistical values except for mean bias between monthly HCHO_{OMI} and monthly HCHO_{VAL} are better than those between monthly HCHO_{OMI} and monthly HCHO_{MRM} .

3.2. Estimation of Daily HCHO Column Density

We also evaluated the performance of the MREs in Table 3 by comparing daily HCHO estimates using the MREs and those using an interpolation method. Table 3 shows the MREs with regression coefficients determined using daily data of dependent and independent variables for 728 days period from 2005 to 2006. Table 3 shows the MREs with regression coefficients determined using daily data of dependent and independent variables for 728 days period from 2005 to 2006. In BPA, $\text{FRP}_{\text{MODIS}}$ and $\text{LST}_{\text{MODIS}}$ are excluded in the MRE for South Borneo due to high p -values (0.576 and 0.166, respectively) while the least square fitting is optimized when all of indepen-

dent variables are included in the MRE in Indochina. In APA, all of independent variables are selected in the final form of MREs in east central China, Beijing, and Seoul. However, in Tokyo, $\text{LST}_{\text{MODIS}}$ and FFC_{NO_2} are excluded in the MRE due to high p -values, which are 0.413 and 0.252, respectively. The R between daily HCHO_{OMI} and daily HCHO_{MRM} is generally smaller than that between monthly HCHO_{OMI} and monthly HCHO_{MRM} . It may be explained by the use of 8 and 16 days data for $\text{FRP}_{\text{MODIS}}$ and $\text{EVI}_{\text{MODIS}}$, respectively in the estimation of daily HCHO_{MRM} . We then estimated daily HCHO (daily HCHO_{VAL}) using the MREs in Table 3 with the inputs of daily independent variable data for 364 days in 2007. A simple interpolation method is also used to estimate HCHO column density (daily $\text{HCHO}_{\text{Interpol}}$) for 364 days in 2007. In this present study, twenty five percent of the total number of pixels within each target area is randomly removed and linearly interpolated with the value of near pixels for each area (Kythe and Wei, 2011).

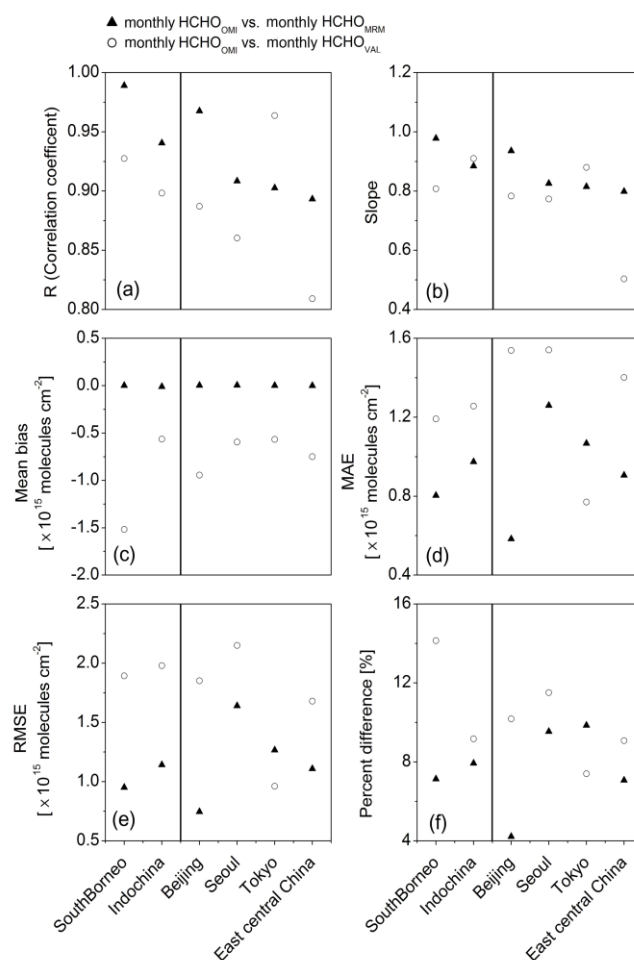


Figure 8. (a) R , (b) slope, (c) mean bias, (d) mean absolute error (MAE), (e) the root mean square error (RMSE), and (f) percent difference between monthly HCHO_{OMI} against HCHO_{MRM} and HCHO_{VAL} . Solid triangles represent the values for the regression between monthly HCHO_{OMI} in 2005 and 2006 and HCHO_{MRM} . Open circles represent those between monthly HCHO_{OMI} and HCHO_{VAL} in 2007.

Table 3. Final Form of the MRE and Correlation Coefficient (R) Obtained from the Linear Regression between Daily HCHO_{MRM} and Daily HCHO_{OMI} for 728 Days Period from 2005 through 2006

Type	Area	VIF*	p-value**	Multiple regression equation [$\times 10^{15}$ molecules cm^{-2}]***	Comparison with daily HCHO _{OMI} R****
BPA	South Borneo (4° S - 0° N, 108° - 116° E)	–	FRP (0.576) LST (0.166)	$4.59 \times \text{BB}_{\text{NO}_2} - 32.80 \times \text{EVI} + 10.61$	0.87
	Indochina (18° - 22° N, 98° - 102° E)	FRP (14.8)		$1.92 \times \text{BB}_{\text{NO}_2} + 0.81 \times 1.04^{\text{LST}} + 14.98 \times \text{EVI} + 0.09 \times \text{FRP} - 5.14$	0.71
APA	East-central China (36° - 40° N, 113° - 119.75° E)	–		$0.09 \times \text{FFC}_{\text{NO}_2} + 6.41 \times 1.01^{\text{LST}} + 9.69 \times \text{EVI} + 0.04$	0.55
	Beijing (39.5° - 40.5° N, 115.5° - 117.5° E)	–		$0.15 \times \text{FFC}_{\text{NO}_2} + 0.06 \times \text{LST} + 21.2 \times \text{EVI} + 5.78$	0.55
	Seoul (37° - 38° N, 126.5° - 127.5° E)	–		$0.07 \times \text{FFC}_{\text{NO}_2} + 2.07 \times 1.02^{\text{LST}} + 22.37 \times \text{EVI} + 5.26$	0.51
	Tokyo (35.5° - 36.5° N, 139° - 140° E)	–	LST (0.413) FFC _{NO2} (0.252)	$34.90 \times \text{EVI} + 3.27$	0.47

* The independent variable candidates that do not satisfy the VIF criterion ($\text{VIF} < 10$) are indicated.

** The independent variable candidates that do not satisfy the p -value criterion ($p\text{-value} < 0.05$) are indicated.

*** FRP is excluded from independent variable candidates for MREs in APA.

**** Correlation coefficient between daily HCHO_{OMI} and HCHO_{MRM}.

Figure 9 presents the (a) R , (b) slope, (c) mean bias, (d), (e) RMSE, and (f) percent difference between daily HCHO_{OMI} in 2007 and daily HCHO_{Interpol}, and those between daily HCHO_{OMI} in 2007 and daily HCHO_{VAL}. There are generally better MAE, RMSE, and percent difference between daily HCHO_{OMI} and daily HCHO_{VAL} than those between daily HCHO_{OMI} and daily HCHO_{Interpol}. Especially, in South Borneo, Indochina, Beijing and east-central China, the average MAE, RMSE, and percent difference between daily HCHO_{OMI} and daily HCHO_{VAL} are 2.78×10^{15} molecules cm^{-2} , 3.76×10^{15} molecules cm^{-2} , and 21.1%. These statistical values show the good performance of the MREs, which is similar with those between monthly HCHO_{OMI} and monthly HCHO_{VAL}. However, a simple interpolation method does not show a good performance of the HCHO estimates as the average MAE, RMSE, and percent difference between daily HCHO_{OMI} and daily HCHO_{Interpol} are 5.41×10^{15} molecules cm^{-2} , 7.62×10^{15} molecules cm^{-2} , and 53.9% in South Borneo, Indochina, Beijing and east-central China. In terms of mean bias, the mean biases between daily HCHO_{OMI} and daily HCHO_{Interpol} are close to zero since the data points in scatter plots between daily HCHO_{OMI} and daily HCHO_{Interpol} are symmetrically and widely scattered in most of regions of interest. Thus, the mean biases between daily HCHO_{OMI} and daily HCHO_{Interpol} are small, there are poor agreements between daily HCHO_{OMI} and daily HCHO_{Interpol}. In Seoul and Tokyo, We found that the average MAE, RMSE, and percent difference between daily HCHO_{OMI} and daily HCHO_{VAL} are 3.65×10^{15} molecules cm^{-2} , 5.45×10^{15} molecules cm^{-2} , and 29.9% whereas those between daily HCHO_{OMI} and daily HCHO_{Interpol} are 5.00×10^{15} molecules cm^{-2} , 7.10×10^{15} molecules cm^{-2} , and 48.6%. Although daily HCHO_{VAL} shows the good agreements with daily HCHO_{OMI} than daily HCHO_{Interpol}, the performance of both the MRM and interpolation method for HCHO estimates is not good enough to provide the HCHO data.

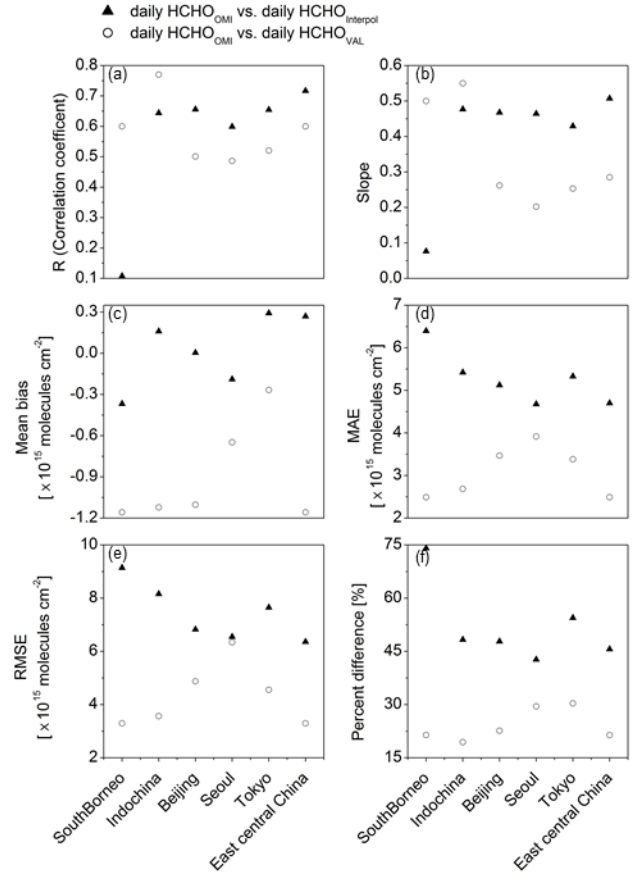


Figure 9. (a) R , (b) slope, (c) mean bias, (d) MAE, (e) RMSE, and (f) percent difference between HCHO_{OMI} against daily HCHO_{Interpol} and HCHO_{VAL}. Solid triangles represent the values for the regression between daily HCHO_{OMI} in 2007 and HCHO_{Interpol}. Open circles represent those between daily HCHO_{OMI} in 2007 and HCHO_{VAL}.

4. Conclusions

For the first time, we have estimated HCHO VCD in East Asia and Southeast Asia using a MRM with OMI and MODIS data. There is better agreement between monthly HCHO_{OMI} and monthly HCHO_{MRM} than that between daily HCHO_{OMI} and HCHO_{MRM} in the BPA and the APA. The agreement between monthly HCHO_{OMI} and monthly HCHO_{MRM} was compared with that between monthly HCHO_{OMI} and monthly HCHO_{VAL}. In addition, the agreement between daily HCHO_{OMI} and daily HCHO_{MRM} was compared with that between daily HCHO_{OMI} and HCHO_{Interpol}.

The statistical values (*R*, slope, mean bias, MAE, RMSE, and percent difference) indicate that the agreement between monthly HCHO_{OMI} and monthly HCHO_{MRM} is generally comparable to that between monthly HCHO_{OMI} and monthly daily HCHO_{VAL}. There are generally better MAE, RMSE, and percent difference between daily HCHO_{OMI} and daily HCHO_{VAL} than those between daily HCHO_{OMI} and daily HCHO_{Interpol}. However, a simple interpolation method does not show a good performance of the HCHO estimates. Although daily HCHO_{VAL} shows the good agreements with daily HCHO_{OMI} than daily HCHO_{Interpol}, the performance of both the MRM and interpolation method for HCHO estimates is not good enough to provide the HCHO data.

In the present study, we found that estimation of HCHO using MRM can be useful in areas where the relationship between HCHO and the independent variables is strong, and the variables have explicit seasonal cycles. The applicable areas are limited and the temporal resolution poorer than a CTM such as GEOS-Chem, which usually provides global coverage with high temporal resolution. However, as part of the efforts to provide HCHO data, MRM can be used to produce HCHO over the areas with poor measurement conditions such as poor signal to noise ratio at UV at high solar zenith angles which reduce the availability of satellite HCHO data. Similar studies can be done to produce HCHO data for other satellite sensors (e.g., SCIAMACHY, GOME-2). For future researches, the HCHO estimation can be attempted in higher time resolution with a variety of environmental variables as inputs for various statistical models.

Acknowledgement. This work was supported by the Pukyong National University Research Fund in 2013 (C-D-2013-1345).

References

- Abdul-Wahab, S.A., Bakheit, C.S., and Al-Alawi, S.M. (2005). Principal component and multiple regression analysis in modelling of ground-level ozone and factors affecting its concentrations. *Environ. Modell. Softw.*, 20(10), 1263-1271. <http://dx.doi.org/10.1016/j.envsoft.2004.09.001>
- Ahmad, S.S., Biiker, P., Emberson, L., and Shabbir, R. (2011). Monitoring nitrogen dioxide levels in urban areas in Rawalpindi, Pakistan. *Water Air Soil Poll.*, 220(1-4), 141-150. <http://dx.doi.org/10.1007/s11270-011-0741-9>
- Andreae, M.O. and Merlet, P. (2001). Emission of trace gases and aerosols from biomass burning. *Global Biogeochem. Cy.*, 15(4), 955-966. <http://dx.doi.org/10.1029/2000GB001382>
- Arlander, D., Brüning, D., Schmidt, U., and Ehhalt, D. (1995). The tropospheric distribution of formaldehyde during TROPOZ II. *J. Atmos. Chem.*, 22(3), 251-269. <http://dx.doi.org/10.1007/BF00696637>
- Baek, K.H., Kim, J.H., Park, R.J., Chance, K., and Kurosu, T.P. (2014). Validation of OMI HCHO data and its analysis over Asia. *Sci. Total Environ.*, 490, 93-105. <http://dx.doi.org/10.1016/j.scitotenv.2014.04.108>
- Barkley, M.P., Palmer, P.I., De Smedt, I., Karl, T., Guenther, A., and Van Roozendael, M. (2009). Regulated large-scale annual shutdown of Amazonian isoprene emissions? *Geophys. Res. Lett.*, 36(4). <http://dx.doi.org/10.1029/2008GL036843>
- Bey, I., Jacob, D.J., Yantosca, R.M., et al. (2001). Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation. *J. Geophys. Res. Atmos.*, 106(D19), 23073-23095. <http://dx.doi.org/10.1029/2001JD000807>
- Boeke, N.L., Marshall, J.D., Alvarez, S., et al. (2011). Formaldehyde columns from the Ozone Monitoring Instrument: Urban versus background levels and evaluation using aircraft data and a global model. *J. Geophys. Res. Atmos.*, 116(D6). <http://dx.doi.org/10.1029/2010JD014870>
- Chance, K. (2002). OMI algorithm theoretical basis document, volume IV: OMI trace gas algorithms. *Accessed*, 12(10), 2009.
- Cuiping, L., Chuangzhi, W., and Haitao, H. (2004). Study on the distribution and quantity of biomass residues resource in China. *Biomass Bioenergy*, 27(2), 111-117. <http://dx.doi.org/10.1016/j.biombioe.2003.10.009>
- De Smedt, I., Müller, J.F., Stavrou, T., van der A, R., Eskes, H., and Van Roozendael, M. (2008). Twelve years of global observations of formaldehyde in the troposphere using GOME and SCIAMACHY sensors. *Atmos. Chem. Phys.*, 8, 4947-4963. <http://dx.doi.org/10.5194/acp-8-4947-2008>
- De Smedt, I., Van Roozendael, M., Stavrou, T., Müller, J.-F., Chance, K., and Kurosu, T. (2012). Intercomparison of 4 Years of Global Formaldehyde Observations from the GOME-2 and OMI sensors, *ESA-ATMOS 2012 Atmospheric Science Conference*, edited by L. Ouwehand, Eur. Space Agency, Bruges, Belgium.
- Freeborn, P.H., Wooster, M.J., Roy, D.P., and Cochrane, M.A. (2014). Quantification of MODIS fire radiative power (FRP) measurement uncertainty for use in satellite-based active fire characterization and biomass burning estimation. *Geophys. Res. Lett.*, 41(6), 1988-1994. <http://dx.doi.org/10.1002/2013GL059086>
- Fu, Tzung May, Daniel J. Jacob, Paul I. Palmer, Kelly Chance, Yuxuan X. Wang, Barbara Barletta, Donald R. Blake, Jenny C. Stanton, and Michael J. Pilling. (2007). Space-based formaldehyde measurements as constraints on volatile organic compound emissions in east and south Asia and implications for ozone. *J. Geophys. Res. Atmos.*, 112(D6). <http://dx.doi.org/10.1029/2006JD007853>
- Giglio, L., Csiszar, I., and Justice, C.O. (2006). Global distribution and seasonality of active fires as observed with the Terra and Aqua Moderate Resolution Imaging Spectroradiometer (MODIS) sensors. *J. Geophys. Res. Biogeo.*, 111(G2). <http://dx.doi.org/10.1029/2005JG000142>
- Gonzi, S., Palmer, P.I., Barkley, M.P., De Smedt, I., and Van Roozendael, M. (2011). Biomass burning emission estimates inferred from satellite column measurements of HCHO: Sensitivity to co-emitted aerosol and injection height. *Geophys. Res. Lett.*, 38(14). <http://dx.doi.org/10.1029/2011GL047890>
- Guenther, A., Geron, C., Pierce, T., Lamb, B., Harley, P., and Fall, R. (2000). Natural emissions of non-methane volatile organic compounds, carbon monoxide, and oxides of nitrogen from North America. *Atmos. Environ.*, 34(12), 2205-2230. [http://dx.doi.org/10.1016/S1352-2310\(99\)00465-3](http://dx.doi.org/10.1016/S1352-2310(99)00465-3)
- Gupta, P. and Christopher, S.A. (2009). Particulate matter air quality assessment using integrated surface, satellite, and meteorological products: Multiple regression approach. *J. Geophys. Res. Atmos.*, 114(D14). <http://dx.doi.org/10.1029/2008JD011496>

- Irie, H., Boersma, K., Kanaya, Y., Takashima, H., Pan, X., and Wang, Z. (2012). Quantitative bias estimates for tropospheric NO₂ columns retrieved from SCIAMACHY, OMI, and GOM-E-2 using a common standard for East Asia. *Atmos. Meas. Tech.*, 5(10), 2403-2411. <http://dx.doi.org/10.5194/amt-5-2403-2012>
- Kim, W., Lee, H., Kim, J., Jeong, U., and Kweon, J. (2012). Estimation of seasonal diurnal variations in primary and secondary organic carbon concentrations in the urban atmosphere: EC tracer and multiple regression approaches. *Atmos. Environ.*, 56, 101-108. <http://dx.doi.org/10.1016/j.atmosenv.2012.03.076>
- Kutner, M.H., Nachtsheim, C., and Neter, J. (2004). *Applied linear regression models*, McGraw-Hill/Irwin.
- Kythe, P. and Wei, D. (2011). *An introduction to linear and nonlinear finite element analysis: A computational approach*, Springer Science & Business Media.
- Langner, A., Miettinen, J., and Siegert, F. (2007). Land cover change 2002-2005 in Borneo and the role of fire derived from MODIS imagery. *Glob. Change Biol.*, 13(11), 2329-2340. <http://dx.doi.org/10.1111/j.1365-2486.2007.01442.x>
- Laothawornkitkul, J., Taylor, J.E., Paul, N.D., and Hewitt, C.N. (2009). Biogenic volatile organic compounds in the Earth system. *New Phytol.*, 183(1), 27-51. <http://dx.doi.org/10.1111/j.1469-8137.2009.02859.x>
- Miura, T., Huete, A.R., and Yoshioka, H. (2000). Evaluation of sensor calibration uncertainties on vegetation indices for MODIS. *IEEE Trans. Geos. Remote Sens.*, 38(3), 1399-1409. <http://dx.doi.org/10.1109/36.843034>
- Nilsson, J. A., X. Zheng, K. Sundqvist, Y. Liu, L. Atzori, A. Elfving, K. Arvidson, and R. C. Grafström. (1998). Toxicity of formaldehyde to human oral fibroblasts and epithelial cells: influences of culture conditions and role of thiol status. *J. Dent. Res.*, 77(11), 1896-1903. <http://dx.doi.org/10.1177/00220345980770110601>
- Olivier, J., Peters, J., Granier, C., Petron, G., Müller, J. F., & Wallens, S. (2003). Present and future surface emissions of atmospheric compounds. *POET Rep. 2, EU Proj. EVK2-1999*, 11.
- Palmer, P.I., Abbot, D.S., Fu, T.M., Jacob, D.J., Chance, K., Kurosu, T.P., Guenther, A., Wiedinmyer, C., Stanton, J.C., Pilling, M.J. and Pressley, S.N. (2006). Quantifying the seasonal and interannual variability of North American isoprene emissions using satellite observations of the formaldehyde column. *J. Geophys. Res. Atmos.*, 111(D12). <http://dx.doi.org/10.1029/2005JD006689>
- Palmer, P.I., Jacob, D.J., Fiore, A.M., Martin, R.V., Chance, K., and Kurosu, T.P. (2003). Mapping isoprene emissions over North America using formaldehyde column observations from space. *J. Geophys. Res. Atmos.*, 108(D6). <http://dx.doi.org/10.1029/2002JD002153>
- Pang, X., Mu, Y., Zhang, Y., Lee, X., and Yuan, J. (2009). Contribution of isoprene to formaldehyde and ozone formation based on its oxidation products measurement in Beijing, China. *Atmos. Environ.*, 43(13), 2142-2147. <http://dx.doi.org/10.1016/j.atmosenv.2009.01.022>
- Peters, E., Wittrock, F., Großmann, K., Frieß, U., Richter, A., and Burrows, J. (2012). Formaldehyde and nitrogen dioxide over the remote western Pacific Ocean: SCIAMACHY and GOME-2 validation using ship-based MAX-DOAS observations. *Atmos. Chem. Phys.*, 12(22), 11179-11197. <http://dx.doi.org/10.5194/acp-12-11179-2012>
- Russell, A., Valin, L., and Cohen, R. (2012). Trends in OMI NO₂ observations over the United States: effects of emission control technology and the economic recession. *Atmos. Chem. Phys.*, 12(24), 12197-12209. <http://dx.doi.org/10.5194/acp-12-12197-2012>
- Sellke, T., Bayarri, M., and Berger, J.O. (2001). Calibration of p values for testing precise null hypotheses. *Am. Stat.*, 55(1), 62-71. <http://dx.doi.org/10.1198/000313001300339950>
- Sims, D.A., Rahman, A.F., Cordova, V.D., El-Masri, B.Z., Baldocchi, D.D., Bolstad, P.V., Flanagan, L.B., Goldstein, A.H., Hollinger, D.Y., Misson, L. and Monson, R.K. (2008). A new model of gross primary productivity for North American ecosystems based solely on the enhanced vegetation index and land surface temperature from MODIS. *Remote Sens. Environ.*, 112(4), 1633-1646. <http://dx.doi.org/10.1016/j.rse.2007.08.004>
- Stavrakou, T., Müller, J.F., Smedt, I.D., Roozendaal, M.V., Van Der Werf, G.R., Giglio, L. and Guenther, A. (2009a). Evaluating the performance of pyrogenic and biogenic emission inventories against one decade of space-based formaldehyde columns. *Atmos. Chem. Phys.*, 9(3), 1037-1060. <http://dx.doi.org/10.5194/acp-91037-2009>
- Stavrakou, T., Müller, J.F., Smedt, I.D., Roozendaal, M.V., Van Der Werf, G.R., Giglio, L. and Guenther, A. (2009b). Global emissions of non-methane hydrocarbons deduced from SCIA-MACHY formaldehyde columns through 2003-2006. *Atmos. Chem. Phys.*, 9(11), 3663-3679. <http://dx.doi.org/10.5194/acp-9-3663-2009>
- Steck, T., Glatthor, N., Clarmann, T.V., Fischer, H., Flaud, J.M., Funke, B., Grabowski, U., Höpfner, M., Kellmann, S., Linden, A. and Perrin, A. (2008). Retrieval of global upper tropospheric and stratospheric formaldehyde (H₂CO) distributions from high-resolution MIPAS-Envisat spectra. *Atmos. Chem. Phys.*, 8(3), 463-470. <http://dx.doi.org/10.5194/acp-8-463-2008>
- Steinbacher, M., Dommen, J., Ordonez, C., Reimann, S., Grüber, F.C., Staehelin, J. and Prevot, A.S.H. (2005). Volatile organic compounds in the Po Basin. part A: Anthropogenic VOCs. *J. Atmos. Chem.*, 51(3), 271-291. <http://dx.doi.org/10.1007/s10874-0053576-1>
- Timm, N.H. (2002). *Applied multivariate analysis (Vol. 26)*, Springer New York. <http://dx.doi.org/10.1007/b98963>
- M. Trainer, E. J. Williams, D. D. Parrish, M. P. Buhr, E. J. Allwine, H. H. Westberg, F. C. Fehsenfeld and S. C. Liu. (1987). Models and observations of the impact of natural hydrocarbons on rural ozone. *Nature*, 329, 705-707. <http://dx.doi.org/10.1038/329705a0>
- Wan, Z. (1999). *MODIS land-surface temperature algorithm theoretical basis document (LST ATBD)*. Institute for Computational Earth System Science, Santa Barbara, 75.
- Zhang, X., Friedl, M.A., Schaaf, C.B., and Strahler, A.H. (2004). Climate controls on vegetation phenological patterns in northern mid-and high latitudes inferred from MODIS data. *Glob. Change Biol.*, 10(7), 1133-1145. <http://dx.doi.org/10.1111/j.1529-8817.2003.00784.x>