First MAX-DOAS observations of atmospheric formaldehyde and glyoxal concentrations at Southeast and South Asian sites

MAX-DOAS 法による東南アジア及び南アジアにおける大気中のホルムアルデヒドとグリオキサール濃度の初観測

January 2019

Hossain Mohammed Syedul Hoque
Graduate School of
Advanced Integration Science
CHIBA UNIVERSITY

（千葉大学審査学位論文）

First MAX-DOAS observations of atmospheric formaldehyde and glyoxal concentrations at Southeast and South Asian sites

MAX-DOAS 法による東南アジア及び南アジアにおける大気中のホルムアルデヒドとグリオキサール濃度の初観測

January 2019

Hossain Mohammed Syedul Hoque
Graduate School of
Advanced Integration Science
CHIBA UNIVERSITY

Abstract

Volatile organic compounds (VOC) exhort significant changes in the atmosphere. Although biogenic emissions are the largest source of VOCs, significant contributions are also attributed to biomass burning and anthropogenic emissions. Formaldehyde (HCHO) and glyoxal (CHOCHO) are two oxidized products of VOCs and are important tracers of VOCs in the atmosphere. HCHO and CHOCHO are one the most abundant aldehyde and dicarbonyl compounds, respectively, in the atmosphere. The glyoxal to formaldehyde ratio (R_{GF}) has been used as an indicator of changes in the VOC emissions. However, the change of R_{GF} values under different VOC emission scenarios is not well understood. Disagreement remain among studies evaluating the R_{GF} response from various measurement platforms (i.e., ground based, satellite). This study focuses on HCHO and CHOCHO observations using the multi-axis differential optical absorption spectroscopy (MAX-DOAS) technique at two SKYNET sites in Thailand and India and the changes of R_{GF} values under different VOC emission sources are evaluated.

A MAX-DOAS instrument was installed at Phimai (15.18 °N, 102.56 °E), a rural site in central Thailand and the continuous observations of aerosols and trace gases started in September 2014. Here, simultaneous measurements of HCHO and CHOCHO from October 2014 to October 2016 are evaluated. During the dry season (January – April) enhanced HCHO and CHOCHO concentrations with peak around March were observed. Such enhanced concentrations were related to pronounced biomass burning events during this period. The enhancements of HCHO during the dry seasons were also observed in the satellite observations and model simulations. The annual mean R_{GF} estimated from the HCHO and CHOCHO observations at Phimai was 0.032 ± 0.005. The R_{GF} values were lower during the dry season (~0.03), whereas higher R_{GF} values (~0.04) were observed during the wet season. Such change of R_{GF} values was consistent with satellite observations and few field studies, suggesting higher R_{GF} for dominant biogenic VOC emissions. The lower R_{GF} in the dry season due to biomass burning was confirmed using nitrogen dioxide (NO_2) as a tracer. Evaluating the diurnal variation of R_{GF}, it was found that diurnal R_{GF} values in the wet season were higher than the dry season. Despite the diurnal variability, higher R_{GF} values under strong biogenic VOC emissions were observed. However, such response of R_{GF} was in contrast to some field studies reporting lower R_{GF} values for strong biogenic emission scenarios.

Using a similar MAX-DOAS instrument, continuous observations of HCHO and CHOCHO have been performed for the first time at Pantnagar (29.03° N, 79.47° E), a semi-rural site in the Indo Gangetic Plain (IGP) region in India. The HCHO and CHOCHO concentrations were high during spring and autumn due presumably to biomass burning. From the whole observation period (January – November) in Pantnagar the mean R_{GF} was estimated to be 0.029 ± 0.006. The mean R_{GF} values at Pantnagar were mostly lower than 0.04 throughout the year. This was potentially due to the combined influence of pyrogenic and anthropogenic emissions at Pantnagar. The R_{GF} values in Pantnagar were compared with the R_{GF} values at Phimai during the dry season. Despite being two different sites with different climatic conditions, mean R_{GF} value was less than 0.04 under the influence of biomass burning and anthropogenic emissions. This finding was consistent with few studies based on field observations and satellite measurements.

Table of Content

Chapter 1: Scientific Background

Chapter 2: Absorption Spectroscopy

Chapter 3: MAX-DOAS observations of formaldehyde and glyoxal in Phimai, Thailand

Chapter 4: MAX-DOAS observations of formaldehyde and glyoxal in Pantnagar, India

Chapter 5: Comparison with satellite measurements and model

Chapter 1: Scientific Background

The information provided in this chapter are based on the textbooks by *Wayne* [2000], *Platz and Stutz* [2008], and Progress and problems in atmospheric chemistry, Volume 3. Additionally, information was also taken from PhD thesis of *Alvarado* [2015], *Peters* [2013], and *Wittrock* [2006]. Information in section 1.3.1 and 1.3.2 were mostly taken from www.inchem.org .

1.1 The formation of our atmosphere

The atmosphere of the Earth, also commonly known as air, is a narrow structured gaseous layer enclosing the Earth and retained by the Earth's gravitational force. The evolution of today's atmosphere started around 4 billion years ago, which was devoid of oxygen. The primitive atmosphere was formed mainly due to volcanic activities, as scientists believe that intensive volcanic activities occurred during the first billion years of the Earth's evolution. The main gases emitted from such volcanic activities were water vapor and carbon dioxide. After the earth started cooling down, the water vapor condensed and formed the water bodies (*i.e.*, sea, ocean, lakes etc.). Oceans started to absorb the large amount of carbon dioxides. Microorganisms in the ocean converted the absorbed carbon dioxides to oxygen through photosynthesis process. Oxygen as a byproduct of such biological processes started to change the composition of the atmosphere, allowing life to evolve. The oxygen concentration also facilitated the formation of the ozone layer, which prevents most of the harmful ultraviolet sunlight from reaching the surface. The composition of the current atmosphere is given in Table 1.1

Table 1.1: The man constituents of the atmosphere. (*Platz and Stutz* [2008])

Gas	Chemical formula	Mixing ratio (%)
Nitrogen	N_2	78.08
Oxygen	O_2	20.95
Argon	Ar	0.93
Carbon dioxide	CO_2	0.037
Water vapor	H_2O	0-4
Neon	Ne	0.0018
Helium	He	0.00052
Methane	CH_4	0.00017
Krypton	Kr	0.00011
Xenon	Xe	0.00009
Hydrogen	H_2	0.00005
Dinitrogen oxide	N_2O	0.00003

Nitrogen, oxygen, and argon are the three most abundant constituents which constitute 99.9% of the air. Water vapor, carbon dioxide and other noble gases also have comparable abundances in the atmosphere. The constituent of the atmosphere is regarded fairly stable in terms of its main components. But atmospheric chemistry is mainly concerned with "trace elements/gases" of the atmosphere. Trace elements are species at much lower abundances in the atmosphere such as hydrogen, methane, ozone, carbonyl sulfide, methyl chloride etc. Despite their low abundances, trace gases have noticeable impact on the atmosphere.

Besides, aerosols are also another important component of the atmosphere. Aerosols are not gaseous elements but suspended particles in the atmosphere. Aerosols can impact the atmosphere in various ways including heating or cooling of the atmosphere, impact on cloud formation and other atmospheric phenomenon. The impact of aerosols on the atmospheric radiation budget is still significantly uncertain due to their high spatial and temporal variation.

1.2 The atmospheric layers

The earth's atmosphere can be divided into four primary regions depending on the vertical temperature change of the atmosphere. These regions are the troposphere, the stratosphere, the mesosphere, and the ionosphere. The troposphere is the lowest layer of the atmosphere which is relatively well mixed and compromises the majority (~80%) of the atmosphere's mass. The altitude of the troposphere depends on latitude and season and can prolong from the surface to 8-18 km. The temperature in the troposphere decreases with altitude which allows the hot air masses to rise and convection processes lead to the vertical mixing of the air masses in the troposphere. Moreover, due to evaporation from the water bodies, the troposphere is dominated by water vapor, which contributes to the formation of clouds. The troposphere can be further divided into two parts. Firstly, the atmospheric boundary layer (BL). Secondly, the free troposphere, extending from the top of BL to the tropopause. The BL is the lowest layer of the troposphere and is strongly influenced by the surface. Mixing and movement of air masses in the BL are strongly influenced by friction on the surface. The height of the BL changes with temperature. Over the land, the BL varies from a few hundred meters at night to around 2 km or more during the day. Aerosols and trace gases emitted from different sources are mostly concentrated in the boundary layer due to strong mixing.

The stratosphere extends from the tropopause (~ 10-15 km) to about 50 km altitude and is a poorly mixed layer. The temperature gradient in the stratosphere is positive, which means the temperature increases with altitude in the stratosphere. This is due to the ultra-violet (UV) light absorption by ozone. The common subdivision of the stratosphere includes: the tropical tropopause layer, the lower and middle stratosphere, and the upper stratosphere.

The mesosphere extends from the tropopause (~ 45-55 km) to ~ 85 km. Due to the presence of less ozone, heating due to light absorption is less important in the mesosphere. The temperature in the mesosphere decreases with altitude (i.e., negative temperature gradient) and the vertical mixing is relatively rapid.

The thermosphere is the coldest region of the atmosphere, which is separated by the mesopause (~80- 90 km) from the mesosphere. This region is characterized by high temperature and a positive temperature gradient. This is mainly due to absorption of short-wave solar radiation by molecular nitrogen and oxygen. Dynamical effect such as winds, gravity waves, atmospheric tides causes rapid vertical mixing in the thermosphere. However, the composition of the atmosphere is fairly constant with respect to the abundant species, despite the vertical temperature and pressure variation.

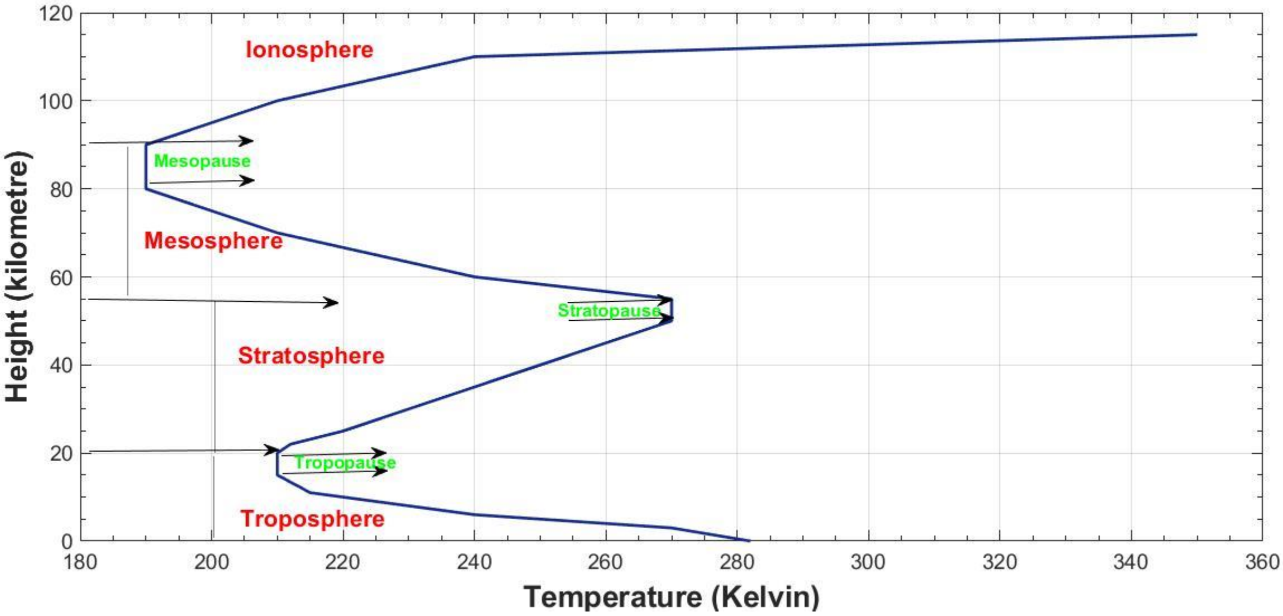


Figure 1.1: Temperature profile of the earth's atmosphere

1.3 Tropospheric Chemistry

(This section is based on the review article of *Atkinson and Arey* [2003])

Numerous organic compounds in the atmosphere take part in different atmospheric chemical reactions. Volatile organic compounds (VOCs) emitted from biogenic, anthropogenic, biomass burning and other sources alter the composition of the atmosphere. These include changing the oxidization capacity of the atmosphere, contribution in secondary organic aerosols (SOAs) and ozone formation, and influencing the production of cloud condensation nuclei [*Houweling et al.*, 1998; *Liakakou et al.*, 2007; *Monks,* 2005; *Poisson et al.*, 2000; *Volkamer et al.*, 2006*; Vrekoussis et al.,* 2007, 2010]. VOCs compromises non-methane hydrocarbons (NHMCs) and oxygenated NMHCs emitted from various sources such as oceans, vegetation, biomass burning, fossil fuel usage, and geochemical processes [*Alvarado et al*. 2014; *Kansal* 2009]. Although VOCs are emitted from different sources, the biogenic emissions are the strongest sources of VOCs. According to *Atkinson and Arey* [2003], globally the biogenic emissions of VOCs are higher than the anthropogenic emissions by a factor of 10. Annually, ~1150 TgC of biogenic VOCs are emitted worldwide [*Guenther et al* 1995]. The estimated emissions of anthropogenic VOCs are ~161-186 Tg C yr^{-1} [*Stavrakou et al*. 2009]. The pyrogenic contribution of VOC is estimated to be ~ 33-49 Tg C yr^{-1} [*Andreae and Marlet*, 2001]. Lastly, ~0.5-5 Tg C yr^{-1} of VOCs are emitted from oceans [*Alvarado*, 2016]. The potential loss and transformation processes of VOCs in the atmosphere are dry and wet deposition, photolysis, reaction with hydroxyl radical (OH), nitrate (NO_3), and ozone(O_3). Removal by reaction with chlorine (Cl) can also occur in marine environments [*Oum et al.*, 1998]. The dry and wet deposition removal processes might be less significant for biogenic VOCs, but could be important for long lived VOCs such as methanol and few biogenic VOC reaction products [*Atkinson and Arey*, 2003].

As biogenic sources are the dominant sources of global VOC concentrations, a bit more details about biogenic VOCs are discussed in this section.

The dominant form of biogenic VOC emitted to the atmosphere is isoprene and monoterpenes. According to *Guenther et al* [1995] vegetation contributes ~ 44% of the global VOC burden and isoprene is the single largest non-methane hydrocarbon emitted from vegetation to the atmosphere. A similar study later by *Guenther et al* [2000] reported that, the biogenic VOCs in North America compromises of 30% isoprene, 25% terpenoid compounds and the rest 40% are non-terpenoid compounds. Table 1.2 lists few biogenic VOCs and their atmospheric life time.

As seen from Table 1.2, most of the biogenic VOCs have atmospheric life time around few hours. Few biogenic VOC such as α-Terpinene, Terpinolene etc. has lifetime ~ 1 min. Which means, if these VOCs are emitted from vegetation, then they will be removed immediately from the atmosphere even at the “cleanest” region. This also indicates the high reactivity of biogenic VOCs and the challenges during their detection.

When VOCs are emitted from vegetation or other sources, due to their high reactivity it breaks into multiple intermediate products through different branching chemical reactions. For example, both formaldehyde and glyoxal are intermediate products of isoprene. Because numerous VOCs are emitted to the atmosphere, the intermediate products and the branching reactions are not clearly understood yet. But since the last few decades, multiple studies have been conducted to understand the VOC branching reactions in the atmosphere.

The initial reaction of VOCs primarily starts with the hydroxyl radical (OH), nitrate radicals (NO_3), and ozone (O_3). For the biogenic VOCs summarized in Table 1.2, in general there are two types of reaction mechanism.

(a) OH, NO_3, and O_3 adding to the C═C bonds
(b) Abstraction of the hydrogen (H) atom from the C─H bond by OH and NO_3

Table 1.2: Examples and lifetime of few biogenic VOCs (*Atkinson and Arey* [2003])

Biogenic VOCs	Lifetime for reaction with OH	Lifetime for reaction with NO_3	Lifetime for reaction with O_3
Isoprene	1.4 h	1.6 h	1.3 day
Monoterpenes			
Camphene	2.6 h	1.7 h	18 d
2-Carene	1.7 h	4 min	1.7 h
3-Carene	1.6 h	7 min	11 h
Limonene	49 min	5 min	2h
Myrcene	39 min	6 min	50 min
Cis/trans Ocimene	33 min	3 min	44 min
α- Phellandrene	27 min	~ 1 min	8 min
β- Phellandrene	50 min	8 min	8.4 h
α- Pinene	2.6 h	11 min	4.6 h
β- Pinene	1.8 h	27 min	1.1 d
Sabinene	1.2 h	7 min	4.8 h
α- Terpinene	23 min	0.5 min	1 min
γ- Terpinene	47 min	2 min	2.8 h
Terpinole	37 min	0.7 min	13 min

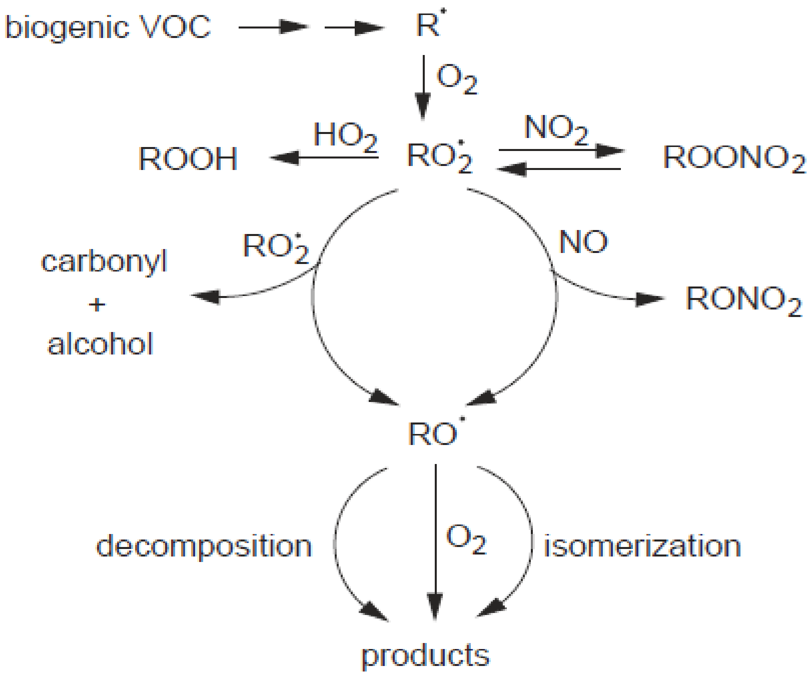
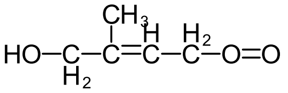
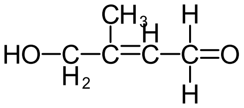
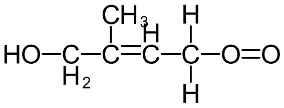


Figure 1.2: The basic reaction mechanism of biogenic VOCs in the atmosphere (Adapted from *Atkinson and Arey* [2003])

Biogenic VOCs compromising C═C bonds in generally prefer the addition of OH and NO_3 over the H atom abstraction. O_3 prefers only the addition to C═C bonds. The abstraction of the

H atom preferably by OH and NO_3 occurs in alkenes, alcohols, ethers, etc. which contain C─H bonds. However, this reaction scheme is not preferable for some VOCs containing C═C bonds such as isoprenes, monoterpenes, ethers etc. But this scheme is important for aldehydes containing C═C bonds. Figure 1.2 shows the general reaction scheme of biogenic VOCs emitted to the atmosphere.

The reaction of the emitted biogenic VOC is initiated by any either mechanism (a) or (b) mentioned above which forms the alkyl radical ($R^\bullet$). $R^\bullet$ can also be termed as substituted alkyl radical. Then the reaction proceeds by forming intermediate radicals called organic peroxy radical ($RO^\bullet_2$) and alkoxy radicals ($RO^\bullet$). From the formation of $RO^\bullet$, intermediate VOC products can be formed through isomerization, decomposition, and reaction with O_2. The pathway from RO to the intermediate products is determined by several factors, such as whether the $RO^\bullet$ has the ability to decompose or isomerize, the preferable pathway of the reaction between $RO_2^\bullet$ and NO, and the level of NO.

An example of the reaction scheme explained is shown in Fig 1.3, using isoprene as an example. The reaction of isoprene is initiated by OH. Following the reaction mechanism shown in Fig.1.2, the intermediate peroxy and alkoxy radicals are  $HO-\underset{H_2}{C}-\overset{CH_3}{C}=\overset{H}{C}-\overset{H_2}{C}-O=O$ and  $HO-\underset{H_2}{C}-\overset{CH_3}{C}=\overset{H}{C}-\underset{H}{\overset{H}{C}}=O$ respectively. From, the alkoxy radical, the intermediate VOC products are produced by isomerization and reaction with O_2. If the isomerization would not be possible in this case, then the intermediate product could be formed by decomposition and reaction with O_2. The reaction of isoprene shown in Fig. 1.3 is based on the condition that, the peroxy radical  $HO-\underset{H_2}{C}-\overset{CH_3}{C}=\overset{H}{C}-\underset{H}{\overset{H}{C}}-O=O$ dominantly reacts with NO. If the concentration of NO is low, then the reaction between the peroxy radical and HO_2 or the peroxy radical and peroxy radical will dominate (refer to Fig. 1.2). This will lead to different intermediate products from isoprene. Thus the intermediate products will depend on the reaction pathways of the primary VOC. Some intermediate products from OH initiated reaction of isoprene is shown in Fig 1.4. According to Fig. 1.4, formaldehyde is a first generation products of OH initiated reaction of isoprene, whereas glyoxal is the third generation product of isoprene reaction scheme with OH.

In case of biomass burning events, isoprene, formaldehyde, and glyoxal can be emitted directly from burning of biomass fuels [*Andreae and Merlet* 2003; *Fu et al* 2008]. Moreover, biomass burning plumes contains high amount of NO and NO_2 which can react with the peroxy radicals (refer to Fig.1.2) to form the intermediate products of VOC reaction schemes. For more detailed description of the VOC reaction in the atmosphere, readers are referred to the review article of *Atkinson and Arey* [2003] and the references therein by.

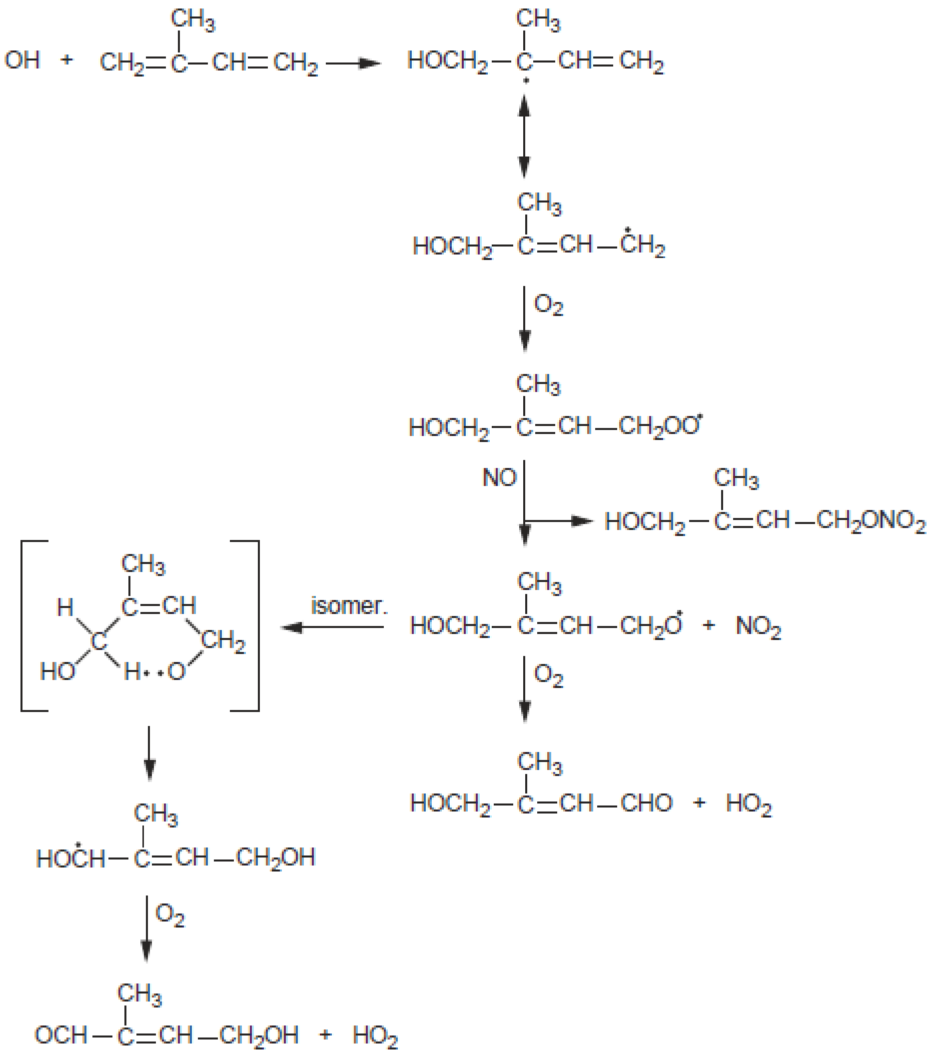


Figure 1.3: Example of OH initiated reaction of isoprene in the atmosphere (Adapted from *Atkinson and Arey* [2003])

The photo oxidation of VOCs in the presence of NO_x causes ozone (O_3) formation in the troposphere. The significant pathway of O_3 formation in the troposphere is the photolysis of nitrogen dioxide.

$$NO_2 + h\upsilon \rightarrow NO + O(^3P) \qquad (1.1)$$

$$O(^3P) + O_2 + M \rightarrow O_3 + M \text{ (M= any other molecule, usually } N_2 \text{ or } O_2\text{)} \qquad (1.2)$$

The photo oxidation of both biogenic and anthropogenic VOCs produce organic peroxy radicals (RO_2) and hydroperoxyl (HO_2) radicals, which react with nitrogen oxide (NO) to form NO_2

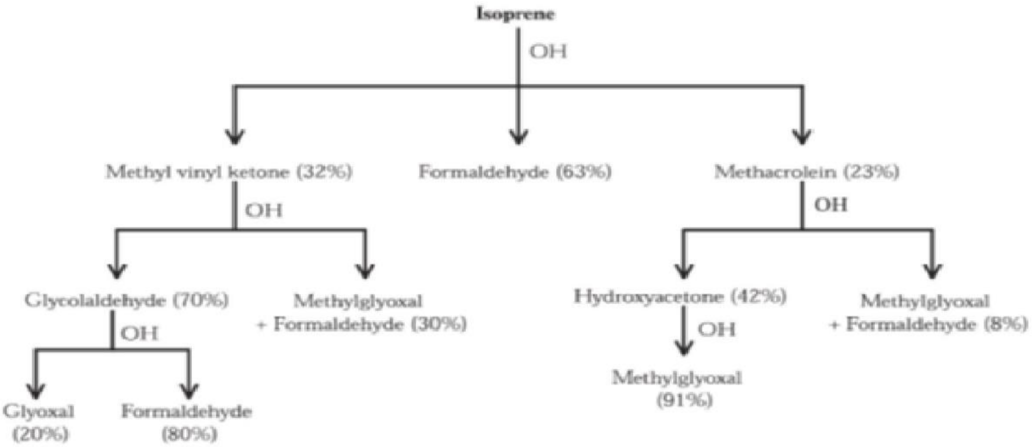


Figure 1.4: Intermediate products from OH initiated reaction of isoprene (Slightly modified from the original picture from *Wittrock* [2006]).

$$RO_2 + NO \rightarrow RO + NO_2 \qquad (1.3)$$

$$HO_2 + NO \rightarrow RO + NO_2 \qquad (1.4)$$

The photolysis of NO_2 then leads to the formation of O_3 through reactions (1.1) and (1.2). Photo oxidation of VOCs is a chain reaction process which is mainly initiated by hydroxyl radicals (OH). The following reactions leads to the formation of O_3 from NO and NO_2

$$VOC + OH \overset{O_2}{\rightarrow} RO_2 + H_2O \qquad (1.5)$$

$$RO_2 + NO \rightarrow NO_2 + RO \qquad (1.6)$$

$$RO + O_2 \rightarrow RCHO + HO_2 \qquad (1.7)$$

$$HO_2 + NO \rightarrow NO_2 + OH \qquad (1.8)$$

After addition of O_2, VOC reacts with OH to produce a RO_2. Similar to RO_2, HO_2 also reacts with NO to produce NO_2. OH is regenerated from the RO_2 product RO. The NO to NO_2 conversion path leading to net formation of O_3 allows the NO_2 to NO ratio to increase and O_3 to increase according to a photo stationary state relation.

1.3.1 Formaldehyde (HCHO)

Formaldehyde (HCHO) is also known as methanal, methylene oxide, oxymethylene, methyl aldehyde, and formic acid. It is produced in large quantities in industries and is used in variety of chemical manufacturing processes. The average rate of global production of HCHO from methane oxidation in the troposphere is 4×10^{11} kg/year and ~ 3.5×10^{9} kg/ year from industrial production. Moreover, HCHO is used as sterilant in medical applications and as preservatives in consumer products (i.e., foods, cosmetics, etc.). Formaldehyde is harmful to human health. Clinical studies and epidemiological surveys observed transient and reversible sensory irritation of the eyes and respiratory tracts in human bodies. In general eye irritation occurs when the airborne HCHO concentrations range from 0.3-0.5 ppmv. HCHO also causes skin irritation and corrosive properties when ingested. However, medical studies found no impact of HCHO on lung functions in either asthmatics or non-asthmatics.

In the atmosphere, HCHO is mainly produced from the photochemical degradation of methane (CH_4) and non-methane hydrocarbons (NMHC). The background HCHO concentration of 0.2 – 1.0 ppbv over remote marine environments are mainly due to methane oxidization [*Burkert et al.,*2001; *Singh et al.,* 2001,2004; *Sinreich et al.,* 2005; *Weller et al.,* 2000]. Additionally, HCHO can be directly emitted from biomass burning, industrial processes, fossil fuel combustion [*Andreae and Merlet*, 2001; *Geiger et al.*, 2002; *Holzinger et al.*, 1999; *Lee et al.*, 1997; *Yokelson et al.*, 1999], and vegetation [*Seco et al.*, 2007]. HCHO is removed from the atmosphere via photolysis at wavelengths below 400 nm, reaction with OH, and wet deposition. This limits the atmospheric lifetime of HCHO to a few hours [*Arlander et al.*, 1995].

The atmospheric HCHO concentration is mainly determined by the oxidation of CH_4 and NMHCs. HCHO is an indicator of photochemical activities and during the oxidation process, almost all VOCs form HCHO as an intermediate product [*Alvarado* 2016]. A simple oxidation process of CH_4 in the atmosphere is given below:

$$OH + CH_4 \rightarrow CH_3 + H_2O \quad (1.9)$$

$$CH_3 + O_2 + M \rightarrow CH_3O_2 + M \quad (1.10)$$

$$CH_3O_2 + NO \rightarrow NO_2 + CH_3O \quad (1.11)$$

$$CH_3O + O_2 \rightarrow HCHO + HO_2 \quad (1.12)$$

$$HCHO + h\upsilon \rightarrow H_2 + CO \quad (1.13)$$

$$HCHO + h\upsilon \rightarrow H + HCO \quad (1.14)$$

$$HCHO + OH \rightarrow HCHO + H_2O \quad (1.15)$$

$$HCO + O_2 \rightarrow HO_2 + CO \quad (1.16)$$

$$HO_2 + NO \rightarrow NO_2 + OH \quad (1.17)$$

Reaction (1.17) is similar to reaction (1.8), showing the role of VOCs in the formation of tropospheric O_3. The CO produced in reaction (16), can yield another HO_2.

$$OH + CO \rightarrow H + CO_2 \quad (1.18)$$

$$H + O_2 + M \rightarrow HO_2 + M \quad (1.19)$$

Thus monitoring of spatial and temporal variability of NMHCs are important for better understanding the process controlling the production and evolution of tropospheric ozone, OH, and SOA. [*DeSmedt et al* 2015].

1.3.2 Glyoxal (CHOCHO)

Glyoxal (CHOCHO) is also known as ethanedial, diformyl, ethanedione, bioformal, and oxal. Similar to HCHO, CHOCHO also has industrial applications such as cross-linking agent in the production of polymers. It is also an intermediate product in pharmaceuticals and dyestuffs. CHOCHO is also harmful to human health. It can cause irritation to eyes and skin and can be harmful by inhalation.

In the atmosphere, CHOCHO is primarily produced from the oxidation of biogenic VOCs such as isoprene and terpenes and additionally from VOCs emitted from anthropogenic sources such as acetylene, benzene, toluene, and xylene [*Fu et al.*, 2008; *Myriokefalitakis et al.*, 2008]. CHOCHO can also be directly emitted from biomass burning [*Fu et al.*, 2008]. Figure 1.4 shows the OH reaction scheme of isoprene. Both HCHO and CHOCHO are byproducts of isoprene. However, CHOCHO is the second or third generation product of the biogenic VOCs, whereas HCHO is the first-generation product. The removal processes of CHOCHO from the atmosphere are similar to those for HCHO, *i.e.*, photo disassociation [*Fu et al.*, 2008], interaction with OH radicals, and dry and wet depositions [*Myriokefalitakis et al.*, 2008]. Due to high solubility, CHOCHO can facilitate SOA formation. The reactions of CHOCHO removal from the atmosphere is as follows:

$$\mathrm{CHOCHO + h\upsilon \rightarrow 2HCO} \tag{1.20}$$

$$\mathrm{CHOCHO + h\upsilon \rightarrow H_2 + 2CO} \tag{1.21}$$

$$\mathrm{CHOCHO + h\upsilon \rightarrow H_2CO + CO} \tag{1.22}$$

$$\mathrm{CHOCHO + h\upsilon \rightarrow H + CO + HCO} \tag{1.23}$$

$$\mathrm{CHOCHO + OH \rightarrow HCOCHO + H_2} \tag{1.24}$$

$$\mathrm{HCOCO + O_2 \rightarrow 2CO + HO_2} \tag{1.25}$$

According to *Myriokefalitakis et al.* [2008], CHOCHO can be used as an indicator of the rate of photochemical VOC processes because CHOCHO is not produced from CH_4 oxidation as HCHO.

1.3.3 Glyoxal to formaldehyde concentration ratio (*R_{GF}*)

(Most of the text in this section are taken from the introduction section of *Hoque et al* [2018])

The atmospheric abundance of HCHO and CHOCHO can potentially reflect the speciation of VOCs to total VOC reactivity due to their similar atmospheric lifetime and different yields from different VOC classes [*Vrekousiss et al* 2010, *DiGangi et al* 2012; *MacDonald et al* 2012; *Li et al* 2013; *Miller et al* 2014; *Kaiser et al* 2015]. The CHOCHO to HCHO concentration ratio (R_{GF}) has been used as a potential tracer of changes in VOC emission sources due to their similar sinks and different sources. The R_{GF} values estimated from satellite observations ranged between 0.04 and 0.06 for regions characterized with strong biogenic emissions, whereas R_{GF} values were < 0.03 in regions influenced by anthropogenic emissions [*Vrekoussis et al.*, 2009, 2010]. From simultaneous multi-axis differential optical absorption spectroscopy

(MAX-DOAS) observations of CHOCHO and HCHO in Cabauw, a rural site in the Netherlands, *Irie et al*. [2011] reported an R_{GF} value of 0.036. *Ortega et al* [2015] conducted MAX-DOAS measurements at an urban site in Mainz Germany, where emissions are mostly anthropogenic and observed R_{GF} values lower than 0.04.

Munger et al. [1995] and *Grosjean et al*. [1996] performed in situ measurements of HCHO and CHOCHO at two rural sites in the U.S and found R_{GF} values of 0.045 and 0.132, respectively. A higher R_{GF} (0.06 ± 0.02) value was observed in a layer ~2.5 km compared with inside the boundary layer (0.027 ± 0.006) during the CalNex and CARES field campaign in 2010, where HCHO and CHOCHO observations were performed using airborne MAX-DOAS [*Baidar et al.* 2013]. During the BEARPEX 2009 and BEACHON-ROCS campaign in the *Pinus Ponderosa* forests, DiGangi *et al.* [2012] conducted simultaneous rapid, in-situ measurements of HCHO and CHOCHO and found higher R_{GF} under dominant anthropogenic emission scenario and biomass burning compared with dominant biogenic VOC emission cases. This result contradicts with the results from satellite observations suggesting a higher R_{GF} in areas with biogenic emissions. During the 2013 SENEX flight campaign over the southeastern U.S., *Kaiser et al.* [2015] found higher R_{GF} values over areas where monoterpene emissions were dominant, and lower R_{GF} values were observed over areas characterized by strong isoprene emissions. This is in agreement with the satellite observational results reported by *Miller et al.* [2014] but contradicts the results of *DiGangi et al.* [2012]. Thus, the R_{GF} response to different types of emissions is still under discussion.

The discrepancy observed in the R_{GF} response to different VOC emission sources among the ground-based and satellite observations is yet under discussion. Despite the discrepancies in the observed R_{GF} trends and absolute values, all the previous studies conclude that R_{GF} at least reflects the part of speciation of VOCs in the air masses sampled [*Kaiser et al* 2015].

1.4 The objective of the study

As seen from the discussion above, the response of R_{GF} to different VOC emission sources is yet under discussion. Moreover, the observations of HCHO and CHOCHO are also limited around the globe. This study primarily focuses on simultaneous MAX-DOAS observations of HCHO and CHOCHO in Thailand and India. These regions are poorly constrained in terms of VOC measurements and the new observations provide the first simultaneous datasets of the two important VOC tracers in these regions. From the continuous observations of HCHO and CHOCHO, the first objective is to estimate the R_{GF} for both sites and compare with literature values. Secondly, because the measurements of HCHO and CHOCHO covers almost all the seasons at both sites, the seasonal change of the R_{GF} values are assessed. As R_{GF} values reflect the VOC speciation, the change in the R_{GF} values is expected to reflect the changes in the VOC emission sources. The outcome of the study is expected to contribute to the ongoing discussion of the R_{GF} response, especially on a seasonal scale. Moreover, the new datasets can be useful in reducing model uncertainties regarding VOC chemistry in the study regions.

References

Alvarado, L. M. A., A. Richter, M. Vrekoussis, F. Wittrock, A. Hilboll, S. F. Schreier, and J. P. Burrows (2014), An improved glyoxal retrieval from OMI measurements, *Atmos. Meas. Tech*, *7*(12), 4133, doi:10.5194 /amt - 7 - 4 133 - 2014.

Alvarado, L. M. A. (2016), Investigating the role of glyoxal using satellite and MAX-DOAS measurements, PhD thesis, University of Bremen, Germany.

Andreae, M. O., and P. Merlet (2001), Emission of trace gases and aerosols from biomass burning, Global *Biogeochem. Cy.*, *15*(4), 955-966, doi:10.1029/2000GB001382.

Atkinson, R., and J. Arey (2003), Gas-phase tropospheric chemistry of biogenic volatile organic compounds: a review, *Atmos. Environ.* , *37*, 197-219, doi:10.1016 / S1352-2310 (03) 00391-1.

Baidar, S., H. Oetjen, S. Coburn, B. Dix, I. Ortega, R. Sinreich, and R. Volkamer (2013), The CU Airborne MAX-DOAS instrument: Vertical profiling of aerosol extinction and trace gases, *Atmos. Meas. Tech.*, *6*(3), 719-739, doi:10.5194/amt-6-719-2013.

Burkert, J., M. D. Andrés-Hernández, D. Stöbener, J. P. Burrows, M. Weissenmayer, and A. Kraus (2001), Peroxy radical and related trace gas measurements in the boundary layer above the Atlantic Ocean, *J. Geophys. Res.*, *106*(D6), 5457-5477, doi:10.1029/2000JD900613.

DiGangi, J., S. Henry, A. Kammrath, E. Boyle, L. Kaser, R. Schnitzhofer, M. Graus, A. Turnipseed, J. Park, R. Weber, R. S. Hornbrook, C. A. Cantrell, R. L. Maudlin III, S. Kim, Y. Nakashima, G. M. Wolfe, Y. Kajii, E. Apel, A. H. Goldstein, A. Guenther, T. Karl, A. Hansel, and F. N. Keutsch (2012), Observations of glyoxal and formaldehyde as metrics for the anthropogenic impact on rural photochemistry, *Atmos. Chem. Phys.*, *12*(20), 9529-9543, doi:10.5194/acp-12-9529-2012.

Fu, T., D. Jacob, F. Wittrock, J. P. Burrows, M. Vrekoussis, and D. Henze (2008), Global budgets of atmospheric glyoxal and methylglyoxal, and implications for formation of secondary organic aerosols, *J. Geophys. Res.*, *113*, D15303, doi:10.1029/2007JD009505.

Geiger, H., J. Kleffmann, and P. Wiesen (2002), Smog chamber studies on the influence of diesel exhaust on photosmog formation, *Atmos. Environ.*, *36*(11), 1737-1747, doi:10.1016/S1352-2310(02)00175-9.

Guenther, A., et al. (1995), A global model of natural volatile organic compound emissions, *J. Geophys. Res.*, *100*(D5), 8873-8892, doi:10.1029/94JD02950.

Holzinger, R., C. Warneke, A. Hansel, A. Jordan, W. Lindinger, D. H. Scharffe, G. Schade, and P. J. Crutzen (1999), Biomass burning as a source of formaldehyde, acetaldehyde, methanol, acetone, acetonitrile, and hydrogen cyanide, *Geophys. Res. Lett.*, *26*(8), 1161-1164, doi:10.1029/1999GL900156.

Hoque, H. M. S., H. Irie, and A. Damiani (2018), First MAX‐DOAS observations of formaldehyde and glyoxal in Phimai, Thailand, *J. Geophys. Res.*, doi:10.1029/2018JD028480.

Houweling, S., F. Dentener, and J. Lelieveld (1998), The impact of nonmethane hydrocarbon compounds on tropospheric photochemistry, *J. Geophys. Res.*, *103*(D9), 10673-10696, doi:10.1029/97JD03582.

Irie, H., H. Takashima, Y. Kanaya, K. Boersma, L. Gast, F. Wittrock, D. Brunner, Y. Zhou, and M. V. Roozendael (2011), Eight-component retrievals from ground-based MAX-DOAS observations, *Atmos. Meas. Tech.*, *4*(6), 1027-1044, doi:10.5194/amt-4-1027-2011.

Kaiser, J., G. M. Wolfe, K. E. Min, S. S.Brown , C. C. Miller, D. J. Jacob, J. A. deGouw, M. Graus, T. F. Hanisco, J. Holloway, J. Peischl, I. B. Pollack, T. B. Ryerson, C. Warkene, R. A. Washenfelder, and F. N. Keutsch (2015), Reassessing the ratio of glyoxal to formaldehyde as an indicator of hydrocarbon precursor speciation, *Atmos. Chem. and Phys.*, *15*(13), 7571-7583, doi:10.5194/acp-15-7571-2015.

Lee, Y. N., X. Zhou, and K. Hallock (1995), Atmospheric carbonyl compounds at a rural southeastern United States site, *J Geophys. Res.*, *100*(D12), 25933-25944, doi:10.1029/95jd02605.

Li, X., T. Brauers, A. Hofzumahaus, K. Lu, Y. Li, M. Shao, T. Wagner, and A. Wahner (2013), MAX-DOAS measurements of NO 2, HCHO and CHOCHO at a rural site in Southern China, *Atmos. Chem. Phys.*, *13*(4), 2133-2151, doi:10.519/acp-13-2133-2013.

Liakakou, E., M. Vrekoussis, B. Bonsang, C. Donousis, M. Kanakidou, and N. Mihalopoulos (2007), Isoprene above the Eastern Mediterranean: Seasonal variation and contribution to the oxidation capacity of the atmosphere, *Atmos. Environ.*, *41*(5), 1002-1010, doi:10.1016/j.atmosenv.2006.09.034.

MacDonald, S. M., H. Oetjen, A. S. Mahajan, L. K. Whalley, P. M. Edwards, D. E. Heard, C. E. Jones, and J. C. Plane (2012), DOAS measurements of formaldehyde and glyoxal above a south-east Asian tropical rainforest, *Atmos. Chem. and Phys*, *12*(13), 5949-5962.

Miller, C., G. Gonzalez Abad, H. Wang, X. Liu, T. Kurosu, D. J. Jacob, and K. V. Chance (2014), Glyoxal retrieval from the ozone monitoring instrument, *Atmos. Meas. Tech.* , doi:10.5194/amt-7-3891-2014.

Monks, P. S. (2005), Gas-phase radical chemistry in the troposphere, in *Chem. Soc. Rev.*, 376-395, doi:10.1039/B307982C.

Munger, J. W., D. J. Jacob, B. C. Daube, L. Horowitz, W. Keene, and B. Heikes (1995), Formaldehyde, glyoxal, and methylglyoxal in air and cloudwater at a rural mountain site in central Virginia, *J. Geophys. Res.*, *100*(D5), 9325-9333, doi:10.1029/95JD00508.

Myriokefalitakis, S., M. Vrekoussis, K. Tsigaridis, F. Wittrock, A. Richter, C. Brühl, R. Volkamer, J. P. Burrows, and M. Kanakidou (2008), The influence of natural and anthropogenic secondary sources on the glyoxal global distribution, *Atmos. Chem. Phys.*, *8*(16), 4965-4981, doi:10.5194/acp-8-4965-2008.

Ortega, I., T. Koenig, R. Sinreich, D. Thomson, and R. Volkamer (2015), The CU 2-D-MAX-DOAS instrument–Part 1: Retrieval of 3-D distributions of NO 2 and azimuth-dependent OVOC ratios, *Atmos. Meas. Tech.*, *8*(6), 2371-2395, doi:10.5194/amt - 8 - 2371 - 2015.

Oum, K. W., M. J. Lakin, D. O. DeHaan, T. Brauers, and B. J. Finlayson-Pitts (1998), Formation of molecular chlorine from the photolysis of ozone and aqueous sea-salt particles, *Science*, *279*(5347), 74-76, doi:10.1126/science.279.5347.74.

Peters, E. (2013), Improved MAX-DOAS measurements and retrievals focused on the marine boundary layer, PhD Thesis, University of Bremen, Germany.

Platt, U., and J. Stutz (2008), *Differential Optical Absorption Spectroscopy*, Springer.

Poisson, N., M. Kanakidou, and P. J. Crutzen (2000), Impact of non-methane hydrocarbons on tropospheric chemistry and the oxidizing power of the global troposphere: 3-dimensional modelling results, *J. Atmos. Chem.*, *36*(2), 157-230, doi:10.1023/A:1006300616544.

Progress and problems in atmospheric chemistry, Volume 3, World Scientific, Singapore

Seco, R., J. Peñuelas, and L. Filella (2007), Short-chain oxygenated VOCs: Emission and uptake by plants and atmospheric sources, sinks, and concentrations, *Atmos. Environ.*, *41*(12), 2477-2499, doi:10.1016/j.atmosenv.2006.11.029.

Singh, H. B., Y. Chen, A. Staudt, D. Jacob, D. Blake, B. Heikes, and J. Snow (2001), Evidence from the Pacific troposphere for large global sources of oxygenated organic compounds, *Nature*, *410*(6832), 1078-1081, doi:10.1038/35074067.

Singh, H. B., L. J. Salas, R. B. Chatfield, E. Czech, A. Fried, J. Walega, M. J. Evans, B. D. Field, D. J. Jacob, and D. Blake (2004), Analysis of the atmospheric distribution, sources, and sinks of oxygenated volatile organic chemicals based on measurements over the Pacific during TRACE-P, *J. Geophys. Res.*, *109*, D15S07, doi:10.1029/2003JD003883.

Sinreich, R., U. Frieß, T. Wagner, and U. Platt (2005), Multi axis differential optical absorption spectroscopy (MAX-DOAS) of gas and aerosol distributions, *Faraday Discuss.*, *130*, 153-164, doi:10.1039/B419274P.

Stavrakou, T., J.-F. Müller, I. DeSmedt, M. Van Roozendael, G. R. Van Der Werf, L. Giglio, and A. Guenther (2009), Evaluating the performance of pyrogenic and biogenic emission inventories against one decade of space-based formaldehyde columns, *Atmos. Chem. and Phys.*, *9*(3), 1037-1060, doi:10.5194/acp-9-1037-2009.

Volkamer, R., J. L. Jimenez, F. San Martini, K. Dzepina, Q. Zhang, D. Salcedo, L. T. Molina, D. R. Worsnop, and M. J. Molina (2006), Secondary organic aerosol formation from anthropogenic air pollution: Rapid and higher than expected, *Geophys. Res. Lett.*, *33*, L17811, doi:10.1029/2006GL026899.

Vrekoussis, M., N. Mihalopoulos, E. Gerasopoulos, M. Kanakidou, P. Crutzen, and J. Lelieveld (2007), Two-years of NO_3 radical observations in the boundary layer over the Eastern Mediterranean, *Atmos. Chem. Phys.*, *7*(2), 315-327, doi:10.5194/acp-7-315-2007.

Vrekoussis, M., F. Wittrock, A. Richter, and J. P. Burrows (2010), GOME-2 observations of oxygenated VOCs: what can we learn from the ratio glyoxal to formaldehyde on a global scale?, *Atmos. Chem. Phys.* , *10*(21), 10145-10160, doi:10.5194/acp-10-10145-2010.

Weller, R., O. Schrems, A. Boddenberg, S. Gäb, and M. Gautrois (2000), Meridional distribution of hydroperoxides and formaldehyde in the marine boundary layer of the Atlantic (48° N-35° S) measured during the Albatross campaign, *J. Geophys. Res.* , *105*(D11), 14401-14412, doi:10.1029/1999JD901145.

Yokelson, R. J., J. G. Goode, D. E. Ward, R. A. Susott, R. E. Babbitt, D. D. Wade, I. Bertschi, D. W. Griffith, and W. M. Hao (1999), Emissions of formaldehyde, acetic acid, methanol, and other trace gases from biomass fires in North Carolina measured by airborne Fourier transform infrared spectroscopy, in *J. Geophys. Res.*, *104*(D23), 30109-30125, doi:10.1029/1999JD900817.

Wayne, R. P. (2000), Chemistry of Atmospheres, Oxford University Press

Wittrock, F. (2006), The retrieval of oxygenated volatile organic compounds by remote sensing techniques, PhD Thesis, University of Bremen, Germany.

Chapter 2: Absorption Spectroscopy

The information provided in this chapter is based on the textbooks by *Platz and Stutz* [2008] and *Lenoble* [1993]. Additionally, information was also taken from PhD thesis of *Alvarado* [2015] and *Peters* [2013]. Few information were also taken from the internet.

2.1 The solar spectrum

The sun is the unique source of radiant energy for the earth, as well as for other planets in the solar system. The sun is a gaseous sphere, made up of approximately three parts of hydrogen and one-part of helium. The mass and radius of the sun is ~1.99×10^{30} kg and 6.96×10^{5} km, respectively. The temperatures in the innermost region of the sun ~10^{7} K and the energy is produced by nuclear fusion reactions. The energy is transferred to the outer layer in the form of electromagnetic radiation. The surface of the sun is known as the photosphere, which is ~500 km thick. The photosphere emits most of the solar radiation received on the earth. The solar radiation emitted by the photosphere is a continuum close to a black body radiation at ~ 6000 K. The spectral distribution of a black body is given by Planck's law as follows

$$B\,(\lambda, T) = \frac{2\pi hc^2}{\lambda^5} \times \left(e^{\frac{hc}{\lambda kT}} - 1\right)^{-1} \qquad (2.1)$$

Where c is the speed of light, h and k are Planck and Boltzmann constant, respectively. B is the spectral irradiance (W/m^2/nm). The maximum wavelength of a black-body radiation at a given temperature is expressed by Wien's displacement law,

$$\lambda_{max} = \frac{b}{T} \qquad (2.2)$$

Where, T is the temperature in kelvin and b is the Wien's displacement constant which equals to 0.2879 cm K. Thus the maximum wavelength for the temperature of 6000 K is ~ 480 nm. Figure 2.1 shows an example of the solar spectrum. As seen from the figure, the solar spectrum is not a perfect black-body radiation curve and shows characteristic structures. These structured lines are called Fraunhofer lines. Above the photosphere, a so called solar-atmosphere compromises a cooling layer of a few hundred kilometers of temperature ~5300 K. This reversing layer is responsible for the absorption of energy which is superimposed on the photosphere continuum creating the structured lines in the solar spectrum. When the solar radiation passes through atmosphere, absorbers in the atmosphere causes further absorptions at characteristics wavelengths.

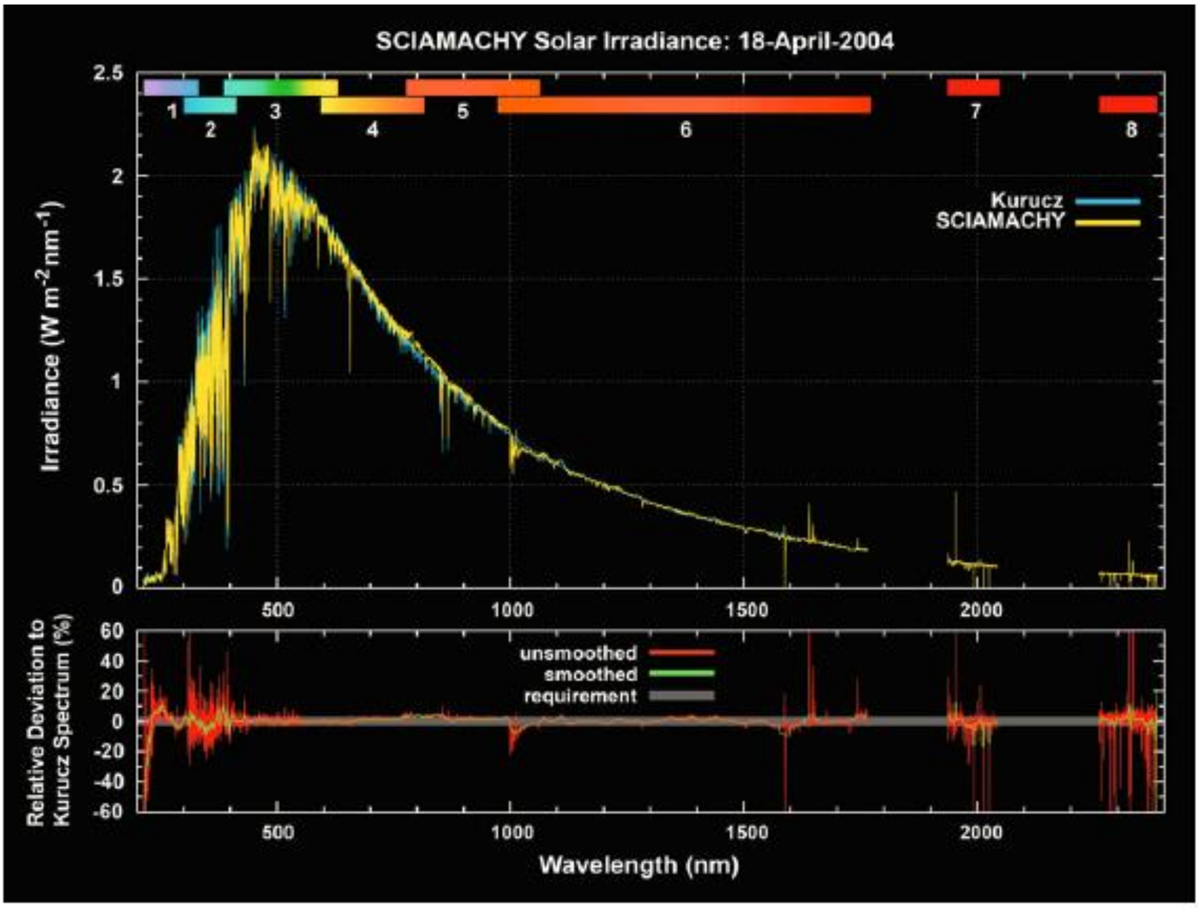


Figure 2.1: Solar spectrum measured by the SCIAMACHY instrument on April 18, 2004. (Adapted from the book SCIAMACHY: Exploring the Changing Earth's Atmosphere, Springer,2011)

2.2 Absorption spectra, energy levels and band structures

The absorption spectrum of gaseous molecules in the atmosphere has composed lines of organized bands. Here a short overview of basic principles of spectroscopy about the absorption lines in the electromagnetic spectrum is provided.

According to Max Planck, absorption and emission of radiation by matter take place in finite quanta energy. The energy E related to absorption and emission is proportional to the frequency hυ

$$E = h\upsilon \tag{2.3}$$

where, h is the Planck constant.

According to Bohr's postulates, electrons in the atoms occupy different energy states or levels. The transition of electrons from one energy state to another causes absorption or emission of energy and the energy associated with such transition equals to hυ

$$E_h - E_l = h\upsilon \text{ (emission)} \tag{2.4}$$

$$E_l - E_h = h\upsilon \text{ (absorption)} \tag{2.5}$$

where, E_h and E_l are the higher and lower energy states respectively. Such transitions give rise to complex lines in spectra. If the transition takes place in the presence of external field, it is called forced transition.

The total energy of a molecule (E) compromises of four modes of energy,

$$E = E_{rot} + E_{vib} + E_{elec} + E_{trans} \tag{2.6}$$

where,

E_{rot} → rotational energy

E_{vib} → vibrational energy

E_{elec} → electronic energy

E_{trans} → translational energy

Except translational energy, the other forms of energy are quantized. Rotational energy levels are spaced closely and rotational transitions result in rotational lines in the microwave or far-infrared region. Vibrational transition also occurs in the near-infrared region. Such transition corresponds to higher change of energy and is also accompanied by rotational transitions. Electronic transition corresponds to frequencies in the visible or ultraviolet region of the electromagnetic spectrum. The electronic bands have a complex structure due to simultaneous vibrational and rotational transitions. Excessive rotational energy can ionize the molecules and excessive vibrational excitation leads to dissociation of molecules. Emission or absorption don't occur when all transitions between quantized energy levels take place. Emission or absorptions are subject to selections rules, which are explained shortly.

2.2.1 Rotational Energy

For a rigid molecule, the rotational energy is given as:

$$E_{rot} = hcF(J) \tag{2.7}$$

where, $F(J)= BJ(J+1)$

B is the rotational constant and is inversely proportional to the moment of inertia. J is the angular momentum quantum number. The selection rule is

$\Delta J= 0, \pm 1$

$\Delta J= \pm 1$ gives absorption in a pure rotational spectrum. The symmetric top molecule has two moments of inertia which leads to two rotational constants A (along the symmetry axis) and B (perpendicular). The rotational energy is given by

$$E_{rot} = hcF(J, K) \tag{2.8}$$

With $F (J, K) = BJ(J+1) + (A-B) * K^2$

Where K is the second quantum number and K <= J. The selection rule is ΔK = 0.

2.2.2 Vibrational Energy

A molecule with N atoms has 3N degrees of freedom. A nonlinear and linear molecule has 3N-6 and 3N-5 modes of vibration, respectively. The vibrational energy for a simple harmonic oscillator is given by:

$$E_{vib} = hcG(v1, v2, \dots \dots ..) \tag{2.9}$$

with $G(v1, v2 \dots) = \sum_i (vi + \frac{1}{2}) \breve{\tilde{v}}_i$

where v_i are the vibrational quantum numbers. The selection rule for a harmonic oscillator are

$\Delta v1 = 1$

$\Delta l = 0, \pm 1$

The only frequencies are fundamentals v1, v2, v3. Some possible transitions can be forbidden by symmetry conditions. Anharmonicity will introduce modifications to the above selection rules.

2.2.3 Electronic Energy

The electronic energy for an atom with single valance electron, in the absence of any external electric or magnetic field is

$$E_{elec} = -\frac{hcRZ^2}{[n+a(l)]^2} \tag{2.10}$$

where Z is the atomic number and R is the Rydberg constant. n and l are the principal and azimuthal quantum numbers. Two magnetic quantum numbers m_l and m_s define the orbital angular momentum and the spin in the direction of an imposed magnetic field respectively. Pauli’s principle determined the lowest-energy electron configuration of an atom. According to Pauli’s principle, two electrons cannot have identical set of quantum numbers n, l, m_l, and m_s. With the increase of atomic number, successive shells specified by increasing number of n and l are occupied with electrons having all possible values of m_l and m_s.

2.2.4 Interactions

In combined transitions involving simultaneous changes of electronic, vibrational, and rotational energy, selection rules involving symmetry rules for the total wave function have to be considered. The vibrational transitions are accompanied by rotational transitions according to the selection rules,

$$\mathrm{J} = \begin{cases} -1 & P\ branch \\ 0 & Q\ branch \\ +1 & R\ branch \end{cases} \tag{2.11}$$

The rotational constant also depends on the vibrational state

$$B = B_e - \sum_i \alpha_i (v_i + \frac{1}{2} d_i) \tag{2.12}$$

i refers to the different vibrations with quantum number v_i and degeneracy d_i. a_i is a constant and Be equals to zero for no vibration. The lines in all the branches are separated, and the spacing in the P and R branch is not constant. Figure shows the energy transition on s vibrational-rotational band. The initial state is characterized by the quantum numbers (v", J") and the final state by (v', J') . If the wavelength number for the fundamental vibrational transition is v_f, then the wavenumbers of the rotation lines are

$$v_{J''}^{J'} = \begin{cases} v_f - (B' + B'')J'' + (B' - B'')J''^2 & P\ branch \\ v_f + (B' - B'')J'' + (B' - B'')J''^2 & Q\ branch \\ v_f + 2B' + (3B' - B'')J'' + (B' - B'')J''^2 & R\ branch \end{cases}$$

In a linear molecule, transitions with Δl = 0 are called parallel bands. If l = 0 in the lower and upper states, ΔJ = 0 is then forbidden, and there will be no Q branch, and a gap v=vf will exist between P and R branches. Again if Δl = 0 and l ≠ 0, a weak Q branch will exist. Transitions with Δl = ±1 are called perpendicular bands and will lead to very strong Q bands.

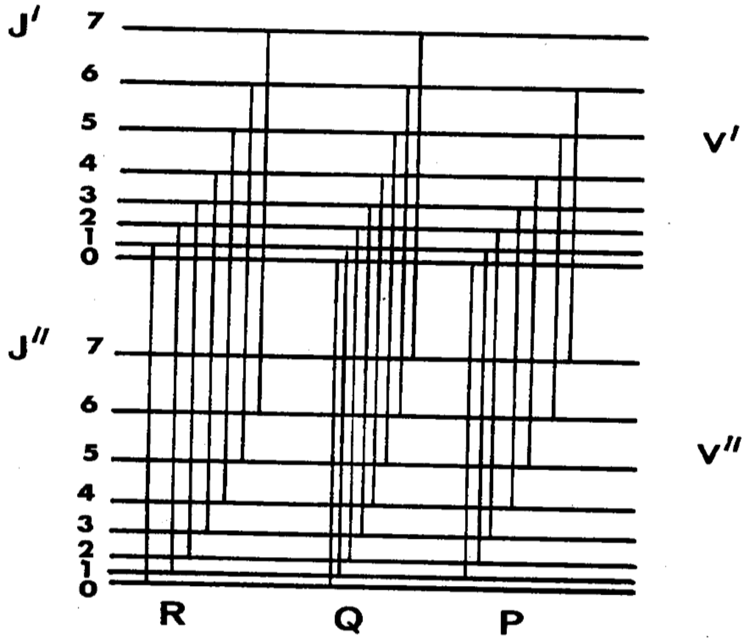


Figure 2.2: Energy transitions in a vibrational-rotational band (Adapted from *Lenoble* [1993])

2.3 Scattering in the atmosphere

Inhomogeneity in the propagation medium produces scattering of the electromagnetic radiation. The scattering of electromagnetic radiation can influence the measured spectrum which is used to retrieve different compounds (aerosols, trace gases, etc.) in the atmosphere. Scattering processes in the atmosphere are briefly explained in this section.

2.3.1 Rayleigh Scattering.

Rayleigh scattering is also known as elastic scattering i.e., scattering by air molecules occurs without the change of the photon energy. Such type of scattering is responsible for the blue color of the sky. The Rayleigh scattering cross-section σ_R was formulated by Rayleigh in 1899,

$$\sigma_R(\lambda) = \frac{8\pi^3}{3\lambda^4 N_{air}^2} * (n_o(\lambda)^2 - 1)^2 * F_K(\lambda) \qquad (2.13)$$

where, λ is the wavelength in cm, n_o is the index of refraction which is also wavelength-dependent, N_{air} is the number density of the air, and $F_k(\lambda) = 1.061$ is the anisotropy correction factor. For simple estimates, σ_R can be rewritten as follows.

$$\sigma_R(\lambda) = \sigma_{Ro} * \lambda^{-4} \qquad (2.14)$$

where $\sigma_{RO} = 4.4 \times 10^{-16}$ cm^2 nm^4 for air. Thus, the smaller the wavelength, the higher the Rayleigh scattering cross-sections. This explains the blue color of the light. The blue region of the visible spectrum has the shortest wavelength, hence scattered the most and the sky appears to be blue to the human eye. The Rayleigh extinction co-efficient ε_R (λ) is given by,

$$\varepsilon_R(\lambda) = \sigma_R(\lambda) * N_{air} \qquad (2.15)$$

And the Rayleigh phase function is given by

$$\phi(cos\theta) = \frac{3}{4}(1 + cos^2\theta) \qquad (2.16)$$

2.3.2 Raman Scattering

As mentioned previously, Rayleigh Scattering is regarded as elastic scattering, where no energy is transferred between the scattering particle and photons. Inelastic scattering occurs if the state of excitation of a scattering particle or molecule changes during the scattering process. In such process, a part of the photon energy is passed from the photon to the molecules or vice versa. In terms of radiant energy, Raman scattering is negligible compared with Rayleigh scattering. However, the Raman spectrum has a line structure characteristic of the molecule, which can be used to identify a component.

If a molecule vibrates or rotates with a frequency υ during the scattering process, the polarizability can be expressed as

$$\alpha = \alpha_o + \alpha_1 cos2\pi\upsilon t \qquad (2.17)$$

For incident radiation at a frequency υ_o, the dipole moment would be

$$P = \xi_o\{\alpha_o cos2\pi\upsilon_o t + \left(\frac{\alpha_1}{2}\right)[cos2\pi(\upsilon + \upsilon_o)t + cos2\pi(\upsilon - \upsilon_o)t]\} \qquad (2.18)$$

The first term on the right-hand side of equation (2.18) corresponds to Rayleigh scattering, and other two terms to Raman scattering. According to quantum theory, the intensity of the Raman line is proportional to the square of the matrix element for the polarizability

$$P_{mn} = \int \psi_m^* \, \alpha \, \psi_n d\tau \qquad (2.19)$$

where ψ_m and ψ_n are the Eigen functions of the Hamiltonian operator, $d\tau$ is the volume element. The asterisk denotes the complex conjugate. For a simple harmonic oscillator, the selection rules for vibrational energy transition is $\Delta v = \pm1$, which leads to frequencies

$\upsilon_1 = \upsilon_o - \upsilon_{vib}$; $\upsilon_2 = \upsilon_o + \upsilon_{vib}$;

The lines υ' around υ_o are similarly explained by the rotation of the molecule; because the polarizability is same after one half rotation. Thus, the selection rules are $\Delta J=\pm2$ instead of $\Delta J=\pm1$ for the rotational energy transitions and for a rigid rotator this lead to

$$|\upsilon' - \upsilon_o| = B(J+2)(J+3) - BJ(J+1) = 4B(J+\tfrac{3}{2})$$

instead of

$$\upsilon_{rot} = B(J+1)(J+2) = 2BJ(J+1)$$

This means, the spacing of lines in the Raman spectrum is twice of that in the infrared rotation spectrum. Similarly, the lines around υ'' around υ_1 corresponds to the rotational structure of the vibrational spectrum. If only the rotational spectrum is affected, then the term rotational Raman scattering is used. If the vibrational state changes with the rotational states, then the scattering is called rotational-vibrational Raman scattering. The rotational – vibrational Raman scattering is an order magnitude smaller than the rotational Raman scattering. Compared to the direct sunlight, rotational Raman scattering has important effect on the scattered sunlight in the atmosphere. As a consequence, the Fraunhofer lines changes in the spectrum of the scattered radiation. This also leads to less deep and strong Fraunhofer lines in the scattered radiation spectrum. Such features are not exclusive to the Fraunhofer structures, but also occur for several observers in the atmosphere, such as ozone. This effect is called the ring effect, which was first formulated by Grainger and Ring (1962) and explained by Brinkman (1968).

2.3.3 Mie scattering

Mie scattering occurs when electromagnetic radiation interacts with matter of dimensions analogous to the wavelength of the incident radiation. This was first described by Gustav Mie. Radiation can be absorbed or scattered by particles. These processes can be described by the absorption and scattering coefficient ε_a (r, λ) and ε_s (r,λ) respectively,

$$dI = -I(\lambda) * \varepsilon_a(r,\lambda)ds \qquad (2.20)$$

$$dI_s = -I(\lambda) * \varepsilon_s(r,\lambda)ds \qquad (2.21)$$

where I(λ) is the radiation flux, dI is the change of the radiation flux after passing through a layer of aerosol of thickness ds. The single scattering albedo A_S and extinction coefficient ε_M is given by

$$A_s = \frac{\varepsilon_s(r,\lambda)}{\varepsilon_a(r,\lambda)+\varepsilon_s(r,\lambda)} \qquad (2.22)$$

$$\varepsilon_M(r,\lambda) = \varepsilon_a(r,\lambda) + \varepsilon_s(r,\lambda) \qquad (2.23)$$

Compared to Rayleigh scattering, the dependence of Mie scattering on wavelength is much weaker (typically proportional to $\lambda^{-1.3}$). The Mie phase function depends on the ratio between radius r and wavelength of the scattered radiation and the size parameter is given by

$$\alpha_s = \frac{2\pi r}{\lambda} \qquad (2.24)$$

The computational effort to calculate the Mie phase function is reduced by the introduction of analytical expression which depends on only few parameters. The most common parameterization is the Henyey-Greenstein approximation:

$$\phi(cos\theta) = \frac{(1-g^2)}{4\pi(1+g^2-2gcos\theta)^{\frac{3}{2}}} \qquad (2.25)$$

Where g is the asymmetry factor and is given by

$$g = \frac{1}{2}\int_{-1}^{1} P(cos\theta) * cos\theta \; dcos\theta \qquad (2.26)$$

For isotropic scattering, ϕ(cosθ) = constant. For complete forward scattering, g = 1. The typical values of g for tropospheric aerosols are ~0.6 - 0.7.

2.4 Absorption in the atmosphere

When a photon is absorbed by a molecule, rotational, vibrational, and electronic transition can occur. The absorption process is described by the classical Lambert – Beer's Law, also often referred to as Bouguer – Lambert law. To formulate the absorption of gaseous absorbers, Lambert – Beer law can be expressed as:

$$I(\lambda) = I_o(\lambda) * \exp(-\sigma(\lambda) * c * L) \qquad (2.27)$$

Here, I_o (λ) is the initial intensity of the radiation emitted from a suitable source, I (λ) is the intensity of the radiation after passing through the medium of thickness L, σ(λ) is the absorption cross-section at the wavelength λ and c is the average concentration of the gaseous absorber. If all the quantities are known, then c can be calculated as follows:

$$c = \frac{\ln(\frac{I_O(\lambda)}{I(\lambda)})}{\sigma(\lambda)*L} = \frac{D}{\sigma(\lambda)*L} \qquad (2.28)$$

where D is called the optical density of a layer of a given absorber. Although equation (2.27) is the basis of absorption spectroscopic application in the laboratory, it is very challenging to use equation (2.27) for atmospheric application. Because, in laboratory applications, I_o (λ) can be measured precisely. For atmospheric applications, all the absorbers in the atmosphere needs to be removed to determine the initial intensity $I_o(\lambda)$, which is practically impossible. The solution lies in so called “differential absorption spectroscopy” which is explained in the next section.

2.4.1 Differential Optical Absorption Spectroscopy (DOAS)

The differential optical absorption spectroscopy (DOAS) was first used by Dobson in 1930 to determine the total column of ozone. Since then this method has been used to measure trace gases in the troposphere and stratosphere. For the application of equation (2.27) in the atmosphere, the processes influencing the intensity of the light in the atmosphere need to be taken into account. Such processes include absorptions of various trace gases with concentration c_j, absorption cross-section $\sigma_j(\lambda)$, Rayleigh ($\varepsilon_R(\lambda)$) and Mie extinction ($\varepsilon_M(\lambda)$), instrumental effects and turbulence. Thus the Beer – Lambert can be written as

$$I(\lambda) = I_o(\lambda) * \exp[-L * (\sum(\sigma_j(\lambda) * c_j) + \varepsilon_R(\lambda) + \varepsilon_M(\lambda))] * A(\lambda) \qquad (2.29)$$

where $A(\lambda)$ summarizes the instrumental effects and turbulence. As mentioned before, in the laboratory, all the factors affecting the intensity of the radiation can be quantified easily. For atmospheric application, the DOAS technique is used to quantify the factors affecting the intensity of the radiation. Aerosol extinction processes, effect of turbulence, and many trace gas absorptions have broad band spectral characteristics. However, certain trace gases exhibit narrowband absorption structures. The DOAS technique separates the broadband - and narrowband spectral structures in the spectrum and isolates the narrowband spectral structures. The separation of the narrowband- and broadband for one absorption band is performed for both absorption cross-section and intensity.

$$\sigma_j(\lambda) = \sigma_{jo}(\lambda) + \sigma_j^{'}(\lambda) \qquad (2.30)$$

where σ_{jo} is the slow varying component with respect to wavelength and σ_j' (λ) is the rapid varying component with respect to wavelength. The terms ‘slow” and “rapid” variation of absorption cross-section depends on the wavelength interval and the width of the absorptions to be detected. Substituting equation (2.30) into equation (2.29)

$$I(\lambda) = I_o(\lambda) * \exp[-L * (\sum(\sigma_j'(\lambda) * c_j)] * \exp[-L * (\sum\left(\sigma_{jo}(\lambda) * c_j\right) + \\ \varepsilon_R(\lambda) + \ \varepsilon_M(\lambda))] * A(\lambda) \qquad (2.31)$$

The first exponential of equation (2.31) describes the effect of structured differential absorption of an absorber and the second exponential refers to the slow varying absorptions as well as the other factors influencing the intensity of the radiation. $A(\lambda)$ describes the broad wavelength dependent transmission of the optical system and turbulence. Now, a quantity I_o' can be defined as the intensity when there is no differential absorption.

$$I_o'(\lambda) = I_o(\lambda) * \exp[-L * (\sum(\sigma_{jo}(\lambda) * c_j) + \varepsilon_R(\lambda) + \varepsilon_M(\lambda))] * A(\lambda) \qquad (2.32)$$

Similarly, the differential optical density D' can be defined as

$$D' = ln\frac{I'_o(\lambda)}{I(\lambda)} = L * \sum \sigma'_j(\lambda) * c_j \qquad (2.33)$$

The trace gas concentration can be then calculated by substituting the differential quantities D' and σ' respectively. The differential absorption cross-sections are determined in the laboratory. Using equation (2.31) and separating the narrowband- and broadband spectral structures requires the measurement of the radiation intensity at multiple wavelengths. Generally, DOAS measurements cover the intensity at 500-2000 individual wavelengths to accurately determine the concentrations of various absorbers. This is an expansion of the DOAS principle used by Dobson, which was based on two or four wavelengths. Using the differential absorptions over extended wavelength range eliminates the necessity of radiometric calibration of the instrument, because the transmission of optical instruments generally shows broadband characteristics. Thus, the elimination of radiometric calibration makes the instrument simple, less expensive, and eligible for long-term observations. Moreover, further advantage of the DOAS technique is the ability to quantify very weak absorbers, corresponding to optical densities ~10^{-4}.

2.4.2 The DOAS equation

For the measurement of the scattered sunlight in the atmosphere, eq (2.31) can be written as

$$I(\lambda) = I_o(\lambda) * \exp\{-\int[\sum_{j=1}^{n} \sigma_j(\lambda, s) * \rho_j(\lambda, s) + \sigma_{Rayleigh}(\lambda, s) * \rho_{Rayleigh}(s) + \sigma_{Mie}(\lambda, s) * \rho_{Mie}(s)]\}ds \qquad (2.34)$$

where n is the number of absorbers affecting the radiation. ρ (λ, s) is the concentration of the absorber along the light path s. In an approximation, the absorption cross-sections are assumed to be similar along the light path. In general, cross-sections are slightly dependent on temperature and altitude. Moreover, the retrieved compound of interest (aerosol, trace gases, etc.) are located at some certain altitude rather than being spread across the whole atmosphere. For some trace gases, cross-sections at different temperatures are used in the retrieval to account for the temperature dependence of the target absorber. Thus, we can define a term called slant column density (SCD) as follows,

$$SCD_j = \int \rho_j(s)ds \qquad (2.35)$$

SCD_j is defined as the integrated concentration of a trace gas along the light path s with unit in molecules/cm^2. Now, applying the DOAS technique (i.e. separating the narrow- and broadband features) and inserting equation (2.35) into equation (2.34),

$$I(\lambda) = I_o(\lambda) * \exp\{-\sum_{j=1}^{n} \sigma'_j(\lambda) * SCD_j - \sum_p a_p\lambda^p\} \qquad (2.36)$$

Where σ' is the differential absorption cross-section and $I_o(\lambda)$is the reference measurement. The polynomial compromises the Rayleigh and Mie scattering, instrumental influence, and turbulence. Taking the natural logarithm and rearranging equation (2.36),

$$\tau(\lambda) = ln\frac{I_o(\lambda)}{I(\lambda)} = \sum_{j=1}^{n} \sigma'_j(\lambda) * SCD_j + \sum_p a_p\lambda^p \qquad (2.37)$$

τ(λ) is the optical depth. Equation (2.36) is called the DOAS equation. This is the basis of analysis of any sort of DOAS measurements. There are variety of arrangement and observation modes of DOAS instruments. The primary classification is based on the light sources as active and passive DOAS. Active DOAS systems use artificial light sources such as laser and passive DOAS systems use natural light bodies such as sun, moon, stars etc. as light sources. Among passive DOAS, multi-axis DOAS (MAX-DOAS) utilizes multiple viewing geometry for observations. MAX-DOAS uses the fact that, at low elevation angles the length of the light path in the troposphere is amplified, thus the absorber concentration in the lower troposphere can be investigated. Additionally, using enough elevation angles, the vertical profile of the trace gases can also be derived. In this thesis, we have utilized the MAX-DOAS system for measuring scattered sunlight. So, from now onwards, the DOAS analysis will be explained mainly focusing on MAX-DOAS observation geometries and measurements.

For ground-based DOAS observations, $I_o(\lambda)$is measured at the zenith direction at lower solar zenith angle (SZA). Because, due to the shorter light path through the atmosphere, the absorption of different absorbers is smaller. Thus, the SCD is not absolute. A term differential SCD is used, which is defined as the difference between the SCD of the reference measurements and the SCD along the light path for other viewing geometry (i.e. MAX-DOAS). Thus the optical depth can be written as

$$\tau(\lambda) = ln\frac{I_o(\lambda)}{I(\lambda)} = \sum_{j=1}^{n} \sigma_j'(\lambda) * \Delta SCD_j + \sum_p a_p\lambda^p \qquad (2.38)$$

where ΔSCD is the differential slant column of the target species.

2.4.3 Vertical Column Density (VCD) and the air mass factor

The SCD strongly depends on the viewing geometry. Thus, for further convenience, the vertical column density (VCD) is calculated, which is defined as the concentration of the absorber integrated along the vertical path,

$$VCD_j = \int \rho_j(z)dz \qquad (2.39)$$

where $\rho_j(z)$ is the concentration of the absorber and dz is the vertical path through the atmosphere.

SCD and VCD are related by a term called air mass factor (AMF). AMF is defined as the ratio of SCD to VCD

$$AMF = \frac{SCD}{VCD} \qquad (2.40)$$

The AMF depends on the radiative transfer of the atmosphere, wavelength, surface albedo, temperature, aerosol and trace gas profiles, clouds, SZA etc. One of the limitations during the computation of the AMF is that, the profile of the absorber $\rho_j(z)$ has to be known beforehand. But in reality, the profile is generally not known or well understood. Moreover, the measurements are performed at different viewing geometries leading to the altitude dependence. To overcome this limitation and characterize the altitude-dependent sensitivity during AMF calculation, the so called Block-AMF (BAMF) concept is used. BAMFs are air mass factors, which correspond to the single layer j of the atmosphere and can be expressed as,

$$BAMF_j = \frac{SCD_j}{VCD_j} \qquad (2.41)$$

where SCD_j and VCD_j are the SCD and VCD in the layer j, respectively. The total SCD then can be defined as follows,

$$SCD = \sum_{j=1}^{TOA} BAMF_j * VC_j \qquad (2.42)$$

Thus, the total SCD is expressed as the sum of the SCDs from the surface layer to the top of the atmosphere (TOA). The advantage of BAMF over AMF is that, BAMF doesn't depend on the absorber profiles. Then, AMF is calculated as,

$$AMF = \frac{\sum_{i=0}^{TOA} BAMF_i * VC_i}{\sum_{i=1}^{TOA} VC_i} \qquad (2.43)$$

2.4.4 Profile Retrieval

MAX-DOAS observations are conducted at different elevation angles. Thus such measurements contain the information of the absorber profile. Here the basic theory of the profile retrieval is explained in brief. The profile information to be retrieved can be expressed with a vector x and the SCD columns calculated from the MAX-DOAS measurements can be expressed with the vector y, then a linear equation can be expressed as

$$y = Kx \qquad (2.44)$$

where K is the weighting function matrix

Equation (2.44) is a linear equation. To retrieve the profile information x, equation (2.44) is required to be inverted. This leads to a so called under-determined system or ill-posed problem, because the number of measurement is less than the number of unknowns (profile information). To solve such ill-posed problems, the optimal estimation method [*Rodgers* 2000] is utilized. This method has wide application in remote sensing field to retrieve the profile of unknown absorbers in the atmosphere. Thus, the profile of interest can be retrieved according to *Rodgers* [2000] as follows,

$$x = x_a + (K^T S_\varepsilon^{-1} K + S_a^{-1})^{-1} K^T S_\varepsilon (y - Kx_a) \qquad (2.45)$$

where,
x_a= the a priori information of the absorber of interest
S_ε = measurement error covariance matrix
S_a= a priori error covariance matrix
Another term called the averaging kernel is calculated to measure the quality of the retrieval. The averaging kernel (A) is defined as

$$A = \frac{\partial x}{\partial x_{true}} = (K^T S_\varepsilon^{-1} K + S_a^{-1})^{-1} K^T S_\varepsilon K \qquad (2.46)$$

where, x is the retrieved profile and x_{true} is the true profile. Thus, the averaging kernel matrix shows the change in the retrieved profile with respect to the true profile. In other words, the averaging kernel matrix shows how the retrieved profile is “smoothed”. For the ideal case, the retrieved profile will be same as the true profile, and the averaging kernel will be an identity matrix.

References

Alvarado, L. M. A. (2016), Investigating the role of glyoxal using satellite and MAX-DOAS measurements, PhD thesis, University of Bremen, Germany.

Lenoble, J. (1993), *Atmospheric Radiative Transfer*, A. DEEPAK Publishing, USA, ISBN 0-937194-21-2

Peters, E. (2013), Improved MAX-DOAS measurements and retrievals focused on the marine boundary layer, PhD Thesis, University of Bremen, Germany.

Platt, U., and J. Stutz (2008), *Differential Optical Absorption Spectroscopy*, Springer.

Rodgers, C. D. (2000), *Inverse methods for atmospheric sounding: Theory and practice*, World Scientific, Singapore.

Chapter 3: MAX-DOAS observations of formaldehyde and glyoxal in Phimai, Thailand

This chapter is primarily based on the research article of *Hoque et al* [2018a].

3.1 Introduction

For the purpose of continuous aerosol and trace gas observations, a MAX-DOAS instrument was installed at Phimai (15.18 °N, 102.56 °E), a rural site in central Thailand. This is the first continuous MAX-DOAS observation in this region. Phimai is a rural site shared by ground-based remote sensing network SKYNET (http://atmos3.cr.chiba-u.jp/skynet/) and the Asian dust and aerosol LIDAR observation Network (AD-Net) (http://www-lidar.nies.go.jp/AD-Net/). SKYNET is an observational network for aerosol-cloud-radiation interaction researches. Although sky-radiometers are the primary instrument of the SKYNET sites, some SKYNET sites also facilitate other observational instruments such as MAX-DOAS, microwave radiometer, cloud camera etc. The continuous operation of the MAX-DOAS instrument in Phimai started on September 2014. In this study, simultaneous measurements of glyoxal (CHOCHO) and formaldehyde (HCHO) from October 2014 to October 2016 are analyzed and utilized to estimate the CHOCHO to HCHO concentration ratio (R_{GF}) for this site. There are two seasons in Phimai: dry (October to May) and wet (June to September) season. Phimai can be affected by biomass burning activities during the dry season [*Campbell et al.*, 2013; *Li et al.*, 2013a; *Sugimoto et al.*, 2015]. Thus, this site provides us the advantage to assess the features of the R_{GF} values under the influence of different VOC emission sources (*i.e.* biogenic emissions and biomass burning). The changes in the R_{GF} values between the seasons are assessed and the reasons for these changes are evaluated.

3.2 Observation Location

The Phimai (15.18 °N, 102.56 °E) site is situated in central Thailand (Fig. 3.1). The capital city Bangkok is located ~260 km southwest of Phimai. As a Metropolitan region strong industrial and vehicular emissions are expected to occur in Bangkok. However, no industries are situated close to Phimai and the measurement site is surrounded by crop land. Around 4 km to the north there is a small town and major roadways. Dry and cool air from the northeast direction affects Phimai from November to mid-February whereas, wet and warm winds blow from the southwest direction from mid-May to September. The rest of the period (i.e., from mid-February to mid-May) is regarded as a transition period as the wind direction starts shifting from northeast to southwest. As mentioned earlier, there are two seasons in Phimai: (i) dry season (compromising eight months from October to May) and (ii) wet season (compromising four months from June to September). For enabling interseasonal comparison, the dry season was reduced to four months. Around March, the influence of biomass burning activities in Phimai is likely to be the strongest [*Sugimoto et al.*, 2015]. Therefore, January to April was regarded as the dry season. Due to some technical issues, no observations were available in July and September, 2015. Thus, during the wet season in 2015 measurements were only available for two months (June and August). The climate classification for the current study stands as: (i) dry season 2015 (January to April 2015), (ii) wet season 2015 (June and August

2015), (iii) dry season 2016 (January to April 2016), and (iv) wet season 2016 (June to September 2016).

Figure 3.1. Location of Phimai, which is ~230 km northeast of Bangkok, the capital of Thailand. The map has been produced using the Google map application.

3.3 Measurements and Retrieval Algorithm

3.3.1 MAX-DOAS

The instrument used in Phimai was similar to that used by *Irie et al* [2011, 2015]. The MAX-DOAS instrument consists of two major parts: an ultraviolet-visible (UV-VIS) spectrometer (Maya2000Pro, Ocean Optics) located indoors and a telescope unit placed outdoors. The spectrometer and the telescope were connected with a fiber-optic bundle cable consisting of seven cores where each core have a diameter of 100µm. The telescope unit has a fixed single telescope and a movable 45° inclined mirror on a rotary actuator for both the reference and off-axis directions.

Figure 3.2. The MAX-DOAS instrument measuring at Phimai, Thailand

Measurement of scattered sunlight was conducted at six elevation angles (ELs) of 2°, 3°, 4°, 6°, 8°, and 70°, and the same sequence of ELs was repeated every 30 min. The telescope unit was placed in a weather shielded quartz tube cap and the telescope was a Plano-convex quartz lens with a focal length of 40 mm and diameter of 25 mm. The spectrometer was equipped with a linear array of charged coupled detectors (CCD) with 3648 pixels (TCD1304AP; Toshiba, Tokyo, Japan). A stepping motor with an angular step of 0.038° was used to control the mirror angle. A laptop was used to control the mirror position and to collect the house-keeping information (such as temperature, offset, etc.). To avoid sedimentation of dust, water droplets etc., a small fan was used under the quartz tube cap of the telescope unit. The field of view (FOV) of the instrument was <1°.

The ELs for spectral measurements were limited to less than 10°. Constraining the off-axis ELs <10° and deriving an elevation angle dependent correction factor for oxygen collision complex (O_4) absorptions, *Irie et al.* [2015] reported good agreement between coincident cavity ring-down spectroscopy, LIDAR and sky radiometer. Such constrains in the ELs are expected to minimize, the possible systematic error in O_4 fitting results, while the measurement sensitivity in the lowest layer of the aerosol and trace gas profiles retrieved remained high. The spectra between 310 and 515 nm were recorded. Wavelength calibration was performed using a high-resolution solar spectrum from *Kurucz et al.* [1984]. The spectral resolution at full width half maximum was approximately 0.4 nm at 357 and 476 nm (Fig 3.4). The wavelength calibration was performed daily to take into account the wavelength shifts and spectral resolution.

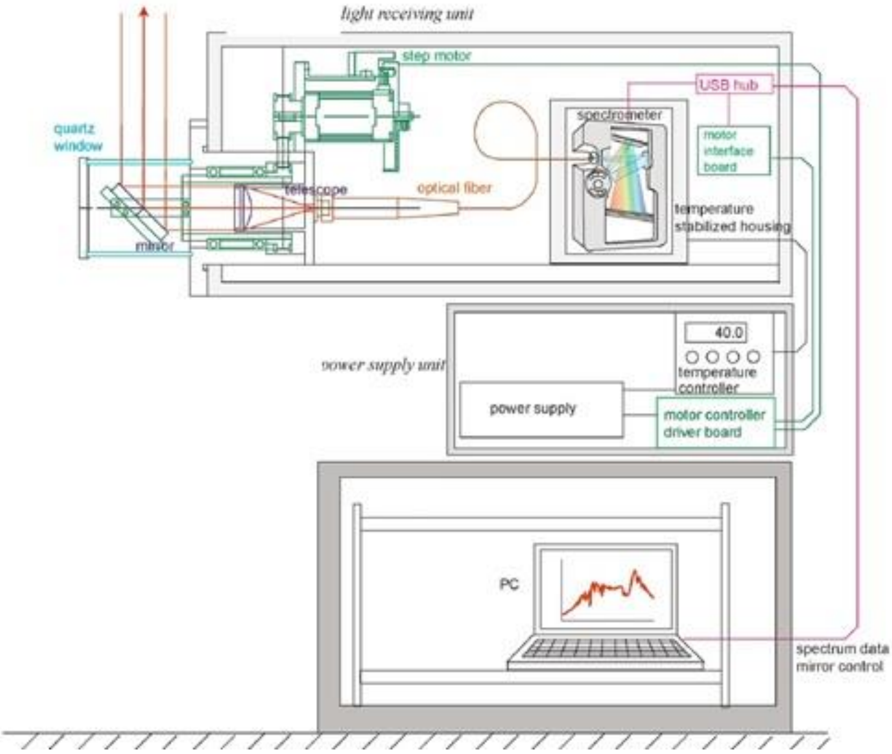


Figure 3.3. Schematic view of the MAX-DOAS instrument. The figure has been slightly modified from the original figure from *Kanaya et al* [2014].

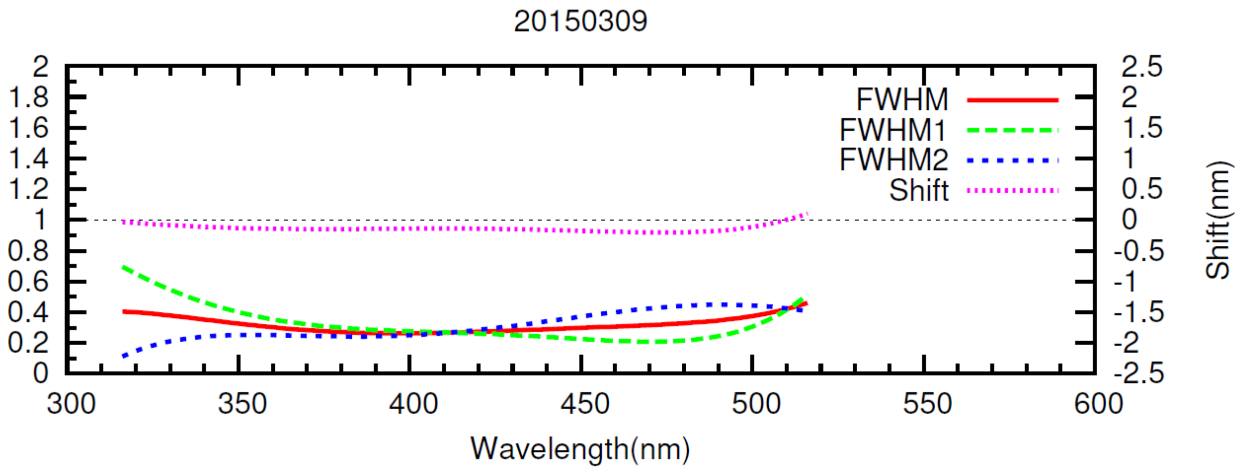


Figure 3.4. Example of wavelenght calibration on March 3, 2015.

3.3.2 Retrieval algorithm

(This section is based on section 2 of *Hoque et al* [2018a]).

The measured spectra were analyzed using the Japanese vertical profile retrieval algorithm version 2 (JM2) [*Irie et al.,* 2011, 2015]. The working procedure of the algorithm is based on three steps which are explained below. The detail formulation of the algorithm is also described by *Irie et al*. [2008a, 2008b, 2011, 2015].

(1) DOAS fit

In the first step, DOAS technique [*Platt*, 1994] was used to analyze the measured spectra and the differential slant column density (ΔSCD) of trace gases were retrieved based on the nonlinear least-square spectral fitting method. The respective fitting windows used for HCHO ΔSCD and CHOCHO ΔSCD retrievals were 336-359 and 436-457 nm. The O_4 ΔSCD values were retrieved from wavelengths around 360 and 477 nm which contains significant O_4 absorptions [*e.g., Irie et al.,* 2009, 2011]. The respective fitting windows for O_4 absorptions were 338-370 and 460-490 nm. The equation used for spectral fitting was:

$$lnI(\lambda) = ln(I_o(\lambda) - c(\lambda)) - \sum_{i=1}^{n} \sigma_i(\lambda) \Delta SCD_i - p(\lambda) \qquad (3.1)$$

where $I_o(\lambda)$ is the reference spectrum measured at time t. $I_o(\lambda)$ was derived from the interpolation of two spectra measured at EL = 70°. The two spectra were measured before and after the off-axis measurement within 30 min. The 30 min is the time to complete a whole sequence of EL (i.e., 5 minutes per EL). This procedure was followed to make the reference measurement under similar atmospheric conditions as the off-axis measurements. So, the difference between the reference measurement and the off-axis measurements should mainly be due to the ELs i.e., path length. If rapid changes occur within the time of one complete EL scan, for example due to clouds, the residuals of the aerosol profiles will be larger and will require data screening. The impact of clouds and data screening will be discussed later in this chapter. ΔSCD is called the differential slant column density. ΔSCD is calculated from the difference between the slant column density along the light path for an off-axis measurement and that for a reference spectrum measurement. $c(\lambda)$ is the wavelength dependent offset and also compromises the influence of stray light. The effect of Rayleigh and Mie scattering $p(\lambda)$ were taken into account by fitting second- and third-order polynomials, respectively. The absorption cross-section data used and the absorbers fitted in the HCHO and CHOCHO fitting windows are listed in Table 3.1. Figure 3.5 shows an example of the fitting results.

Figure 3.6 shows an example of the calculated O_4 ΔSCD, CHOCHO ΔSCD, and HCHO ΔSCD at all ELs, under clear sky condition on April 14, 2015. The clear sky condition was selected from the coincidence LIDAR measurements at Phimai. The error in the ΔSCDs were calculated from two type of statistics. One from the dark current measurements during nighttime and another from the fitting residuals. The corresponding mean error in the CHOCHO ΔSCD and HCHO ΔSCD on April 14, estimated from both type of statistics was ~9×10^{13} and ~2×10^{15} molecules cm^{-2}, respectively.

Table 3.1. Data sources of the absorption cross sections used in this study and the fitted observers in the sources and absorbers fitted in the HCHO and CHOCHO window

Cross-section	**Absorbers fitted**	**Data Source**
O_3		*Bougmil et al.* [2003], 223K
NO_2		*Vandaele et al.* [1996], 295K
BrO		*Fleischmann et al.* [2004], 223K
Ring		*Chance and Spurr* [1997]
H_2O		*Rothman et al.* [2009], 280K
O_4		Hermans et al. [2003], 296K
HCHO	O_3, NO_2, HCHO, BrO, O_4, Ring	*Meller and Moortgart* [2000], 293K
CHOCHO	O_3, NO_2, CHOCHO, H_2O, O_4, Ring	*Volkamer et al.* [2005], 296K

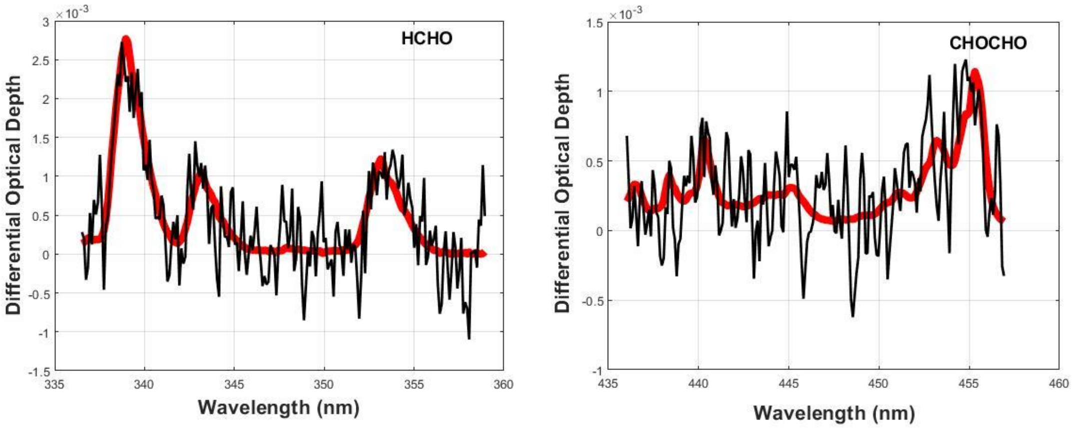


Figure 3.5. An example of spectral fitting of HCHO and CHOCHO using the DOAS technique. The HCHO spectrum was measured at 16:00 and the CHOCHO spectrum was measured at 12:00 local time (LT) on April 14, 2015 at an EL of 2°. The red curve shows the results of the DOAS fit. The black lines indicate the scaled cross-sections plus the fitting residuals.

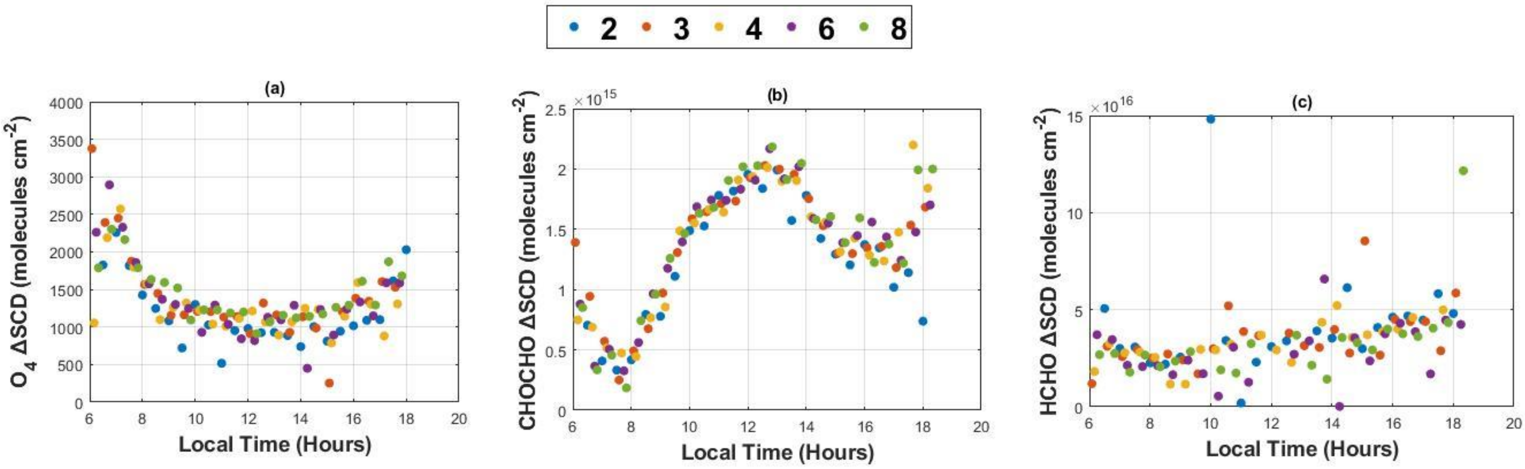
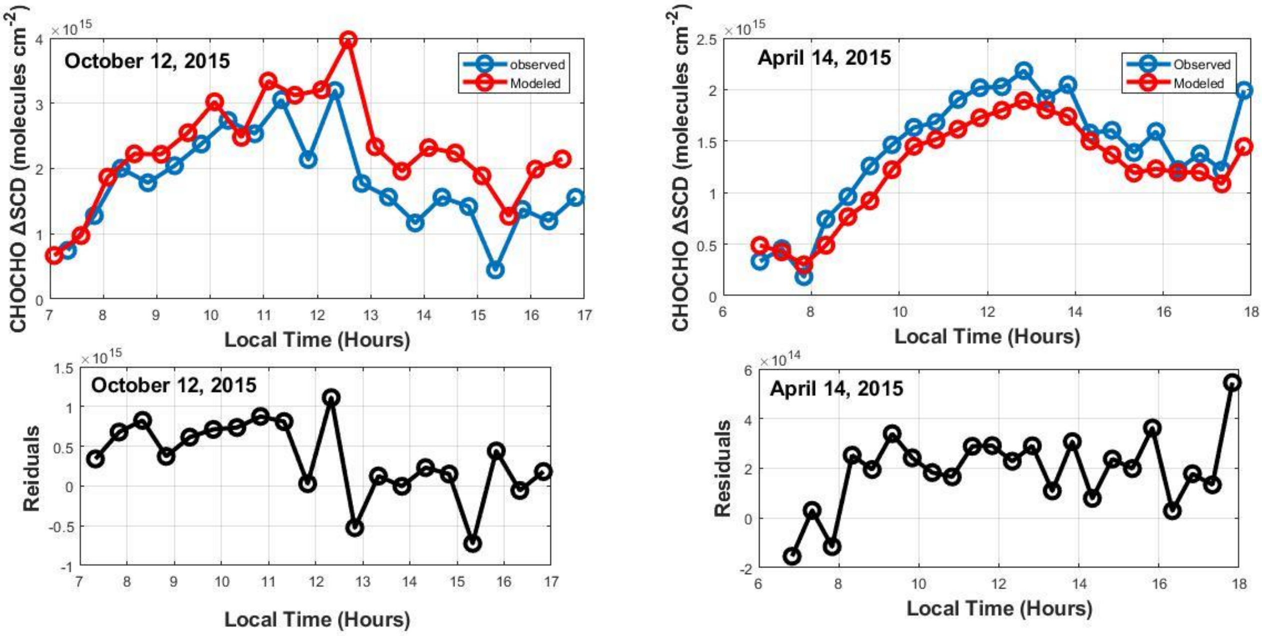


Figure 3.6. Examples of (a) O_4 ΔSCD, (b) CHOCHO ΔSCD, and (c) HCHO ΔSCD at all ELs on April 14, 2015.



.

Figure 3.7. Examples of fiiting residulas of CHOCHO measurements on October 12 and April 14, 2015, at an EL of 8°. Ocotober 12, 2015 is a cloudy day whereas April 14 was a clear sky day. The distinction between clear sky and cloud conditions were based on the coincident LIDAR measurements.

Figure 3.7 shows residuals of CHOCHO ΔSCD at an EL=8° on October 12 and April 14, 2015. These dates are arbitrary examples of cloudy (October 12,2015) and clear (April 14, 2015) days, selected based on the coincident LIDAR observations. The residuals on the cloudy day are higher comparative to the clear sky conditions. Such large residuals are subject to data screening, which will be explained later in this chapter.

(2) Aerosol profile/column retrieval

After the DOAS analysis, the retrieval of aerosol optical depth (AOD) τ and the aerosol extinction coefficient (AEC) k were performed. This step of the algorithm is based on the optimal estimation method [*Irie et al.*, 2008a; *Rodgers*, 2000]. The iteration equation used to solve the nonlinear inversion problem was as follows:

$$x_{i+1} = x_i + (S_a^{-1} + K_i^T S_\varepsilon^{-1} K_i + \gamma_i D)^{-1} \{K_i^T S_\varepsilon^{-1} [y - F(x_i)] - S_q^{-1} [x_i - x_a]\}$$

(3.2)

where, x_{i+1} and x_i are the current and previous state vectors, respectively. y is the measurement vector and x_a is the a priori state vector. Sε and Sa are the error covariance matrices of the measurement and the a priori information. K_i is the weighting function matrix, and F is the forward model which calculates O_4ΔSCD values based on given aerosol information. D is a diagonal scaling matrix, and γ_i is used to optimize the retrieval and is updated at each iteration step.

y (measurement vector) and x (state vector) are defined as:

$$y = (O_4\ \Delta SCD(\Omega_1) \ldots \Delta SCD(\Omega_n))^T \qquad (3.3) \text{ and}$$

$$x = (\tau\ \ F_1\ F_2\ F_3)^T \qquad (3.4),$$

where *n* denotes the number of measurements in 30 min (time for one complete EL scan). Ω is the viewing geometry. Ω compromises of three quantities: the solar zenith angle (SZA), EL, and relative azimuth angle. The *F* values are profile shape determining parameters with values ranging from 0-1. The partial AOD for 0-1, 1-2, and 2-3 km were defined as AOD ·F_1, AOD ·(1-F_1) F_2, and AOD · (1-F_1) (1-F_2) F_3 respectively. Above 3 km, the partial AODs were defined as AOD ·(1-F_1) (1-F_2) (1-F_3). From 3 to 100 km, a continuous exponential AEC profile was determined assuming an AEC value at the top layer and an exponential profile shape. Similarly, continuous profiles for layers below 3 km, *i.e.*, 2-3, 1-2, and 0-1 km were determined. The advantage of such parameterization is that, *F* values can be used instead of the absolute value of the AEC. Moreover, when profile shape determining parameters are used instead of the absolute values of AEC, the relative variation in the averages for the 0-1 km layer is lower, [*Irie et al.* 2008a]. On the other hand, drawback of using such parameterization is that, to determine the vertical resolution and measurement sensitivity of the retrieval, we have to depend on other supporting information. Retrievals and simulations performed by other groups for the similar viewing geometry of our instrument [*i.e., Frieß et al.*, 2006] were taken as reference to overcome the disadvantages of the parameterization. x_a (a priori state vector) and S_a (a priori covariance matrix) were constructed following the similar procedure of *Irie et al.* [2008a] based on two months of LIDAR data. The *F* values used in this study were: AOD = 0.21 ± 3.0, $F_1 = 0.60 \pm 0.05$, $F_2 = 0.80 \pm 0.03$, and $F_3 = 0.80 \pm 0.03$. These values were identical to those of *Irie et al.* [2011].

A lookup table (LUT) of the box air mass factor (A_{box}) vertical profile were created using the radiative transfer model JACOSPAR [*Irie et al.*, 2015]. JACOSPAR is based on the Monte Carlo Atmospheric Radiative Transfer Simulator (MCARaTS) [*Iwabuchi,* 2006]. The results of MCARaTS were compared with other radiative transfer models [*Wagner et al.,* 2007]. To simulate a realistic atmosphere, the altitude of the surface and the MAX-DOAS instrument (212 m a.s.l) was considered. The A_{box} LUT was used to determine the optimal aerosol loading and the A_{box} profiles, which accounted for the O_4 ΔSCD at all ELs. The A_{box} LUT was calculated assuming single values of the single scattering albedo (s = 0.95), asymmetry parameter (g = 0.65, under the Henyey-Greenstein approximation), and surface albedo (a = 0.10). Further sensitivity tests were conducted to examine the sensitivities of the AEC retrievals to changing these parameters (g, s, and a) by ±0.05, which will be discussed later in this chapter.

The retrieval can be characterized by analyzing the averaging kernels. Figure 3.8 shows the mean averaging kernel of the MAX-DOAS aerosol retrievals at 476 nm for the whole observation period in Phimai. The mean value of the degrees of freedom for signal (DOF) [*Rodgers,* 2000] for the aerosol retrieval at 476 nm was 2.54. To assess the contribution from the measurements in the retrieval, the area [*Rodgers*, 2000], was calculated according to the procedure of *Lui et al* [2005]. The calculated areas were 1.0, 0.8, 0.5, and 0.1 for τ, F_1, F_2, and F_3, respectively. The highest values were observed for τ, followed by F_1 and F_2. This indicate that the retrieval was performed by scaling the given a priori profile first, then followed by change in the values of F_1 and F_2. *Irie et al*. [2008a] conducted a sensitivity test to estimate the errors in the retrieved state vectors due to the choice of a priori values and their errors. Such errors of τ, k (0–1 km), k (1–2 km), and k (2–3 km) values at 476 nm were estimated to be 10%, 9%, 34%, and 43%, respectively.

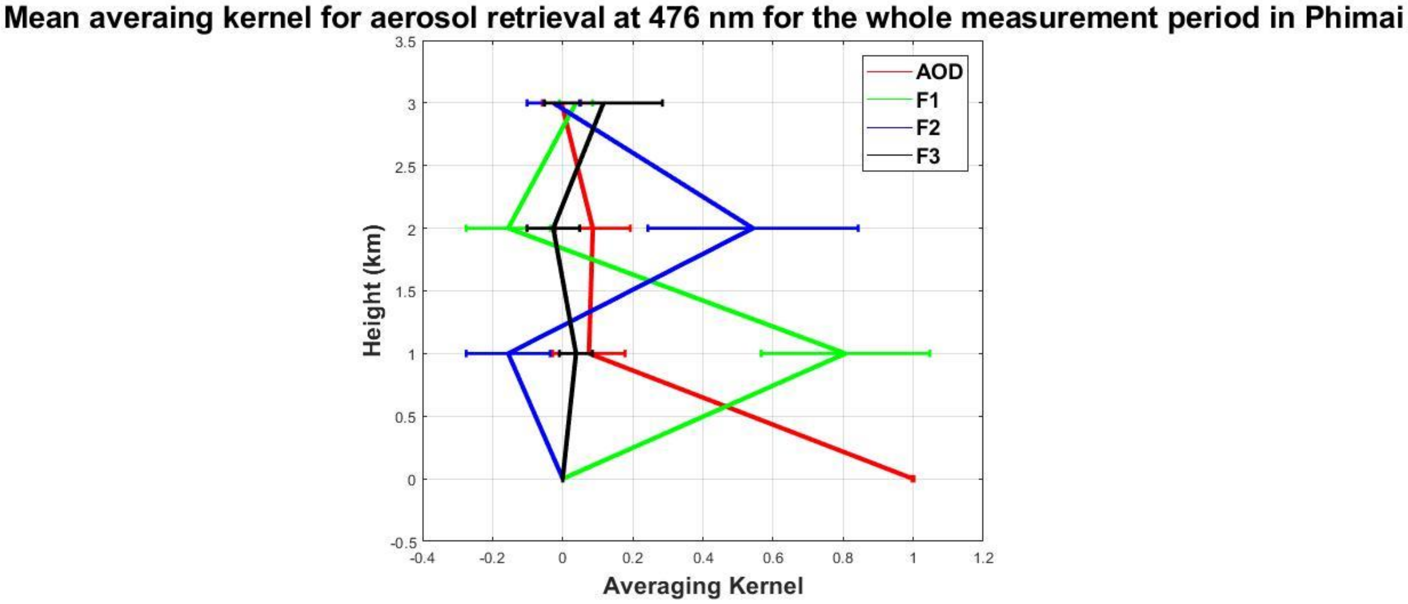


Figure 3.8. Mean averaging kernel for the aerosol retreival at 476 nm for the whole measurement period in Phimai.

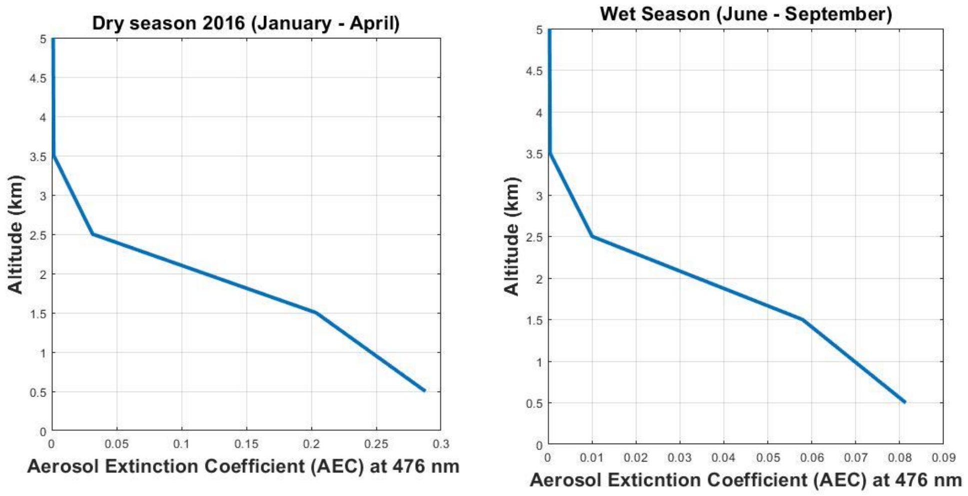


Figure 3.9. Mean AEC profile at 476 nm for the dry (left panel) and wet (right panel) season in 2016.

The random error in the retrieval of the state vector quantities can be characterized by the covariance matrix S

$$S = (K^T S_\varepsilon^{-1} K + S_a^{-1})^{-1} \qquad (3.5)$$

The random error represents the smoothing error and the retrieval error noise [*Rodgers* 2000]. The residuals of O_4 ΔSCD fittings were used to construct the measurement error covariance matrix. This was because residuals were higher than the error in the O_4 ΔSCDs. The estimated random error in the AEC at 476 nm for the 0-1 km layer was ~24% for the whole observation period. The systematic error can have several sources. According to *Irie eta al* [2008a, 2009, 2011], the total systematic errors for the retrieved AEC at 476 and 354 nm are less than 30 and 50% respectively.

Figure 3.9 shows the mean AEC profile at 476 nm during the dry and wet season in 2016, as an example of the retrieved AEC profile using JM2. The higher AEC during the dry season could potentially be due to the influence of biomass burning, because the dry season in Phimai is affected by biomass burning [*Sugimoto et al* 2015].

(3) Trace gas retrieval

In the third step of the retrieval, the trace gas (i.e. HCHO and CHOCHO) ΔSCDs were converted to vertical column densities (VCDs) and to their respective vertical profiles. The trace gas retrieval step also utilizes the prepared LUT and the nonlinear iterative inversion method similar to the aerosol retrieval. The retrieval procedure of HCHO and CHOCHO were similar. Thus, only the CHOCHO retrieval is explained below.

For the trace gas retrievals y (measurement vector) and x (state vector) were defined as

$$y = (CHOCHO\ \Delta SCD(\Omega_1) \ldots CHOCHO\ \Delta SCD(\Omega_n))^T \qquad (3.6) \text{ and}$$

$$x = (VCD\ f_1\ f_2\ f_3)^T \qquad (3.7)\ ,$$

VCD is the vertical column of the trace gas below 5 km. The *f* values are the profile shape determining factors similar to the *F* values used in the aerosol retrieval step. A continuous profile was assumed above 5 km. Similar to the aerosol retrieval step, VCD and *f* values were used to define the partial VCD values for 0-1, 1-2, 2-3, and 3-5 km. From the partial VCD value, we obtain the mean number density of the trace gas in a particular layer. This number density is then converted to volume mixing ratio (VMR) of the trace gas using the U.S. standard atmospheric temperature and pressure data. The temperature and pressure data were scaled to the surface conditions of the Phimai site. The profile shape determining values (*f*) were similar to those used for the aerosol retrieval.The a priori state vector x_a for retrieving the CHOCHO VCDs was constructed from the 20% of the maximum CHOCHO ΔSCD during one complete scan of the ELs (i.e. 30 min). The error of the a priori values were set to 100% of the highest ΔSCD values. Setting the a priori error to 100% made the VCD independent of the a priori data used for retrieval. Thus, all the row vectors in the averaging kernel matrix almost equals to one.

The ΔSCD values were used as a priori from the fact that, ΔSCD contains altitude dependent information of the trace gases. Figure 3.10 shows the mean averaging kenels for the HCHO and CHOCHO retreival. The areas for VCD and f_1 was 1 and between 0.2-0.4, respectively, indicating that, the retrrieval was performed scaling the apriori information similar to the aerosol retrieval. The mean DOF of the signal of the CHOCHO and HCHO retrieval for the whole observation period were 1.46 and 1.30, respectively. Figure 3.11 shows the mean CHOCHO and HCHO profiles for the dry and wet season. Similar to the AEC profile (Fig 3.9), higher concentration of CHOCHO and HCHO are observed during the dry season, which is mostly likely due to the influence of biomass burning.

For the conversion of the calculated CHOCHO vertical profiles to CHOCHO ΔSCD the LUT compromising the A_{box} profiles prepared during the aerosol retrieval was used. The A_{box} profiles in the LUT were calculated for the wavelengths at 357 and 476 nm which are different from the representative wavelength of the trace gases. To overcome such difference in the representative wavelengths, the AOD at the wavelengths of HCHO and CHOCHO was determined from the AOD at the wavelength closer to the HCHO and CHOCHO wavelengths, respectively. For such wavelength conversion of AOD Ångström exponent value of 1.00 was used. The AOD recalculated at the representative wavelength of the trace gas retrievals were used to match the respective A_{box} values in the LUT. Adopting such measures will not significantly impact the A_{box} and trace gas concentration profile in case of optically thin absorbers [*Irie et al.*, 2011; *Wagner et al.*, 2007].

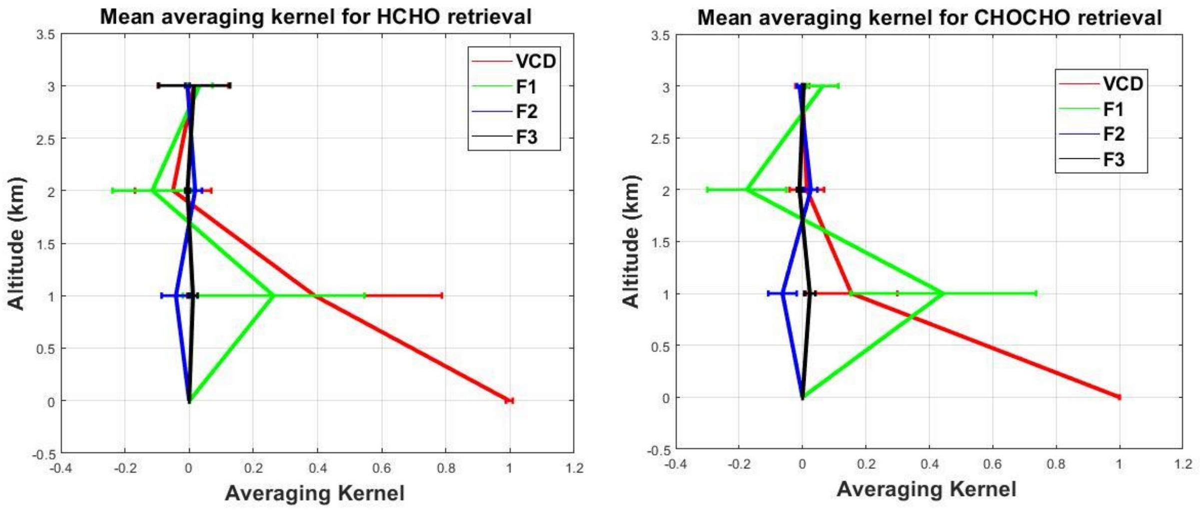


Figure 3.10. Mean averaging kernel of the HCHO (left panel) and CHOCHO (right panel) for the whole observation period.

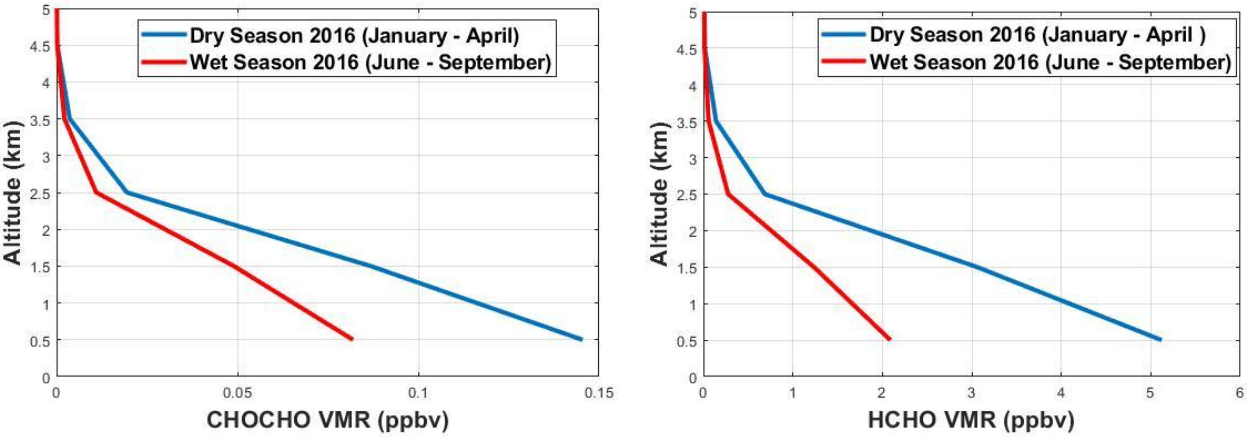


Figure 3.11. Mean CHOCHO (left panel) and HCHO (right panel) profiles for the dry and wet seasons in 2016.

The random error in the retrieved trace gas vertical profiles can be characterized by equation (3.5) similar to the aerosol retrieval. The residuals of HCHO ΔSCD and CHOCHO ΔSCD were used respectively to construct the S_e (measurement error covariance matrix). The systematic error was estimated by varying the AOD by additional ± 30 and ±50% for the CHOCHO and HCHO retrievals respectively. These statistics were based on the uncertainty in the aerosol retrieval in the UV and visible region [*Irie et al*, 2011].

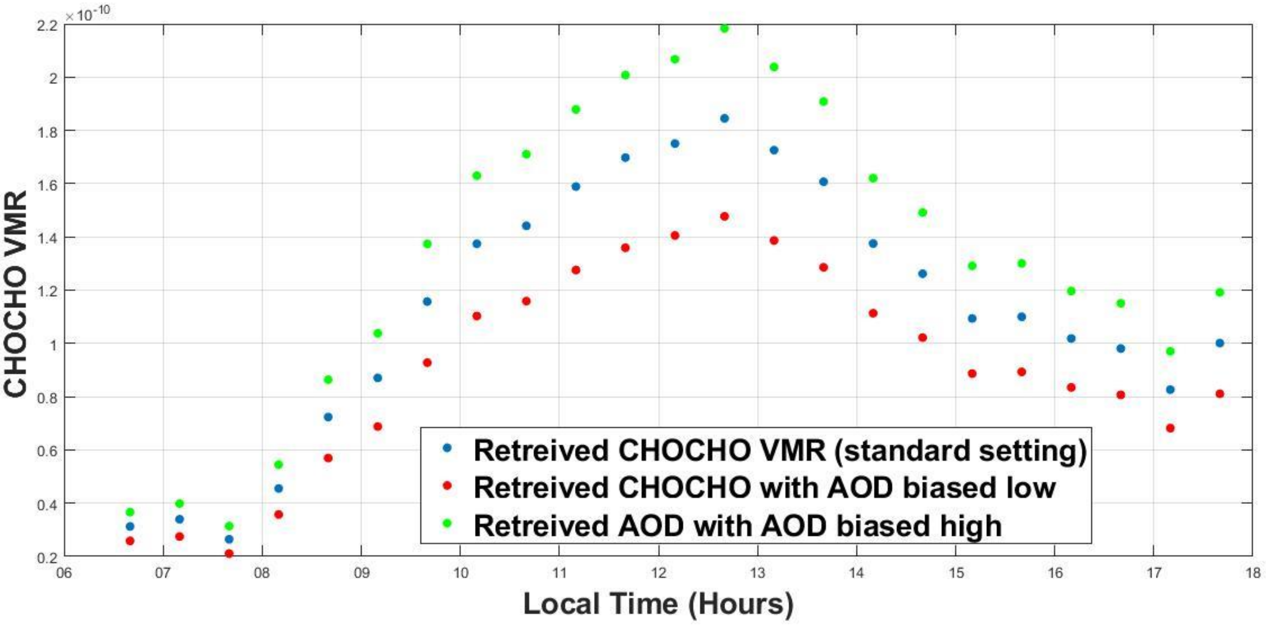


Figure 3.12. CHOCHO retreival on April 14, 2015 (clear sky condition) for different AOD settings. The blue color indicates the standard setting, whereas the green and red colors are the CHOCHO VMRs retreived by changing the AOD by -30 and +30% respectively.

Another source of error in addition to the random and systematic errors estimated previously is the bias in the ELs. During MAX-DOAS observations, the horizon needs to be defined properly for the precision of the ELs. Inappropriate horizon definition can lead to biases in the ELs. In general, defining the horizon precisely is a challenging task during setup of MAX-DOAS instruments. Thus, such errors were estimated in this study. Figure 3.13 depicts the impact of bias in the ELs on the retrieved HCHO and CHOCHO VMRs for the 0-1-km on a selected clear sky day on April 14, 2015. Coincident LIDAR measurements at Phimai were used to judge the clear sky conditions on the respective date. The general setting of the instrument is referred as ‘EL with no bias’ (*i.e.*, the optimal setting of the instrument). New A_{box} profiles were calculated by varying the EL by ±0.5° and the resulting trace gas (i.e. HCHO and CHOCHO) VCDs and vertical profiles were retrieved which are shown in Fig. 3.13. When the EL is varied by +0.5°, higher A_{box} and VMR values are observed. Similarly, varying the EL by -0.5° lower A_{box} and VMR values are observed. Varying the EL by +0.5° yields that the instrument is pointing below the original horizon position. The positive shift in the EL will result in an increase of the light path length and lead to higher A_{box} and VMR values. High A_{box} values is an indication of low aerosol loading in the atmosphere. Thus the light path through the absorber layer (trace gas) will be increased and the resulting VMR values will be higher. The sensitivity test shows that the difference between the retrieval based on the standard settings and assumed additional biases is ~ 5% for both HCHO and CHOCHO. Thus the inaccuracy in defining the horizon during the instrument step which might lead to biases is expected to have low impact on the final retrieved traced gas VMR values. The bias in the ELs can also impact the retrieved VCDs. The sensitivity test showed that, such biases impact the retrieved VCDs and VMR values in the same order of magnitude. Table 3.2 summarizes the estimated errors for the MAX-DOAS observations of HCHO and CHOCHO in Phimai.

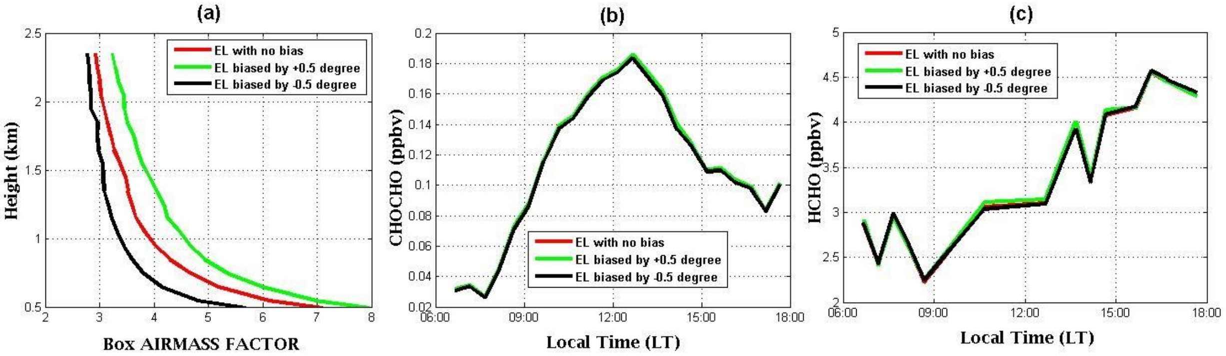


Figure 3.13. Plot showing the additional AMF calculations and retreivals of HCHO and CHOCHO VMRs , prformed to investigate the influence of possible bias in the ELs during our instrumental setings. (a) The box AMFs shown here are for SZA of 45°, a RAA of 180°, and an EL of 2°. Assesing the influence of potential bias in the EL on the trace gas retreival. Different cases of the box AMF calculations are shown in different colours. The standard setting of the instrument (i.e., our current setting) is reffered as” EL with no bias”, plotted in

red color. The green and black color plot indicate the AMF calculations for additional bias by +0.5°and -0.5°respectively. (b) The retrieval of CHOCHO concentrations for the 0-1 km layer corresponding to the cases of AMF calculations in plot (a). (c) The retrieval of HCHO concentrations for the 0-1 km layer corresponding to the cases of AMF calculations in plot (a)

Table 3.2. The random, systematic, and estimated bias form the ELs in the retrieved HCHO and CHOCHO concentrations for the 0-1 km layer. The errors are expressed in percentage.

Component	Random error	Systematic error	Error due to bias in EL	Total error
CHOCHO	10	16	5	20
HCHO	17	24	5	30

3.3.3 Other sources of error

(1) Error induced due to fixed values of g, s, and a.

To calculate the A_{box} LUT single values of the single scattering albedo (s = 0.95), asymmetry parameter (g = 0.65, under the Henyey-Greenstein approximation), and surface albedo (a = 0.10) were used. These could lead to error in the retrieved aerosol and trace gas products. Thus, further sensitivity tests were performed to estimate the errors due the fixed values of these parameters. For this, retrievals were performed by varying g, s, and a by ±0.05. The results of the sensitivity test are summarized in table 3.3.

Table 3.3. Estimated error due the fixed values of g, s, and a

Error source	Test	AOD	AEC (0-1 km)	CHOCHO VCD	CHOCHO VMR (0 -1 km)	HCHO VCD	HCHO VMR (0 -1 km)
g	+0.05	8%	9%	<2%	<2%	<2%	<2%
g	-0.05	7%	8%	<2%	<2%	<2%	<2%
s	+0.05	2%	2%	<1%	<1%	<1%	<1%
s	-0.05	<1%	<1%	<1%	<1%	<1%	<1%
a	+0.05	<2%	<2%	<1%	<1%	<1%	<1%
a	-0.05	2%	<2%	<2%	<2%	<2%	<2%

As seen from the table, the errors induced due to the fixed values of the parameters were within the range of the total estimated error for the retrieval.

(2) Impact of other absorbers in the fitting window

For the retrieval of the trace gases, additional absorbers were fitted in the fitting window of the target trace gas. In case the target trace gas is a weak absorber, it can be sensitive to the other strong absorbers in the fitting window. For example, weak absorber CHOCHO can be sensitive to the H_2O cross-section. To assess such influence on the CHOCHO retrieval, the retrieval of CHOCHO were tested using different H_2O cross-section versions and temperature.

Figure 3.14 shows the impact of using different H_2O cross-sections on the CHOCHO fit and retrieval. H_2O cross-sections form HITRAN database 2004 and 2009 at 280 K were used for the test. The main retrieval was performed using the HITRAN 2009 database for H_2O cross-sections. No significant influence of different H_2O cross-sections on the CHOCHO retrieval were observed. Due to different versions of H_2O cross-section data, the mean difference in the optical depth was ~ 7%. In some cases, such differences were ~18 %, but within the range of our estimated error for CHOCHO.

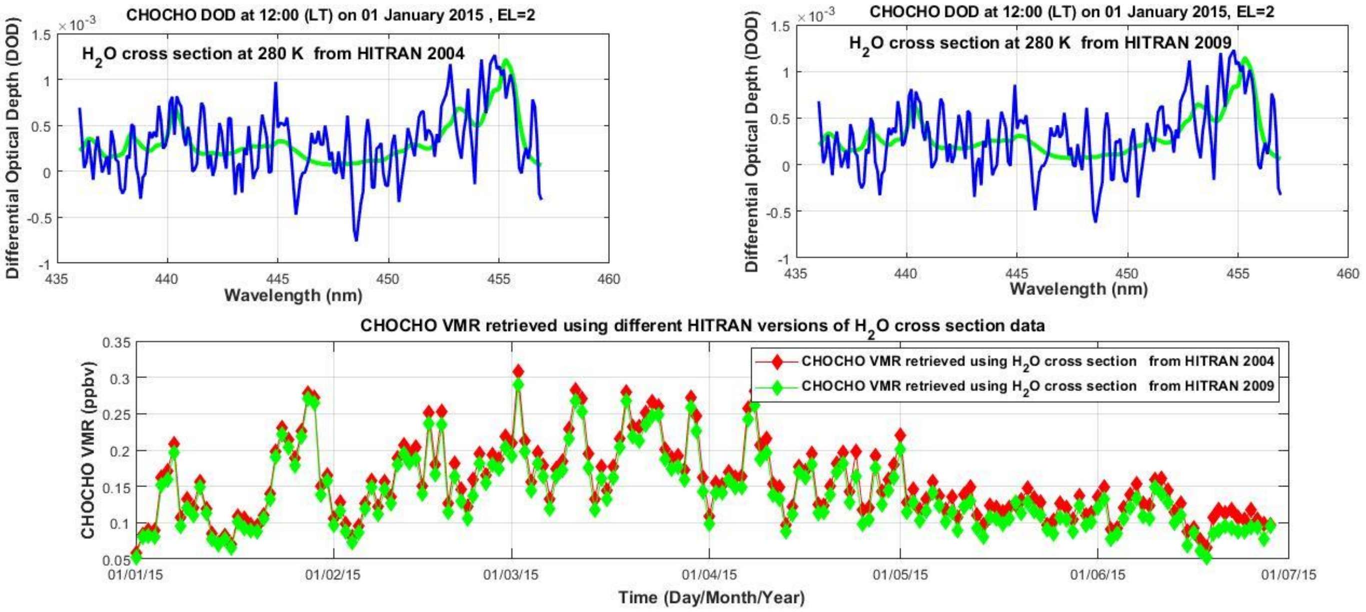


Figure 3.14. (upper panel) Fiitings of CHOCHO using H_2O cross section from HITRAN 2004 (left panel) and 2009 (right panel). The temperature of both the H_2O cross section versions were 280 K. (lower panel) CHOCHO VMR retrieved for the 0-1 km layer using different H_2O cross section.

Figure 3.14 also shows the CHOCHO VMR retrieved using two different versions of H_2O cross-sections. In most of the cases, the differences between the two retrievals were within 10 %. Therefore, the error which might have incurred due to the difference in H_2O cross-section data, were within the range of our estimated total error for CHOCHO.

Figure 3.15 shows the CHOCHO retrieval using two different H_2O cross-section temperature. The HITRAN 2009 H_2O cross-sections at 280 and 296 K were used for the test. The standard retrieval was performed using the HITRAN 2009 cross-sections at 280 K. In most of the cases, the difference between the two retrievals were within 6%. Therefore, the error, which might have incurred due to the difference in H_2O cross-section temperature, will be within the range of the estimated total error for CHOCHO.

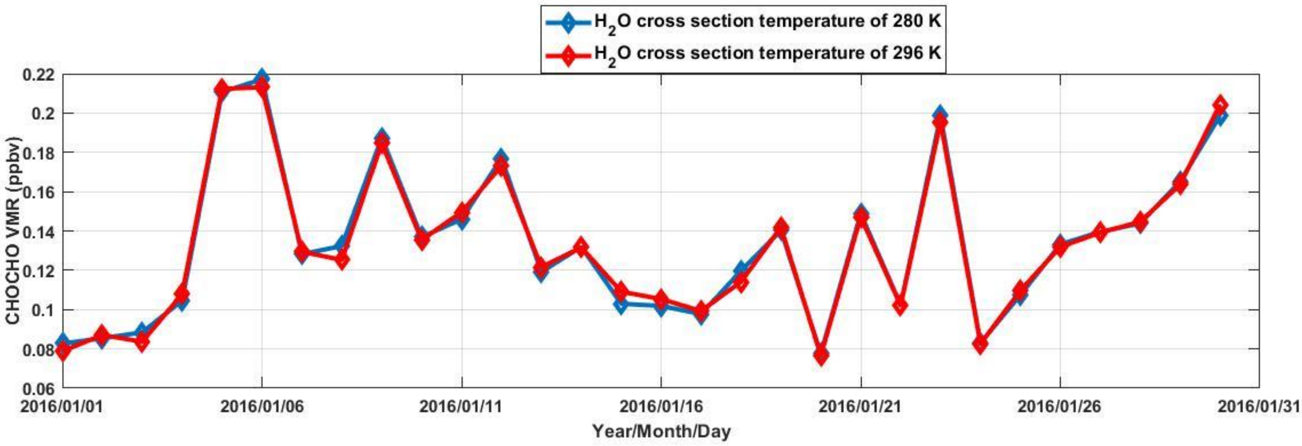


Figure 3.15. CHOCHO retrieved using two different water vapor cross section temperatures of 280 K and 296 K. The example has been shown for the measurements on January 2016.

3.3.4 Cloud influence on the MAX-DOAS retrievals

MAX-DOAS measurements can be influenced by clouds which can bias the retrieve products. The discrimination between clouds and aerosols are yet challenging. We performed data screening to minimize the number of measurements influenced by clouds. Two steps were followed. In the first step, retrievals for retrieved AOD greater than 3 were excluded. AOD = 3 is the highest value in the LUT. Optically thick clouds are primarily responsible for such large optical depths. In the second step, screening was performed based on residuals of the O_4 ΔSCD, trace gas ΔSCDs, and the DOF. The threshold for the residuals of O_4 ΔSCD, CHOCHO ΔSCD, and HCHO ΔSCD were < 10%, <50%, and <50%, respectively. The DOF threshold was set to DOF> 1.02.

Figure 3.16 shows an example of the impact of data screening on the retrieval. The clear and cloudy days were distinguished using coincident sky view camera images (Fig. 3.17) and LIDAR observations (Fig. 3.18). Both June 27 and 28 were rainy days. Compared to the clear sky days, there were very few retrievals on the cloudy days. This means, the data screening implemented, that limits the ranges of the retrieved AOD, the ΔSCD fitting residuals yielded in the vertical profile retrieval, and the DOF for the signal has excluded most data influenced by clouds present below ~1 km.

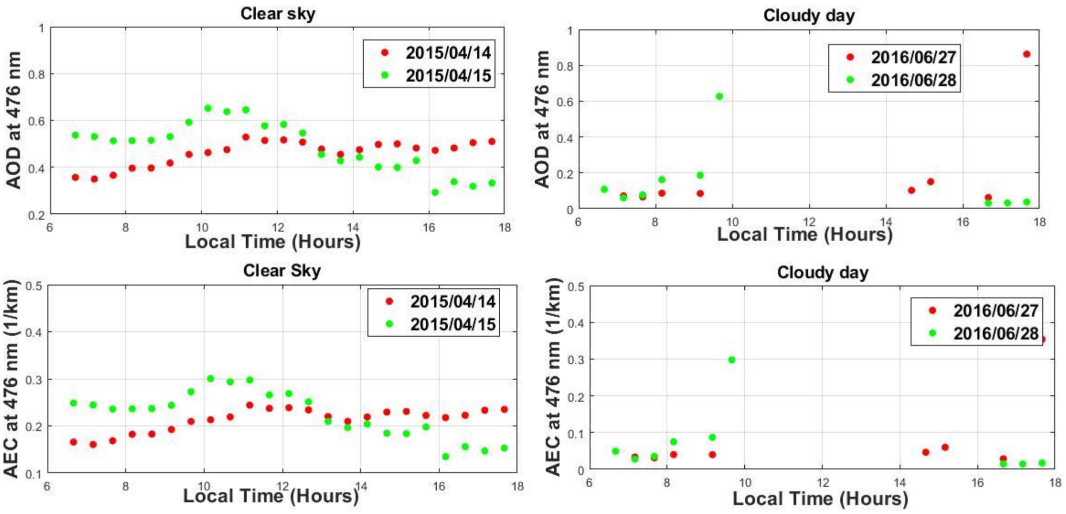


Figure 3.16. (top panel) AOD at 476 nm on two clear sky (left panel) and cloudy days (right panel). (bottom panel) AEC at 476 nm on the same clear sky (left panel) and cloudy days (right panel). The clear sky and cloudy days were distinguished using the sky view camera images and coincident LIDAR observations.

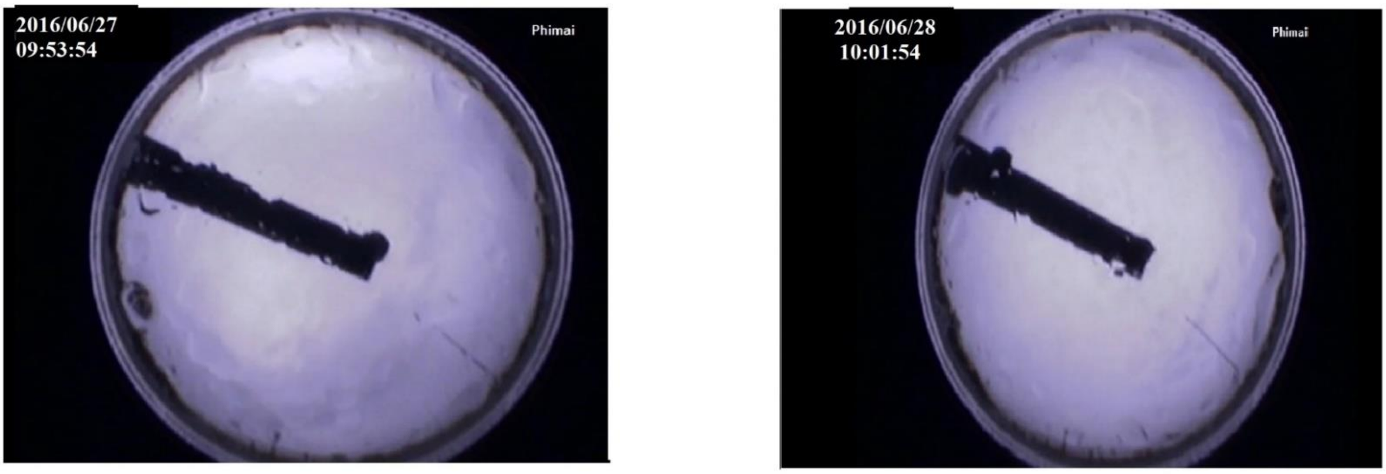


Figure 3.17. Skyview images on June 27 and 28, 2016 showing overcast and rainy conditions.

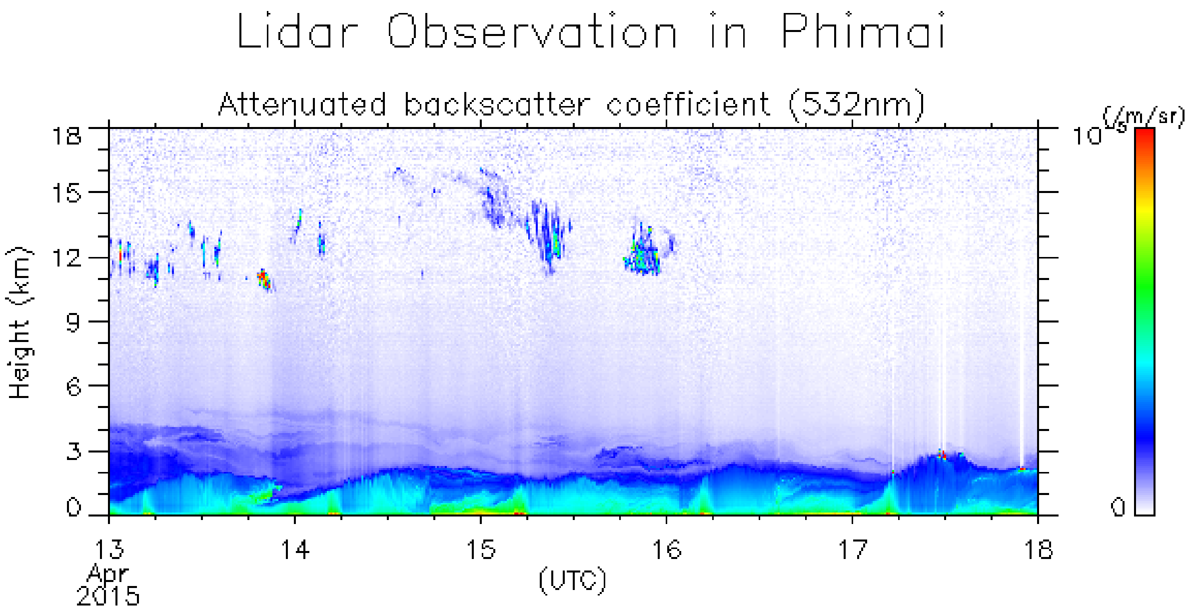


Figure 3.18. Coincident LIDAR measurements in Phimai showing clear sky conditions on April 14 and 15, 2015.

To assess the influence of clouds above 2km, we compared our water vapor measurements (H_2O) with the microwave radiometer (MWR) water vapor measurements. The idea behind such comparison is that, water vapor profile is well known and can be measured with high precision with different instrument. In the first step of the comparison, the MAX-DOAS precipitable water content (PWC) (to be explained in the next section) is compared with the MWR PWC for a particular period. Secondly, some cloudy and non-cloudy days are selected within the same period. The cloudy and non-cloudy days were distinguished utilizing the coincident LIDAR measurements. In the next step MAX-DOAS PWC is compared with the MWR water vapor columns for both the cases (cloudy and non-cloudy). The H_2O retrieval was performed simultaneously with the other trace gases. The fitting window of 495-415 nm was used for the retrieval of H_2O VMR values [*Irie et al.*, 2011]. For this comparison we used the MAX-DOAS measurements in Chiba (35.62° N, 140.1°E). The MAX-DOAS measurements were conducted using the similar settings used for Phimai. The limiting factor of this approach is that, good agreement between MAX-DOAS and MWR is observed when there is less water vapor above 2km, above which MAX-DOAS observation is weak. Measurements at our Phimai site did not match such criterion. So, we have used the MAX-DOAS measurements in Chiba for such evaluation.

The microwave radiometer used in this study is Radiometrics WVR-1125. Such instruments measure down welling radiation at 23.8 GHz, which is on the wing of the 22.3 GHz water vapor absorption line [*Perez-Ramirez et al* 2014]. The detail description of the instrument, absorption and emission bands, calibration and retrieval can be found elsewhere [*Cadeddu et al* 2013; *Westwater* 2013]. The WVR-1125 measurements were validated against other two microwave radiometers (MP1502, MP1504) during an intensive campaign in Chiba. Data of

MWR are output as value of precipitable water content and units in centimeters (cm). The MAX-DOAS H_2O VCD were converted to PWC using the following formula

$$PWC[cm] = \frac{N[molecules.cm^{-2}]*M[g.mol^{-1}]}{N_A[molecules.mol^{-1}]*\rho[g.cm^{-3}]} \tag{3.8}$$

where N is the MAX-DOAX VCD, M is the molar mass of water (18gmol^{-1}), N_A is the Avogadro's number and ρ is the density of water (1gcm^{-3}). Figure 3.19 shows the PWC measured by MAX-DOAS and WVR-1125 from December 27, 2015 to January 1, 2016. During this period, qualitatively good agreement between the two H_2O products with high correlation (*R* =0.94) and the mean absolute difference of 0.07±0.08 cm was observed. The difference can occur if there are more water vapor above 2 km, as MAX-DAOS measurements are sensitive up to 2 km. Besides, the approximation of the H_2O a priori above 2 km can also lead to such difference.

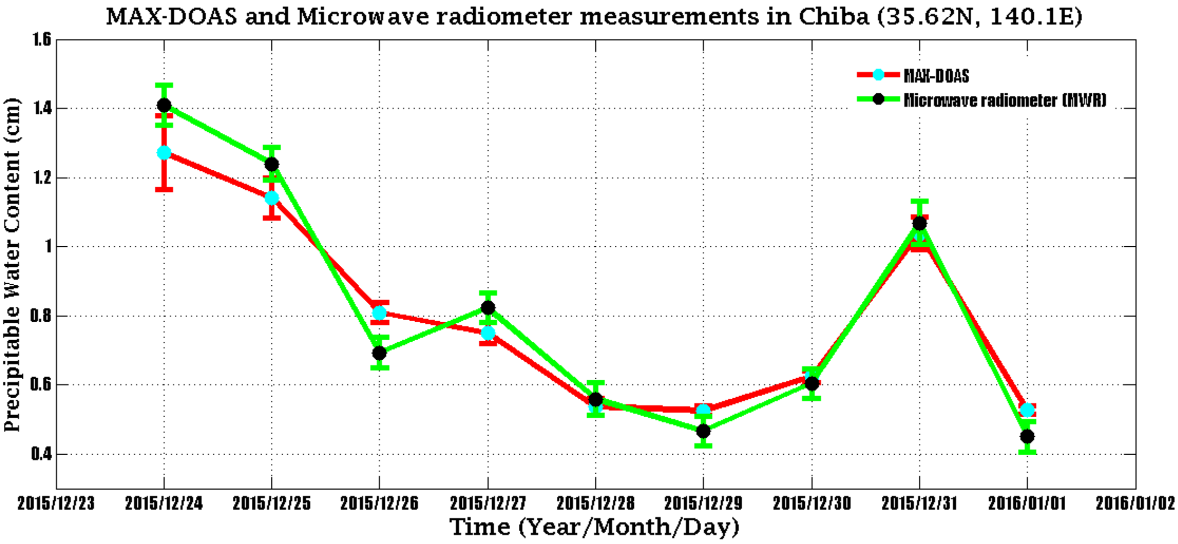


Figure 3.19. Precipitable Water Content (PWC) measured from MAX-DOAS and microwave radiometer in Chiba from December 25, 2015 to January 01, 2016

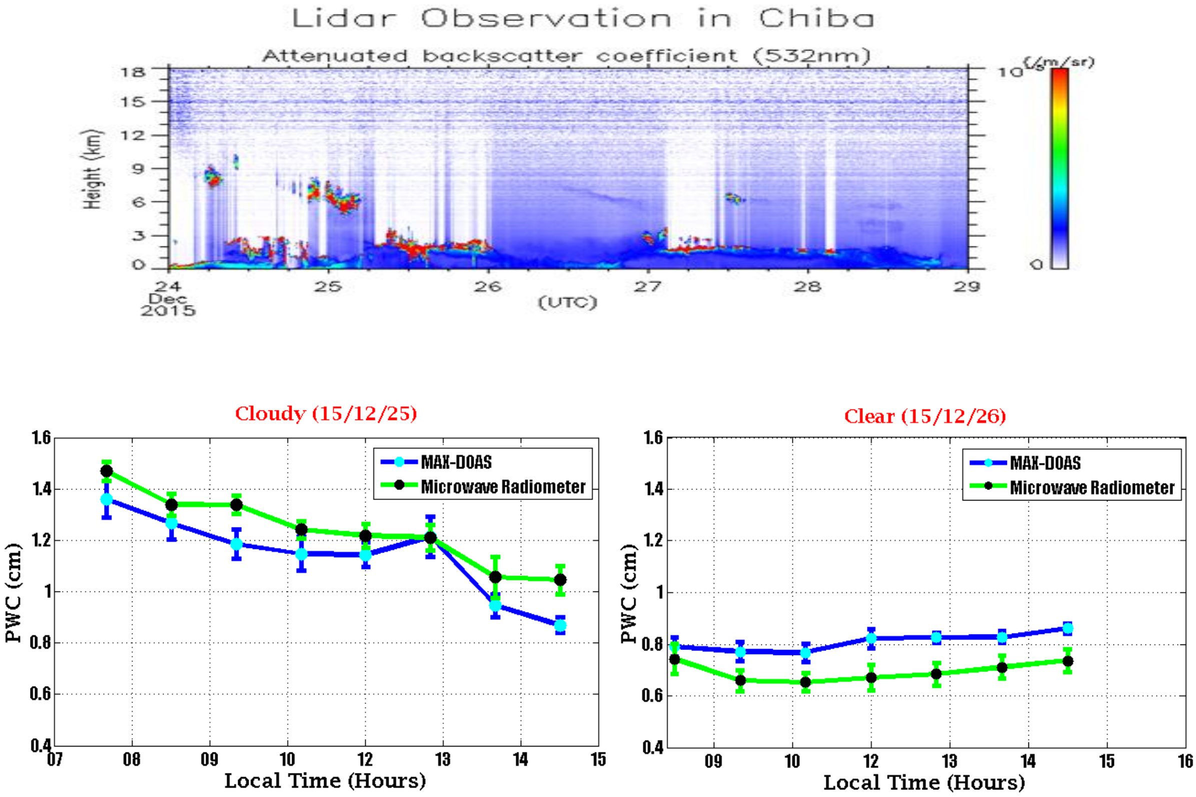


Figure 3.20. Selection of Cloudy and clear day based on coincident LIDAR measurements. The red dots indicate the cloud base heights.

Figures 3.20 and 3.21 shows the agreement between the two measurements on cloudy and non-cloudy conditions selected based on coincident LIDAR measurements in Chiba. On 25th December there were clouds above 3 km and 26th December was a clear day (refer to Figure 3.20). Similarly, 28th December was almost clear day with little clouds in the morning and on 31st December clouds were detected at a height of 3 km (Refer to Figure 3.21). Evaluating these cases, we find a good agreement between the two measurements even under cloudy conditions. The maximum difference occurring under cloudy condition is ~16% which is in the range of the uncertainty estimated for our retrieval. This indicate that, even for clouds above 2 km our trace gas has been retrieved in a reasonable manner. Thus, the data screening method has minimized the influence of the clouds on the retrieved products. Although, complete removal of cloud impact is not possible, but the error which can occur from the cloud influence is expected to be in the range of the total error estimated for the retrieved products.

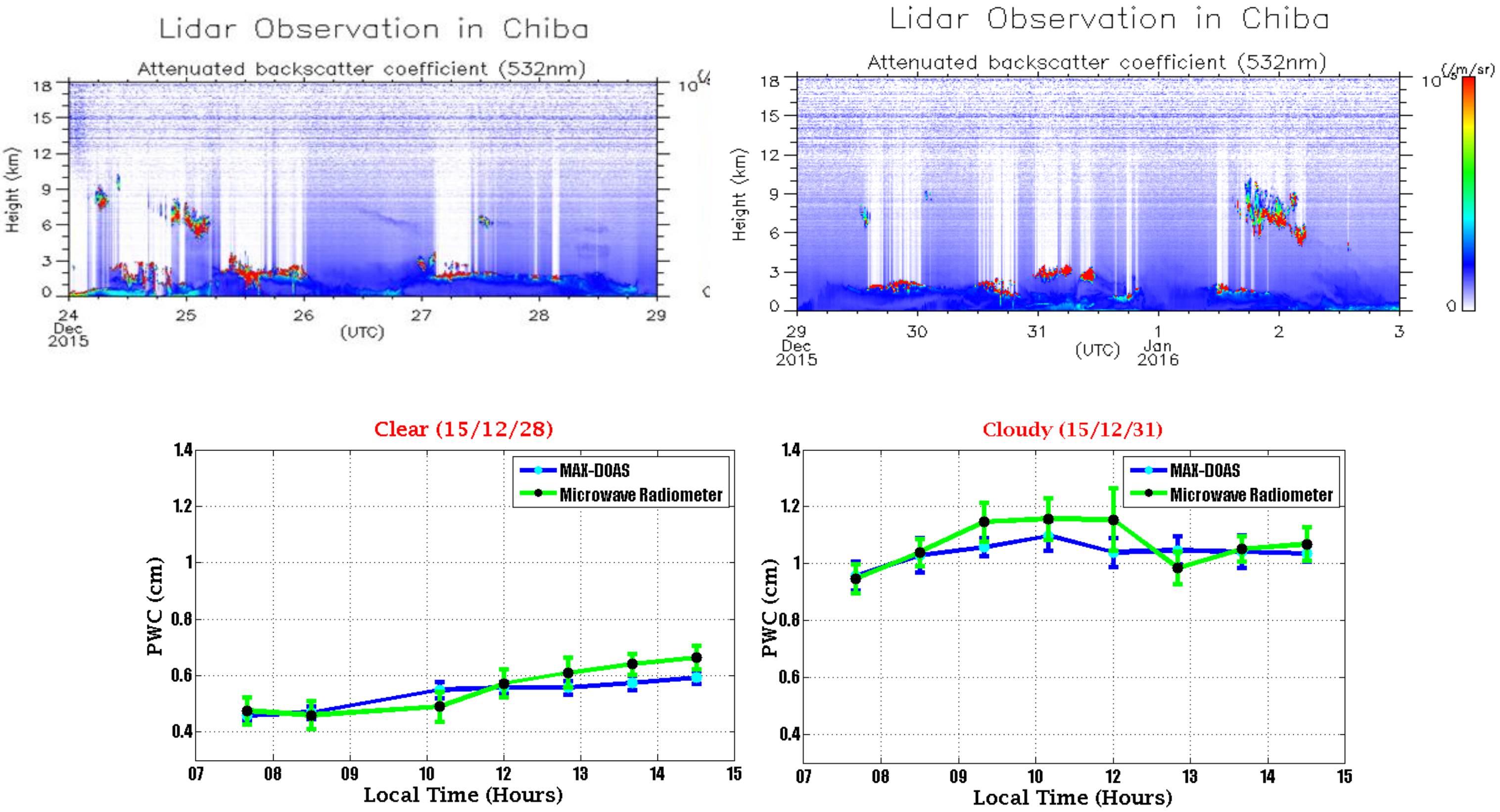


Figure 3.21. Selection of Cloudy and clear day based on coincident LIDAR measurements. The red dots indicate the cloud base heights.

3.4 Satellite Observations

(This section has been taken from section 2.1 of *Hoque et al* [2018a])

We used various satellite observations in this study to support our analysis. The satellite observations used in this study are: (1) level-3 monthly gridded carbon monoxide (CO) total column (TC) observed from the atmospheric infrared sounder (AIRS) aboard NASA's Aqua satellite [*McMillan et al.*, 2005]. The spatial resolution of the dataset was 1° × 1°, (2) fire radiative power (FRP) [*Justice et al.*, 2002], and (3) level-3 monthly gridded enhanced vegetation index (EVI) [*Justice et al.*, 1998; *Running et al.*, 1994] with a resolution of 0.05° × 0.05°. FRP and EVI products are derived from the Moderate Resolution Imaging Spectrometer (MODIS) observations. The CO TC (v0006) and EVI (v0005) were downloaded from NASA's earth data website (https://giovanni.gsfc.nasa.gov/giovanni/). FRP (MODIS Collection 6) data were downloaded from the NASA fire archive (https://firms.modaps.eosdis.nasa.gov).

CO is mainly emitted from incomplete combustion processes and is also produced from methane and VOC oxidization [*Crutzen et al.*, 1979; *Logan et al.*, 1981]. Pyrogenic and biogenic VOCs contribute to ~50% of global CO concentrations [*Logan et al.*, 1981; *Thompson et al.*, 1994]. The annual variation of CO is mainly governed by biomass burning events, extensive wildfires, and forest fires [*Yurganov et al.*, 2008]. The relationship between CO and biomass burning has been studied by various researchers. [*Galanter et al.,* 2000; *McMillan et al.,* 2005; *Warneke et al.,* 2006; *Yurganov et al.,* 2008].

The MODIS FRP product is a measure of emitted radiant energy from fires [*Justice et al*., 2002]. The MODIS algorithm utilizes the 4-µm channel measuring the brightness temperature of fire pixels and the background surface temperature of the neighboring pixel [*Kaufman et al*., 1998]. Over strong biomass burning areas, there is a strong link between the seasonal variation of tropospheric NO_2 and FRP values [*Schreirer et al.* 2014].

Vegetation index (VI) products from MODIS are used to extract information on spatial and temporal variation of global vegetation conditions, which are mainly used to monitor photosynthetic activities [*Justice et al.,* 1998; *Running et al.,* 1994]. The normalized difference vegetation index (NDVI) is more sensitive to chlorophyll amounts in the plants. EVI is more responsive to the structural variation of the vegetation areas [*Gao et al.,* 2000]. A good positive correlation among EVI, HCHO, and CHOCHO VCDs were reported by *Vrekoussis et al.* [2010].

3.5 Results and discussion

3.5.1 Seasonal variations of HCHO, CHOCHO, and NO_2

In this section, NO_2 data are also analyzed and discussed in addition to the HCHO and CHCOHO measurements. To evaluate the impact of anthropogenic emissions on R_{GF}, NO_2 was used as a tracer. NO_2 was retrieved simultaneously with the VOC tracers (i.e., HCHO and CHOCHO) The fitting window of 460-490 nm was utilized to retrieve the NO_2 VMR values [*Irie et al.,* 2011].

The seasonal variation in the trace gas (HCHO, CHOCHO, and NO_2) concentrations for the 0-1-km layer is shown in Fig. 3.22. The seasonal variations in HCHO and CHOCHO were identical with higher values during the dry season (January – April) and lower values during the wet season (June-September). The highest HCHO concentrations (~ 8 ppbv) were observed in March and the peak during the month of March was consistent in 2015 and 2016. The highest CHOCHO concentrations (~2 ppbv) were also observed during March in both the years. There was not much variation in the HCHO and CHOCHO concentrations during the wet seasons and the mean concentrations were ~ 2 and ~0.05 ppbv, respectively. NO_2 seasonal variations also showed higher values during the dry period and lower during the wet period. The highest concentration of NO_2 in Phimai was ~1.3 ppbv during the dry seasons. During the wet season, the mean NO_2 concentration was ~0.70 ppbv.

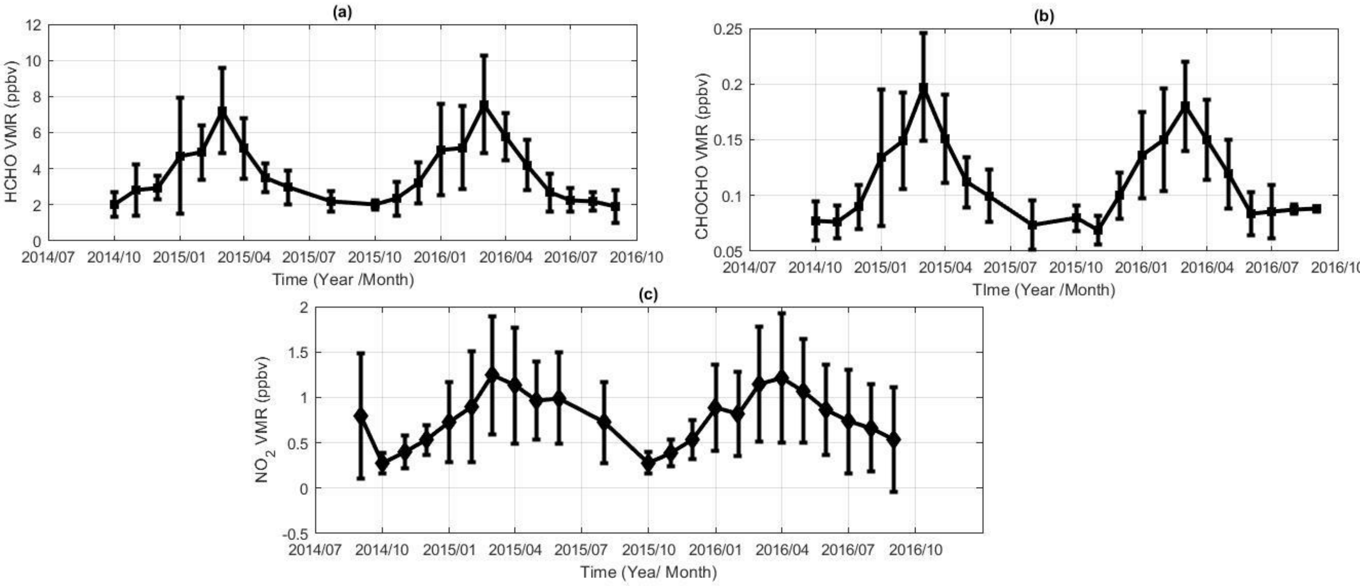


Figure 3.22. The seasonal variations of (a) HCHO, (b) CHOCHO, and (c) NO_2 VMR data for the 0-1 km layer. The error bars in the plots are showing the standard deviation of the mean values in the plot.

Several studies [*i.e. Campbell et al.,* 2012; *Li et al.,* 2013b; *Sugimoto et al.,* 2015] have found evidence of the biomass burning events in Phimai during the dry season. To further confirm the biomass burning influence, CO TCs from AIRS observation over central Thailand covering our observation period was used, which are shown in Fig. 3.23. The CO column over Phimai during the dry season was higher compared to the wet season. Because CO has a strong link with biomass burning [*Galanter et al.,* 2000; *McMillan et al.,* 2005], higher CO columns during the dry season is likely indicative of strong fire activities in Phimai. In addition to the CO TCs, FRP values from MODIS observations were also used which is shown in Fig 3.24. The FRP values within the area of 100 km from the center of Phimai site were used to calculate the daily mean values. During both dry seasons higher FRP values were observed. Within the area of 100 km centered at Phimai, ~580 and ~650 fires were detected in 2015 and 2016, respectively. The higher number of fires during the dry seasons were consistent with the higher CO columns during the similar periods. Utilizing satellite measurements, Alvarado *et al.* [2014] found strong correlation between CHOCHO VCDs and FRP values over region where the dominant source of CHOCHO was pyogenic emissions. Adapting the similar procedure, the correlation between CHOCHO VCD and FRP values were calculated for the Phimai site and a good positive correlation (*R* = 0.80) was observed (Fig. 3.24(b)), which was consistent with the results of Alvarado *et al.* [2014].

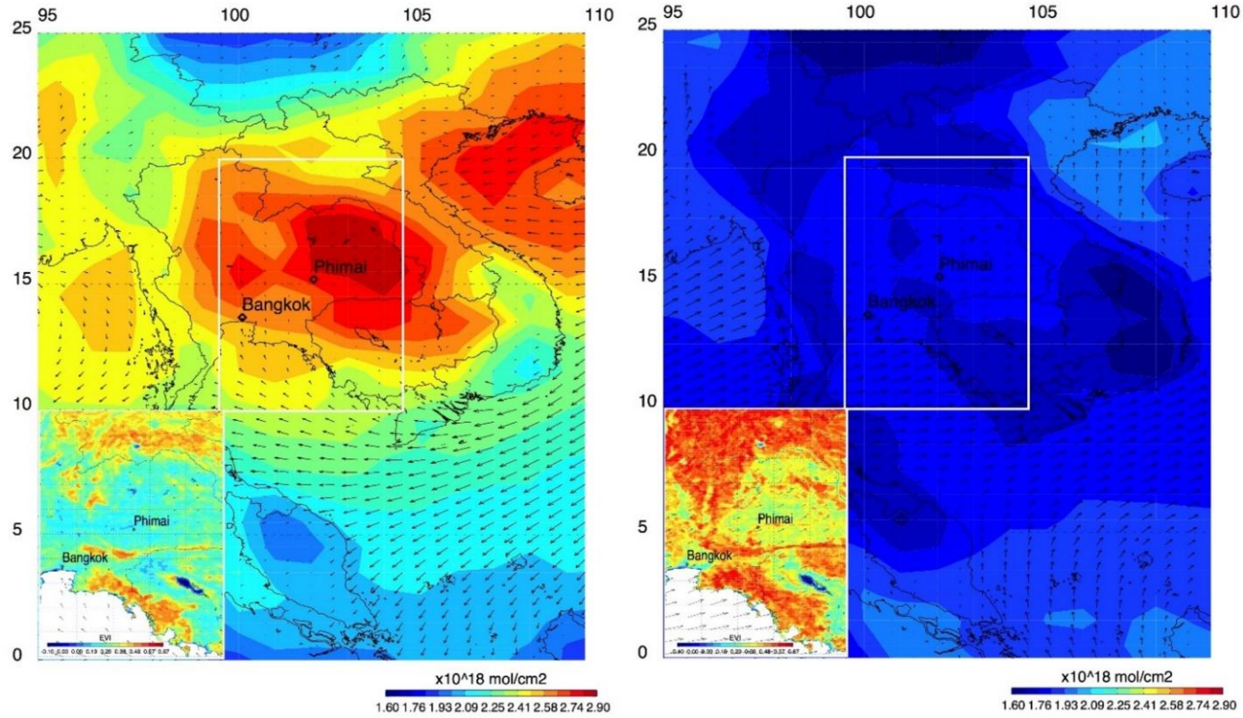


Figure 3.23. The AIRS satellite observations of carbon monooxide (CO) total columns (TC) over central Thailand covering the period of our MAX-DOAS observations in Phimai. The plot in the left panel are showing the CO TCs during the dry seasons and the right panel displaying the CO TCs for the wet season. The Enhanced vegetation index (EVI) data from MODIS observations for the dry and wet seasons are shown in the insets of the left and right panel, respectively. The black arrows in the plots are showing the wind direction and intensity at Phimai during the dry and wet seasons.

A study by *Schreirer et al.* [2014] utilized satellite observations and showed strong relationship between the FRP values and NO_2 VCD with correlation coefficient between 0.6 – 0.8 over southeast Asia. Similar to the study of *Schreirer et al.* [2014] the correlation between NO_2 VCD and FRP values were calculated and a moderate correlation of 0.61 (Fig. 3.24 (c)) was observed which was in the range of the reported literature values. The CO TCs, FRP values and the correlation among FRP, CHOCHO VCD, and NO_2 VCD provide evidences of biomass burning in Phimai during the dry seasons.

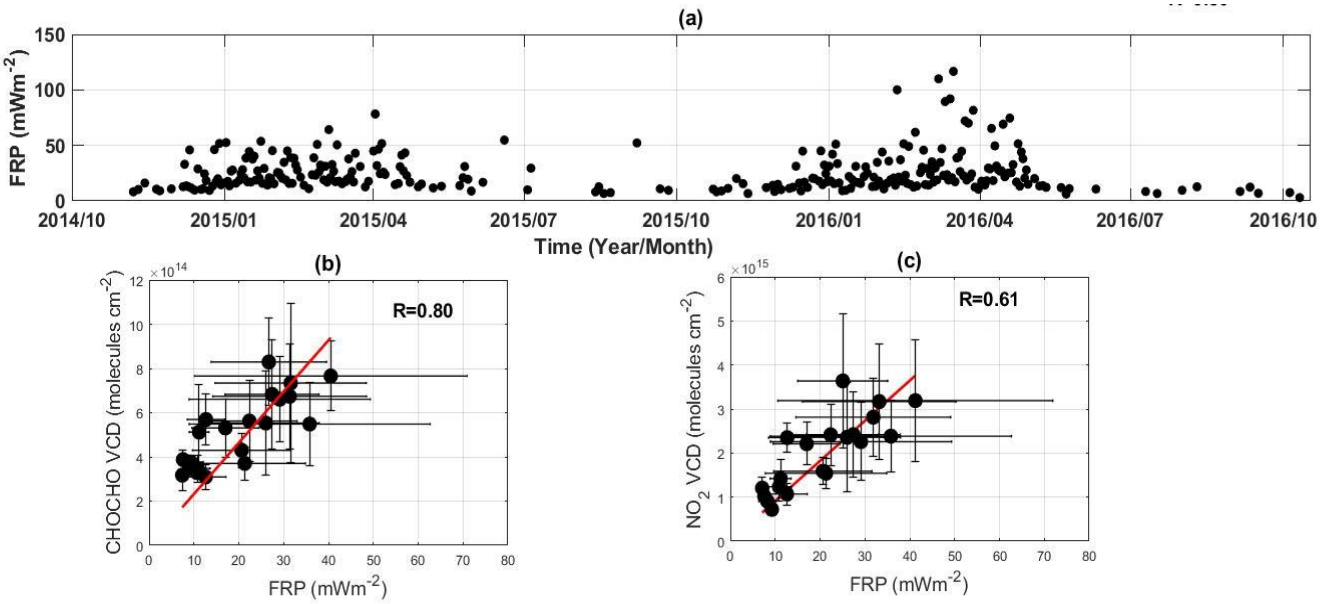


Figure 3.24. (a) Time series of the daily mean FRP values over Phimai covering the period of our MAX-DOAS observations (i.e., Ocotober 2014- September 2016). The FRP values were averaged within an area of 100 km centring at the Phimai station. (b) Correlation plots between FRP and CHOCHO vertical column densities. (c)Correlation plots FRP and NO_2 vertical column density. In both the cases of (b) and (c) the FRP values were averaged within an area of 100 km centring at the Phimai station. The red lines in the correlation plots are indicating the regression line

3.5.2 Backward trajectories

Additional analysis using backward trajectories was performed to further support the link between biomass burning and the higher trace gas concentrations during the dry season in Phimai. The influence of the emissions in Bangkok was also assessed using the trajectory analysis. The online trajectory model HYSPLIT [*Stein et al.*, 2015] was used for such analysis. Figure 3.25 depicts backward trajectories and the trace gas concentrations on selected days during the dry and wet season in Phimai. The trajectory calculations during the dry season were initiated at 10:00 UTC at an altitude of 200 m above the surface of the Phimai site and was calculated for 24-h in the backward direction. The trajectory calculations during the dry season were performed for January 11 and March 22, 2016. The concentrations of HCHO, CHOCHO, and NO_2 on the respective dates are shown in the lower panel of Fig. 3.25. The trajectory calculations for the wet season were also similar except the trajectories were calculated for 48-h in the backward direction. This was mainly done to assess the influence of emissions in Bangkok at Phimai. The example dates for the wet season were June 11 and 12, 2016. The selection of the example dates of the dry season was based on the mean FRP values and the intensity of fire activities with an area of 100 km around Phimai. During the wet season, the example dates were selected depending on the air masses travelling close to or over Bangkok or not. On March 22, 2016, several fires with mean FRP values ~ 22 mWm^{-2} were observed within 100 km of Phimai. With winds blowing from the southwest direction, prior to the arrival at Phimai, the trajectories are seen to travel over the fire hotspots regions.

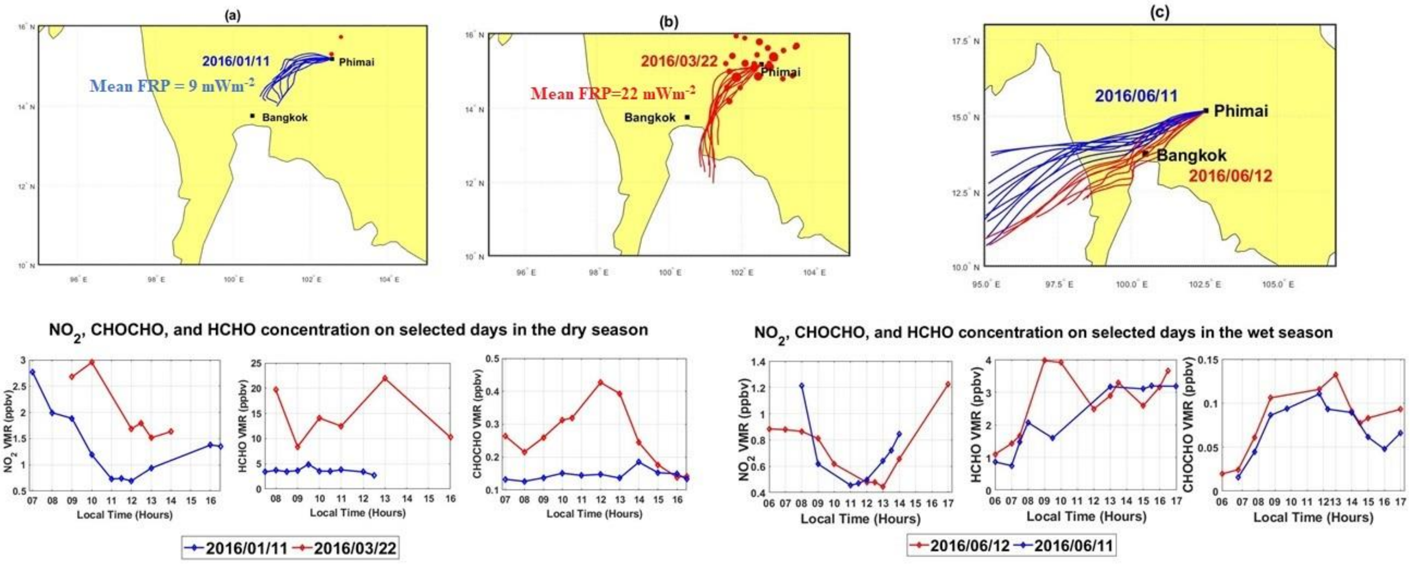


Figure 3.25. (top panels) (a) The backward trajectories initiated at Phimai site on 11 January, 2016, at 10:00 UTC at an altitude of 200 m above the surface. The trajectories were run for 24 hours in the backward direction. (b) The similar backward trajectory simulation conducted for March 22, 2016. The red mark dots in (a) and (b) are indication of the location of fires within an area of 100 km from the Phimai site. (c) The similar backward trajecotry simulation on June 11 and 12, 2016. These trajectories were simulated for 48 hours in the backward direction to assess the influence of the emissions in Bangkok at Phimai. (lower panels) The dirunal mean concnertration of NO_2, CHOCHO, and HCHO on the respective selected dates of the dry and wet season in 2016.

The higher concentrations of the trace gases (HCHO, CHOCHO, and NO_2) on March 22, 2016 are also potentially linked to the air masses travelling over fire active regions. On January 11, 2016, there were only few fires around Phimai. The mean FRP value on this day was ~9 mWm^{-2}. Westerly winds prevailed and likely due to weaker winds on this day the air masses from short distance arrived at Phimai. The lower HCHO, CHOCHO, and NO_2 concentrations on this date compared to March 22, 2016 were congruent with the few amount of fires. Similar response of the HCHO, CHOCHO, and NO_2 concentrations to the fire activities were also seen in the similar type of example shown in Fig 3.26. Thus, the enhancements observed in the trace gas concentrations in Phimai were likely due to biomass burning.

On the selected example dates of the wet season, the dominant winds were mostly blowing from the southwest direction. On June 11 and 12, 2016, air masses travelled near and over Bangkok, respectively, prior to arrival at Phimai. Comparing the concentrations of CHOCHO, HCHO, and NO_2 on the selected dates, no significant differences were observed.

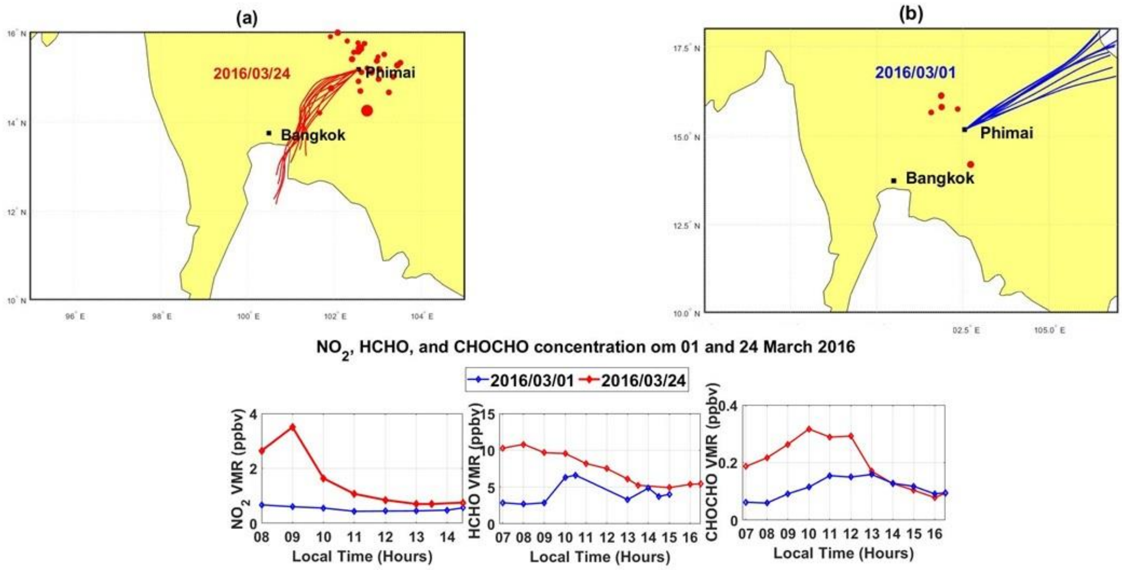


Figure 3.26. (a)) The backward trajectories initiated at the Phimai site on March 24, 2016, at 10:00 UTC at an altitude of 200 m above the surface. The trajectories were run for 24 hours in the backward direction. (b) The similar backward trajectory simulation conducted for March 1, 2016. The red mark dots in (a) and (b) are indicating the location of fires within an area of 100 km from the Phimai site. The concentration of HCHO, CHOCHO, and NO$_2$ on the respective days are shown in the lower panel. The dates were selected based on the FRP and fire location within 100 km centred at Phimai. The mean FRP values on March 1 and 24 were ~14 and ~ 35 mWm^{-2} respectively.

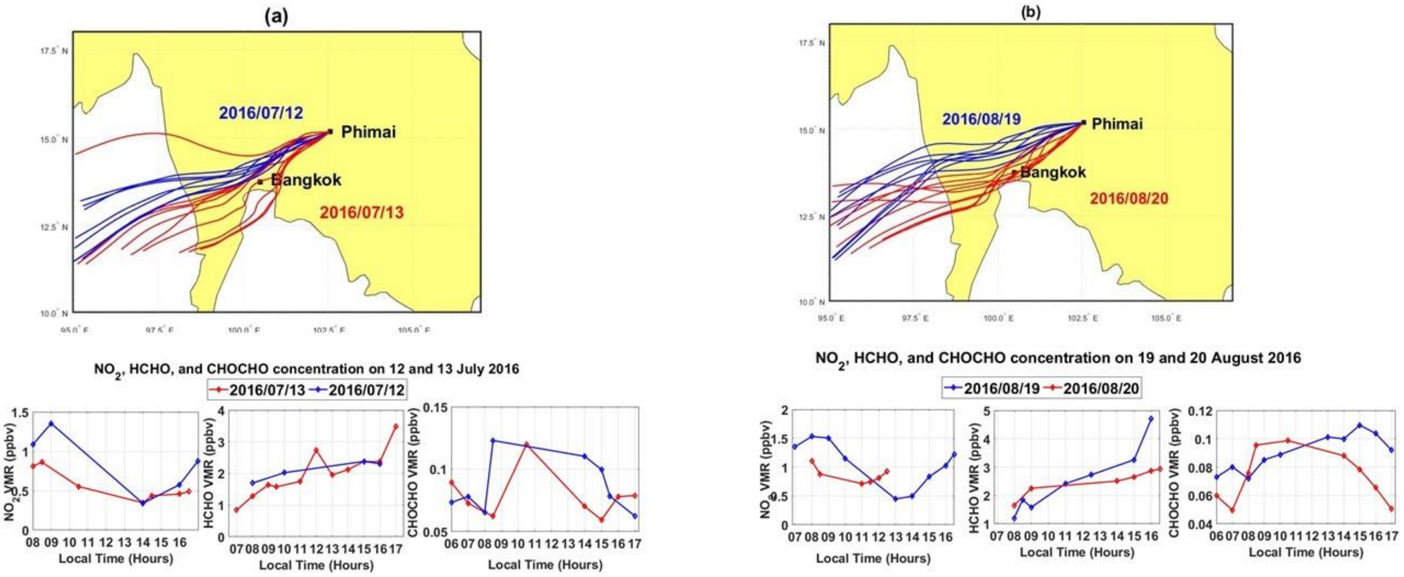


Figure 3.27. 48 hours backward trajectories initiated from the Phimai site at 10:00 UTC at an altitude of 200 m from the surface on (a) June 12 and 13, 2016 and (b) August 19 and 20, 2016. Hourly concentrations of CHOCHO, HCHO, and NO$_2$ on the respective days of (a) and (b) are shown in the lower panels. The trajectories were selected based on whether the air

massess crossed over Bangkok or not before reaching Phimai. In (a) on June 13 airmasses crossed over Bangkok, whereas on June 12 the airmasses travelled away from Bangkok before reaching Phimai. Similarly in (b)on August 19 airmasses were far away from Bangkok before reaching the Phimai station, whereas on August 20 trajectories tarvelled over Bangkok.

Fig. 3.27 shows similar example of the trajectories travelling over and close to Bangkok prior to their arrival at Phimai and no significant differences in the trace as concentration were observed. Therefore, the emissions in Bangkok are expected to have low impact on the trace gas concentration level at Phimai. During the transition period in the dry season, wind could also blow from the southwest direction. Because during this period the dominant northeasterly wind started changing the direction toward southwest and air masses from Bangkok could arrived at Phimai during the dry season. However, as significant fire events were observed around Phimai during the dry season, the fire events are expected to have stronger influence on the trace gas concentrations compared to the emissions in Bangkok.

3.5.3 Potential influence of different VOC sources

HCHO and CHOCHO are primarily emitted from biogenic sources. Thus the impact of biogenic emissions on the concentrations of HCHO and CHOCHO during both season in Phimai was assessed. For such assessment EVI data over Phimai during the dry and wet season obtained from MODIS observations were used which are plotted in the inset of Fig. 3.23. High EVI index values indicate large amount of vegetation. As seen from Fig 3.23, the EVI index values are higher during the wet season compared to the dry season. This indicates that, the amount of vegetation in Phimai during the dry season is low compared to the wet season. The seasonal variation of the EVI index is opposite to the seasonal behavior of HCHO and CHOCHO concentrations. Moreover, there were almost no fire events during the wet season in Phimai. Thus, very low fire activities and higher EVI values during this period were indicative of a dominant biogenic sources of HCHO and CHOCHO.

HCHO and CHOCHO are produced from photolysis of VOCs. Thus the impact of the temperature on the concentration of the both trace gases was investigated. The variation of monthly mean temperature in Phimai is shown in Fig. 3.28. During the month of April, the temperature was highest, ~30°C. The mean temperature during the wet season was ~25°C whereas the mean temperature during the dry season was slightly higher ~27°C. Thus the mean temperature contrast between the two seasons in Phimai is not very high. Moreover, poor correlations among CHOCHO and HCHO concentrations and temperature were observed (Fig. 3.28). *Stavrakou et al.* [2014] estimated the isoprene (a VOC precursor) emissions over Southeast Asia using an isoprene

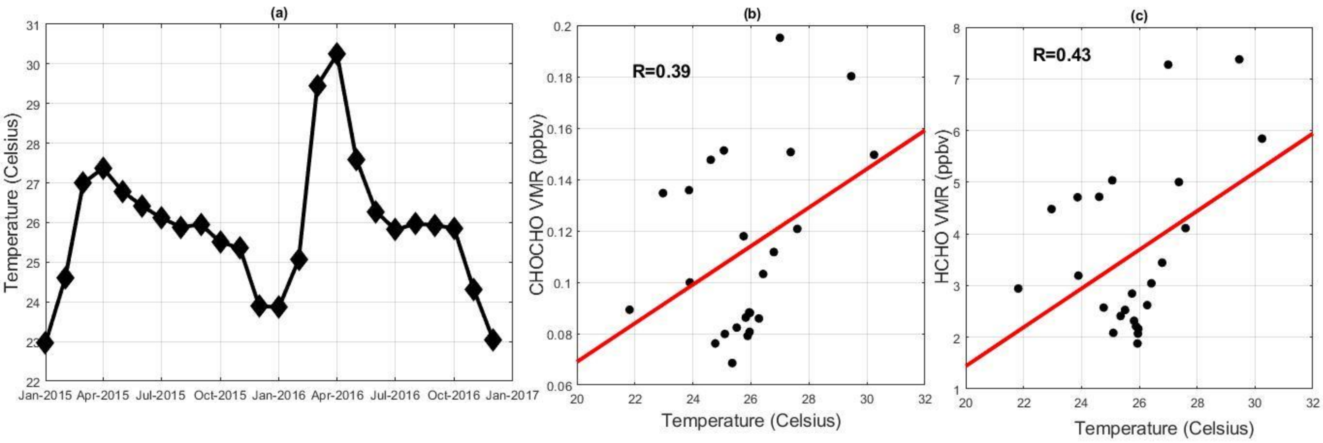


Figure 3.28. (a) The monthly variation of temeprature at the Phimai site. (b) Correlation plots showing the correlation between temperature and CHOCHO cocnetration for the 0-1 km layer. (c) Correlation plots showing the correlation between temperature and HCHO cocnetration for the 0-1 km layer. All the data covering our MAX-DOAS observation period are used in the correlation plots. The red lines are indicating the regression lines.

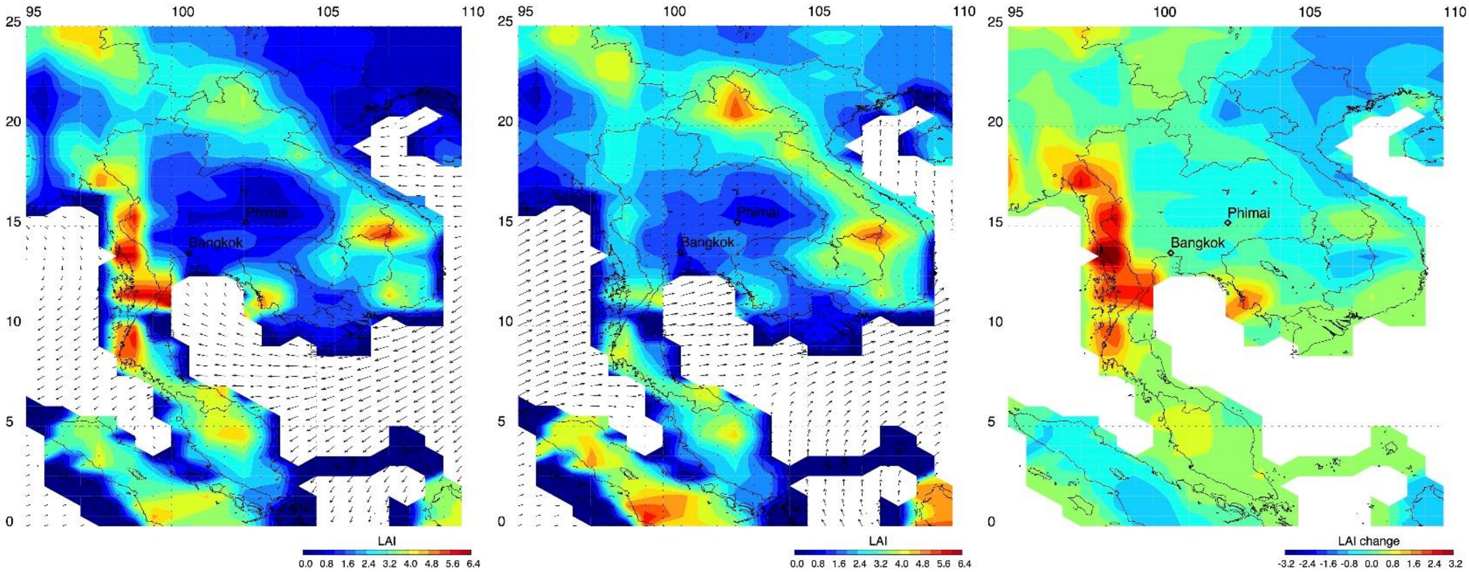


Figure 3.29. Leaf area index (LAI) over Phimai in dry and wet season. The left and middle panel shows the LAI in the dry and wet season over Phimai. The right panel shows the difference in the absolute LAI values between the two the seasons.

emission model, and found that the isoprene emissions in Thailand were strongly correlated with temperature. Especially the high isoprene emissions showed strong correlation with higher temperatures during the dry season. However, the model results showed almost similar isoprene emissions throughout the year around Phimai indicating similar biogenic emission strength during both seasons. Similarly, the simulations from global model MERRA-2 [*Bosilovich et al.*, 2015] also showed no significant variations in the leaf area index throughout the year in Phimai (Fig. 3.29), indicating almost similar amount of vegetation. The similarity of the isoprene emissions during the dry and wet seasons appears to be in agreement with the low temperature contrast between the two seasons in Phimai. In principle, this contradicts with the observed seasonality in EVI over Phimai, which shows significant differences between the two seasons. Thus the LAI and the EVI index in Phimai show contradictory results. However, various studies [*i.e., Cohen et al.,* 2003; *Davi et al.*, 2006; *Son et al.,* 2013; *Wang et al.,* 2005] have reported that the response of NDVI, EVI, and LAI can be different depending on the phenological conditions of the plants in response to different climate conditions. According to the simulation results of the isoprene emission model (i.e., *Stavrakou et al.,* 2014), the biogenic emission sources were almost similarly strong throughout the year in Phimai. On the other hand, the CO columns, FRP values, and trajectory analysis showed that pyrogenic sources (biomass burning) have strong influence on the HCHO and CHOCHO concentrations in Phimai. As the strength of the biogenic emissions is almost similar and the temperature contrast between the seasons are low, the enhanced concentrations of HCHO and CHOCHO during the dry seasons are likely to be linked to biomass burning. Such enhancements due to biomass burning are likely to be responsible for the low correlation among HCHO, CHOCHO, and temperature.

3.5.4 Diurnal variations of HCHO and CHOCHO

In addition to the seasonal variation, the diurnal variation of the trace gases was also analyzed. The diurnal variations in the trace gas (CHOCHO, HCHO, and NO_2) concentrations for the 0-1-km layer during the dry and wet seasons are shown in Fig. 3.30. The diurnal variation of CHOCHO showed lower values in the early morning and late afternoon during both seasons. During the both dry seasons, the peak concentration (~0.24 ppbv) of CHOCHO was observed around 11:30 LT and the daily mean concentration was mostly above ~0.1 ppbv. During the wet seasons, peaks in diurnal variation in CHOCHO concentrations were also observed, but at different time. Due to the lack of measurements during the wet season in 2015, it was not possible to confirm the coherency in the peaks during the wet season. Thus, the diurnal variations in CHOCHO concentrations for individual months during the wet season in 2016 were analyzed. Figure 3.30 (c) shows the diurnal variations in CHOCHO concentrations during the months of the wet season in 2016.

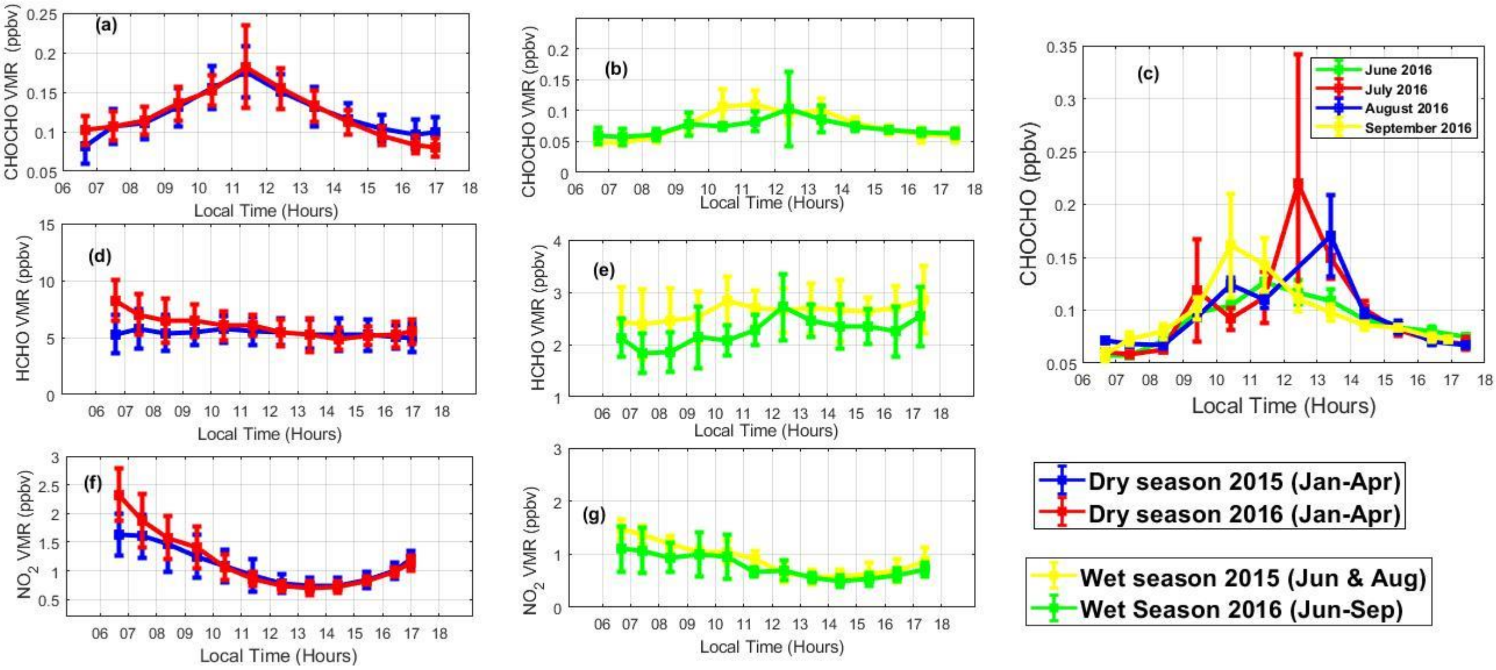


Figure 3.30. The diurnal variation of NO_2, CHOCHO, and HCHO concentrations for the 0-1 km layer in Phimai. The error bars show the standard deviation of the mean values. (a) The diurnal variation of CHOCHO during the dry seasons in 2015 and 2016, (b) the diurnal variation of CHOCHO during the wet seasons in 2015 and 2016, (c) the diurnal variation of CHOCHO during the month of the wet season in 2016, (d) the diurnal variation of HCHO during the dry seasons in 2015 and 2016, (e) the diurnal variation of HCHO during the wet seasons in 2015 and 2016, (f) the diurnal variation of NO_2 during the dry seasons in 2015 and 2016, and (g) the diurnal variation of NO_2 during the wet seasons in 2015 and 2016.

As seen from Fig 3.30 (c) the occurrence of the peak in the CHOCHO diurnal variation was different during the months. The reason behind the indistinct peaks in the CHOCHO diurnal variation during the months of wet season 2016, might be the less number of measurements under clear sky condition. After performing data screening as explained in section 3.3.4, on average ~580 and ~336 measurements were available during the months of the dry and wet seasons in 2016, respectively. During the month of March, the highest number (N = 645) of data points was obtained whereas, the lowest number of data points (N = 264) was seen in July, 2016. This indicates that due to rain and overcast conditions in the wet season, the number of measurements under clear sky conditions was small. The hourly mean CHOCHO concentrations during the dry seasons were higher than that of the wet seasons. This indicates the CHOCHO emission sources were stronger during the dry season. This is likely due to the significant fire events around Phimai during the dry season.

During the two dry seasons, the hourly HCHO concentration did not show any significant variation throughout the day and the concentration level was ~5 ppbv (Figs. 3.30(d) and 3.30(e)). However, during the early morning hours of the dry season in 2016, high HCHO concentrations of ~8 ppbv were observed. But such high HCHO concentrations during the early morning hours were not observed in the dry season of 2015. The reason of such high concentrations in the early morning hours is not clear but could be potentially linked to the fire actives occurring during the night time. During the wet seasons, the hourly mean HCHO concentrations were mostly below 3ppbv and the HCHO diurnal concentration did not show any significant variations. Similar to CHOCHO the hourly HCHO concentration in the dry season was higher (~ 2-3 times) compared to the wet season. This is also likely due to the impact of the frequent fire activities during the dry season.

3.5.5 Probable sinks of HCHO and CHOCHO in Phimai

Photolysis and OH oxidization are likely to be the dominant sink of HCHO and CHOCHO. During the dry season, wet deposition could have occurred due to higher rainfall, in addition to photolysis and OH oxidization. According to *Volkamer et al.* [2007] CHOCHO is more likely to be uptaken by aerosols during the nighttime compared with the daytime. The aerosol deposition rate as well as the Henry's law coefficient of CHOCHO is higher compared with that of HCHO [*Zarzana et al.,* 2017 and the additional references mentioned]. Several studies [i.e. *Akagi et al.,* 2011, 2012; *Zarzana et al.,* 2017; *Yokelson et al.,* 2009] have found that particles emitted from biomass burning events are less hygroscopic compared to the other types of aerosol particles. Because the dominant CHOCHO emission sources during the dry season is biomass burning, aerosol deposition might not be a strong sink of CHOCHO during the dry season.

3.5.6 Diurnal Variation of NO_2

The diurnal variation of NO_2 (higher values in the morning hours and lower in the afternoon) was similar during the dry and wet seasons. Such diurnal variations are commonly observed in rural sites and mostly related to the change in the boundary layer height (Figs. 3.30 (f) and 3.30 (g)). The hourly NO_2 concentration during the noon time (13:25 LT) was ~ 0.6 and ~0.5 ppbv during the dry and wet seasons. The daily mean NO_2 concentrations were mostly below ~1 ppbv during both the seasons. Using ship-borne MAX-DOAS observation during a cruise over the Pacific Ocean *Takashima et al.* [2012] found NO_2 concentration of ~0.2 ppbv with the highest concentration of ~0.4 ppbv. These concentrations were also reported for the 0-1 km layer identical to the definition used in this analysis. Thus the mean NO_2 concentration observed in Phimai during the dry and wet seasons can be compared with that over the Pacific Ocean. Although the site was affected by biomass burning during the dry season, the mean NO_2 concentration during this period seems to be low. This could also be due to the age of the plumes impacting the site or intercepted by the instrument. An example case is shown in Fig. 3.31, where the FRP values and backward trajectories are used to assess the aging impact of the plumes. Backward trajectories were calculated for March 19 and February 2, 2016. On these dates fires of similar FRP values of ~30 mWm^{-2} were observed within an area of 10 and 30 km, respectively, around the Phimai site. Although on March 19, more fires were observed around Phimai, but the locations of the fires were far away from the site. As seen from Fig. 3.31, prior to the arrival at the site, the trajectories travelled over the locations where the fires were located on both dates. The concentrations of all the three trace gases (HCHO, CHOCHO, and NO_2) on March 19 were higher compared to the concentrations observed on February 2. Although fire of similar power was detected on both dates, the distance of the fires could have driven the difference in the trace gas concentrations on the respective dates.

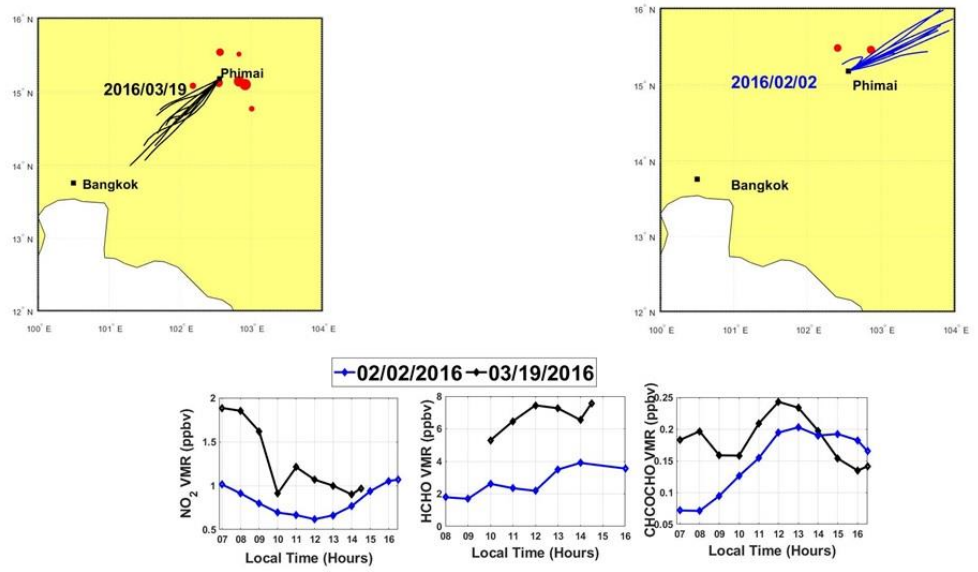


Figure 3.31. The backward trajectories initiated at the Phimai site on March 19, 2016, at 10:00 UTC at an altitude of 200 m above the surface. The trajectories were run for 10 hours in the backward direction. (b) The similar backward trajectory simulation conducted for February 2, 2016. The red mark dots are indicate the location of fires within an area of 100 km from the Phimai site. The concentration of HCHO, CHOCHO, and NO_2 on the respective days are shown in the lower panels.

3.5.7 Estimation of the R_{GF}

The formation mechanisms of HCHO and CHOCHO are different, which can be used to characterize the VOC emission sources, by calculating the R_{GF} value (CHOCHO to HCHO ratio). Because R_{GF} compromises the contribution of both the HCHO and CHOCHO concentrations. The annual mean R_{GF} value estimated from the observations of HCHO and CHOCHO in Phimai was 0.032 ± 0.005. The annual mean R_{GF} value was estimated from the daily mean concentrations of HCHO and CHOCHO for the 0-1-km layer. The observations covering our analysis period (*i.e.*, October 2014 – September 2016) were used to calculate the R_{GF} value. The R_{GF} value was also estimated using the monthly mean values of HCHO and CHOCHO, which yielded the similar results. The uncertainty was calculated from the individual uncertainties, i.e., the systematic error, the random error, and the error related to the bias in ELs. According to Table 3.2, the estimated uncertainty related to the bias in the ELs for both the trace gases (HCHO and CHOCHO) was 5%. Thus in the ratio, such error will cancel out. The systematic errors in the both trace gases were estimated from the air mass factor calculations. Similar to the bias in the ELs, the systematic error will also cancel out largely in the ratio due to similar source of the systematic error. The calculated systematic error was approximately 16%, which is equal to the systematic error estimated for the CHOCHO profile retrieval (Table 3.2).

Table 3.4. Overview of R_{GF} values reported for rural sites in literature.

Location	Method	R_{GF}	Reference
Georgia, USA	In situ measurements	0.006	*Lee et al.* [1995]
Pinnacles, USA	In situ measurements	0.045	*Munger et al.* [1995]
San Nicolas, USA	In situ measurements	0.132	*Grosjean et al.* [1996]
Pattstthum, Germany	In situ measurements	0.012	*Moortgat et al.* [2002]
Anadia, Portugal	In situ measurements	0.044	*Cerqueira et al.* [2003]
Montelibretti, Italy	In situ measurements	0.107	*Possanzini et al.* [2007]
Sierra, USA	QCL[a] spectrsocopy	0.004	*Choi et al.* [2010]
Cabauw, Netherlands	MAX-DOAS	0.036	*Irie et al.* [2011]
Southern China	MAX-DOAS	0.062	*Li et al.* [2013b]
Ponderosa forests, USA	LIF[b] /LIP[c]	< 2 or > 2.5 depending on VOC emission sources	*DiGangi et al.* [2012]
SE USA	LIF/ ACES[d]	< 2.5 or > 3 depending on VOC emission sources	*Kaiser et al.* [2015]

[a] Quantam cascade laser . [b] Laser induced fluorescence . [c] Laser-induced phosphorescence. [d] Airborne cavity enhanced spectrometer.

On the other hand, as the random errors are estimated from the residuals of the fittings of the trace gases, their contribution will not be cancelled out, similar to the systematic errors. The random errors will contribute in the same order of magnitude in the ratio. However, as the ratio is calculated over a temporal scale of one year, the random error is also expected to reduce by a factor of $\sqrt{\frac{1}{n}}$, where n denotes the temporal scale (i.e., number of days used to estimate the ratio). Then, the total uncertainty in the calculated ratio will be the root sum square of all the individual errors (i.e., random and systematic error), which was calculated to be approximately 17%. The estimated error is expected to be the maximum uncertainty in the calculated ratio.

After estimating the R_{GF}, we compared our estimated R_{GF} value with values reported in literature for rural sites, summarized in Table 3.4. Comparing our value with other studies, we find that, our estimated value for the Phimai site is in the range of the values reported in the literature. The yield of HCHO and CHOCHO from different VOC emission sources can be different which can lead to the change of the R_{GF} values with the change in the composition of the VOCs. As observed from the seasonal variation in HCHO and CHOCHO, the dominant VOC sources were different during the dry and wet seasons. Thus, analogous to the individual seasonal variation in HCHO and CHOCHO, the seasonal variations in the R_{GF} values were assessed and the estimated R_{GF} values for during the dry and wet seasons are summarized in

Table 3.5. The estimated error in the seasonal mean R_{GF} values are the 2σ standard error and was estimated using the error propagation formula. The seasonal mean R_{GF} during the both dry season (i.e. 2015 and 2016) were 0.030 and 0.034, whereas the values during the wet seasons were 0.028 and 0.040 respectively. Utilizing the GOME-2 global observations, *Vrekoussis et al.* [2010] found R_{GF} values in the range between 0.04 and 0.06 for regions with dominant biogenic VOC emission sources. On the contrary, lower R_{GF} values mostly less than 0.03 were observed in regions where the dominant VOC emission sources were anthropogenic as well as the NO_2 concentrations were high.

Analyzing the CO TCs, FRP values, and backward trajectories at the Phimai site, a clear link was found among the high HCHO and CHOCHO concentrations and biomass burning during the dry seasons, which was also consistent with previous studies. Moreover, in addition to the high HCHO and CHOCHO concentrations during the dry seasons, the NO_2 concentrations were also higher. Therefore, the lower seasonal mean R_{GF} values observed during the dry seasons in Phimai are likely to be linked with the biomass burning events. The lower seasonal mean R_{GF} values in the dry season, also supports the results of *Vrekoussis et al.* [2010] which reported lower R_{GF} values for regions characterized with enhanced NO_2 level. In contrast, due to the dominant biogenic VOC sources during the wet season, higher seasonal mean R_{GF} values were observed during both wet seasons.

The change in the R_{GF} values during different seasons was also assessed during the different months of the dry and wet seasons in Phimai, depicted in Fig. 3.32. The error bars in Fig 3.32 indicates the 2σ standard error of the estimated monthly mean R_{GF} values. As seen in the seasonal variation in R_{GF}, the monthly mean R_{GF} values during the dry season months were significantly lower compared with the wet season months. On average, the difference between the monthly mean R_{GF} values in the dry and wet seasons were approximately 30%.

Table 3.5. R_{GF} estimated for different seasons in Phimai. The errors in the ratio is the 2σ standard error in the estimated seasonal mean R_{GF}

Season	R_{GF}
Over all period	0.032 ± 0.005
Dry season (Jan/2015-April/2015)	0.030 ± 0.001
Wet season (Jun/2015 & Aug/2015)	0.034 ± 0.002
Dry season (Jan/2016-April/2015)	0.028 ± 0.002
Wet season (Jun/2016-Sep/2016)	0.040 ± 0.003

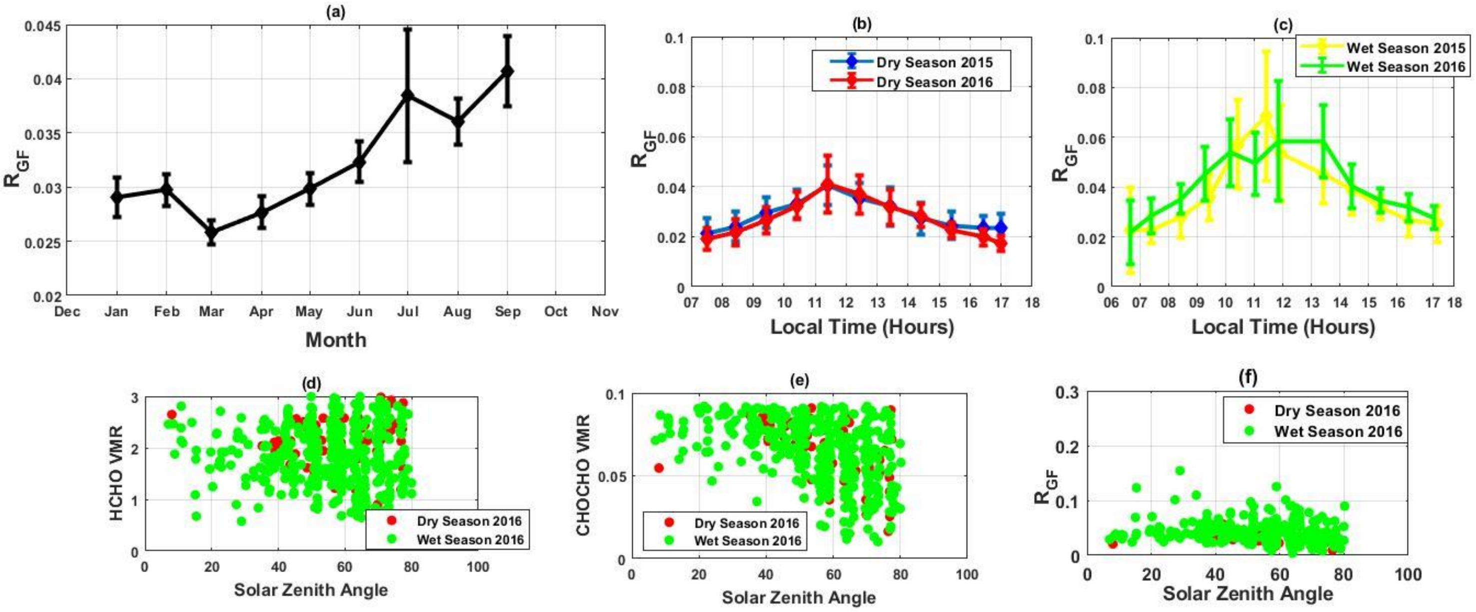


Figure 3.32. The monthly and diurnal variation of R_{GF} values in Phimai (a) The monthly mean R_{GF} during the dry and wet seasons. These monthly mean values include the months of the dry and seasons in both 2015 and 2016. (b) The dirunal variation of R_{GF} during the dry seasons in 2015 and 2016 and (c) The dirunal variation of R_{GF} during the wet seasons in 2015 and 2016.The error bars in (a-c) show the 2σ standard error of the mean values. (d-f) The HCHO and CHOCHO concentrations and the estimated R_{GF} values from the respective HCHO and CHOCHO concentrations are plotted against SZA in during the dry and wet season in 2016.

3.5.8 Diurnal variation in R_{GF}

Figure 3.32 shows the diurnal variations in R_{GF} during the dry and wet seasons in Phimai. The diurnal variations in R_{GF} shows lower values during early morning hours and higher during the afternoon hours, which resembles the typical diurnal variation of CHOCHO concentrations. As seen in Fig. 3.32 (b and c), the hourly mean R_{GF} values ranged between approximately 0.022 and ~0.068 during the wet season, whereas the hourly mean R_{GF} values during the dry season were lower (mostly between 0.017 and 0.040) compared to the wet season values. Thus it is evident that, despite the variation of R_{GF} values throughout the day in Phimai, the values in the wet season were higher than the dry season. Moreover, the diurnal mean R_{GF} values observed during the dry and wet season in Phimai were found to be consistent with the values reported by *Vrekoussis et al.* [2010]. As the R_{GF} values can vary throughout the day, such diurnal variation can impact the estimated seasonal and monthly mean R_{GF} values. But the seasonal and monthly mean R_{GF} values showed significant differences during the dry and wet season, indicating that the diurnal variability of R_{GF} had minimal impact on the estimated seasonal and monthly mean R_{GF} values.

3.5.9 Impact of SZA on R_{GF}

The sun position (i.e. SZA) can change between the dry and wet seasons in Phimai, which can impact the retrieval of HCHO and CHOCHO concentrations and hence the R_{GF} value estimated from the trace gas concentrations. Although, the impact of SZA has been taken into account in the radiative transfer calculation, here such impact is assessed on the final products (i.e., HCHO and CHOCHO concentrations and R_{GF}). The temperature in Phimai does not change significantly between the dry and wet seasons. Thus it is expected that the solar position in Phimai will not change significantly. The difference in the SZAs between the dry and wet seasons was approximately 10°. The HCHO and CHOCHO concentrations and the estimated R_{GF} values for the dry and wet seasons are plotted against the SZA variation in Phimai in Fig 3.32 (d-f). The dry season in Phimai is influenced by biomass burning, resulting in enhanced concentrations in the trace gas. To minimize the enhancement caused by pyrogenic emissions during the dry season, the 40th percentiles of HCHO and CHOCHO concentrations were defined as the threshold value. During both seasons, no strong dependence of the trace gases or R_{GF} on SZA was observed. Thus, the impact of difference in the SZA on the estimated R_{GF} is expected to be minimal.

3.5.10 Impact of wet removal on the estimated R_{GF}

The wet season in Phimai is characterized by rainfall. As wet deposition can be a probable sink of both HCHO and CHOCHO, it can also impact the R_{GF} values estimated during the wet season. To investigate the such impact of rainfall on R_{GF} a technique called superposed epoch analysis was used. For such analysis, the data during the wet season in 2016 was used. Because the 2016 wet season data were available for the all wet season months. The basic theme behind such analysis is to compare the R_{GF} values before and after rainy days. To judge a rainy day, the NCEP reanalysis relative humidity (RH) [*Kalnay et al.,* 1996] data on a temporal scale of six hours were used. A day was judged a rainy day if the hourly maximum was RH greater than 98%. The result of the superposed epoch analysis is shown in Fig. 3.33 (a). The zero, positive, and negative numbers in the x axis of Fig 3.33(a) indicate the rainy day, number of days after rain, and number of days before the rain. The numbers in red indicate the number of days used to calculate the R_{GF}. For example, the R_{GF} on the rainy day represents the mean of 47 rainy days. As seen from the figure the mean R_{GF} values showed no statistically significant differences before and after the rainy days judged based on the reanalysis data. In both cases, i.e., before and after rainy days, R_{GF} values > 0.04 were observed, which match with the results of *Vrekoussis et al* [2010] and *Kaiser et al* [2015].

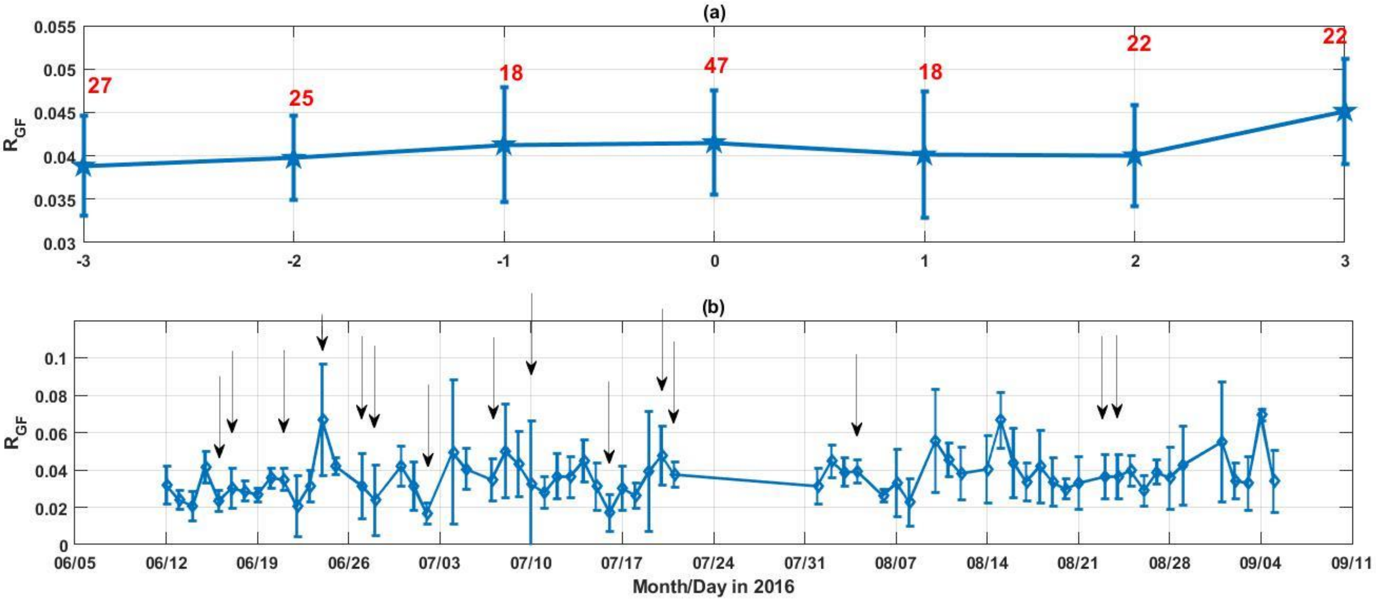


Figure 3.33. (a) The estimated mean R_{GF} values on rainy day, days after rain, and days before rain during the wet season in 2016. The NCEP reanalysis relative humidity (RH) data were used to judge rainy and non-rainy days. If the dirunal mean RH is greater than 98%, then it was judged as a rainy day. The zero, postive, and negative numbers indicate the rainy days, three days after rain, and three days before the rain, respectively. The numbers in red in the plot are the number of days used to calculate the respective mean R_{GF} values. (b) The hourly mean R_{GF} values during the weeks in the wet season of 2016. The black arrows in the plot are indicative of the rainy days judged based on the sky-view camera images in Phimai. The error bars in both (a) and (b) show the 2σ standard error of the mean values.

The Phimai site is also equipped with a sky-view camera. With such provision, individual rainy days can be identified and such individual events were examined in addition to the reanalysis data. The rainy and non-rainy days judged based on the skyview camera images and the respective daily mean R_{GF} values during the wet season in 2016 are shown in Figure 3.33(b). The black arrows in Fig. 3.33(b) indicate the rainy days. The R_{GF} values immediately after the rainy day and few days after rain were assessed. On days immediately after rain, both high (~0.04 or above) and low (< 0.04) R_{GF} values were observed. For example, on June 24 and July 7 which were immediate days after rain, R_{GF} values were ~0.04. Similarly, the R_{GF} values were found to high after few days of rain. From Fig 3.33(b), after rain on August 4, the R_{GF} values ~0.04 or higher were observed on August 14 and 15. Similarly, few days after rain on August 23 and 24, R_{GF} values greater than 0.04 were observed. Thus, comparing the two cases (i.e., immediately after rain and few days after the rain), no significant differences in the R_{GF} values were observed, which indicates minimum impact of wet deposition on the high R_{GF} values during the wet season. As seen from Fig 3.32, the R_{GF} showed strong diurnal variation and such comparison can also be affected by the diurnal variation in R_{GF}. Thus, the impact of wet removal on R_{GF} cannot be exempted completely and more analysis is required to have a qualitative idea about the wet deposition on the R_{GF} values during the wet season in Phimai.

3.5.11 Assessment of anthropogenic influence on the R_{GF}

The impact of pyrogenic activities on the R_{GF} values was assessed using the NO_2 concentration as a tracer of biomass burning. The correlation among NO_2, HCHO, and CHOCHO are shown in Fig. 3.34. Moderate correlation among the NO_2, HCHO, and CHOCHO daily mean concentration was observed. According to *Vrekoussis et al* [2010], when the NO_2 concentrations change, such change can have a systematic influence on the VOC concentrations, which can be expressed as a function of the pollution level in the region of interest. This can be seen, considering extreme cases using the regression lines in Fig. 3.34. When the NO_2 concentration are very low (i.e. close to 0 ppbv), the CHOCHO and HCHO concentrations are ~0.05 and ~1 ppbv, respectively. Then the R_{GF} value estimated from the HCHO and CHOCHO concentrations is 0.05. On the other hand, when the NO_2 concentration reaches the highest level, i.e. approximately 3 ppbv, the respective CHOCHO, HCHO, and R_{GF} values are ~0.28 and ~11 ppbv, and 0.025. This indicates that the NO_2 concentration level in Phimai is likely to have an influence on the VOC concentrations and hence the R_{GF} values. For the assessment of the NO_2 concentration on the estimated R_{GF} values, the NO_2 concentration of 0.8 ppbv for the 0-1 km layer was set as the threshold value. NO_2 concentrations above and below the defined threshold value were regarded as high and NO_2 periods, respectively.

The R_{GF} values estimated for the defined high and low NO_2 periods were denoted as R_{High} and R_{Low}, respectively. The NO_2 concentration of 0.8 ppbv was selected as a threshold because the value was close to the mean NO_2 concentration for the whole observation period in Phimai. Next, to take into account the change of SZA between the dry and wet seasons, small series of days within the same month of the same dry and wet seasons were selected. For example, during the dry seasons in 2015 and 2016, a small series of days (around 3 or 4 days) in March was selected when the respective NO_2 concentration in short series of days are higher and lower than 0.8 ppbv (i.e., the threshold value). Similarly, during the wet season, similar small series of days were considered when the NO_2 concentration was higher and lower than the defined threshold value. Fig. 3.34(c) shows the R_{GF} estimated based on the defined NO_2 threshold value. During the dry season in 2016, the estimated R_{High} and R_{Low} were 0.023 and 0.040, respectively. This means that when the NO_2 concentration was high, the R_{GF} value was low and vice versa.

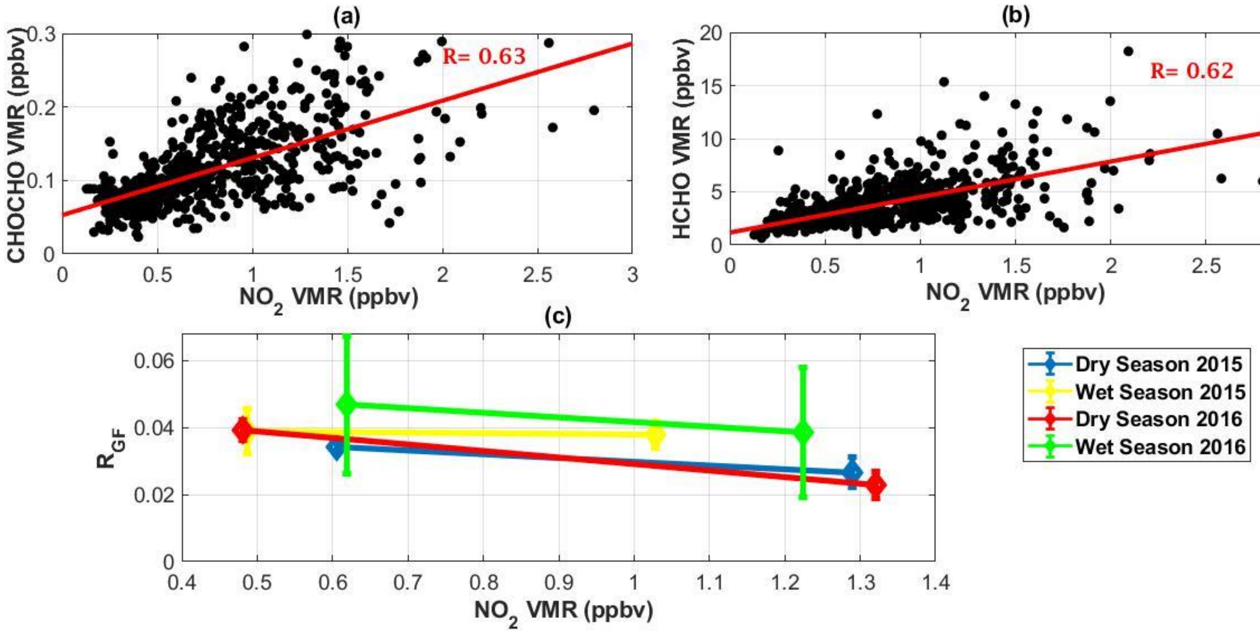


Figure 3.34. Correlation plots showing the correlations among NO_2, HCHO, and CHOCHO in Phimai. (a) Correlation plots showing the correlations between NO_2 and CHOCHO concetrations, (b) Correlation plot showing the correlation between NO_2 and HCHO concentrations. The data for the whole observation period were used to calculate the correlation coeeficeints. The daily mean concntrations of NO_2, HCHO, and CHOCHO were used for the plots. (c) The R_{GF} values estimated for short periods based on the NO_2 concentration of 0.8 ppbv set as threshold value. The error bars in the R_{GF} values display the 2σ standard error.

The estimated R_{High} and R_{Low} values for the dry season in 2015 were also low and high, respectively, similar to the features observed during the dry season in 2016. The conditions for estimating the R_{High} and R_{Low} were same. The R_{High} and R_{Low} were estimated under similar conditions such as the same month of a season, low impact of SZA, and same duration of days for high and low NO_2 cases. This yields that, the difference observed in the estimated R_{High} and R_{Low} values within the same dry season is related to the variation of the NO_2 concentration within the dry seasons.

Conversely, the R_{High} and R_{Low} values estimated for the wet season in 2015 and 2016 were almost similar, i.e., no significant differences were observed. Comparing the R_{High} values during the dry and wet season it has been observed that, the R_{High} values during the dry season was lower and higher during the wet season. This means that, the R_{GF} values estimated for the high NO_2 concentration cases during the dry seasons is likely to be related to the source of higher NO_2 concentration in the dry season. From the discussion in section 3.5.1 and 3.5.2, it was seen that the Phimai site was affected by biomass burning during the dry season and as a result higher NO_2, CHOCHO and HCHO concentrations were observed. Thus, the low R_{High} (lower R_{GF} value) during the dry season can be attributed to the pyrogenic activities during this period. In contrast, biogenic emission sources were the dominant VOC sources in Phimai during the wet season and thus, the R_{High} and R_{Low} values were high (high R_{GF} values).

The choice of the NO_2 concentration selected as the threshold value (i.e., 0.8 ppbv) can also affect the estimated R_{High} and R_{Low} values. Sensitivity tests were performed by estimating R_{High} and R_{Low} values using the NO_2 concentration threshold value of 1.0, 0.9, 0.7, and 0.6 ppbv for the 0-1 km layer. The estimated R_{High} and R_{Low} values changed approximately 20% when the threshold NO_2 concentration value was changed. Such change in the R_{High} and R_{Low} values is likely due to the difference in the number of the days in the series considered for estimating the ratio for different cases during the dry and wet seasons. However, although the absolute values of R_{High} and R_{Low} showed some sensitivity to the NO_2 threshold value, the response of R_{High} and R_{Low} values to different NO_2 concentration cases remained unchanged during the dry and wet seasons.

Our analysis showed that the R_{GF} values were lower during the strong pyrogenic activities in the dry season whereas R_{GF} values were higher when the dominant VOC sources were biogenic during the wet season. Such behavior of the R_{GF} values under different VOC scenarios were consistent with the satellite data based study by *Vrekoussis et al.* [2010]. R_{GF} values in the range of 0.04 – 0.07 under strong biogenic emission scenarios was also observed by *Kaiser et al.* [2015]. However, the R_{GF} response under biomass burning observed in Phimai was contrast to that reported by *DiGangi et al.* [2012]. In Phimai, the R_{GF} values were lower in the dry season due to strong fire activities, whereas *DiGangi et al.* [2012] and *Zarzana et al.* [2017] stated that the R_{GF} values tends to be higher during fire events. The discrepancy observed in the R_{GF} response in Phimai during the dry season and the findings of *DiGangi et al.* [2012] and *Zarzana et al.* [2017] can be potentially due to the type of biomass fuel burnt and also the age of the plumes reaching the observation sites or the measuring instruments. *Zarzana et al.* [2017] found that the R_{GF} values increased when their airborne sampler moved away from the fire source towards the downwind direction. Similarly, *DiGangi et al.* [2012] found that the age of the plume can impact the R_{GF} values under fire conditions. Additionally, different yields of CHOCHO and HCHO from different biomass fuel [*Akagi et al.*, 2011; *Stockwell et al.*, 2015] can also impact the R_{GF} values under fire conditions.

In the case of dominant biogenic VOC emission sources, the R_{GF} value will depend on the different yields of HCHO and CHOCHO from the VOC sources. For example, modelling studies by *Kaiser et al* [2015] showed that the yield of CHOCHO per HCHO is higher from monoterpenes compared to isoprenes, which led to higher R_{GF} values under dominant biogenic emission sources. In contrast, *DiGangi et al.* [2012] found that R_{GF} values were lower when the dominant VOC emission sources were monoterpenes and 2-methyl-3-buten-2-ol. The discrepancy observed between *Kaiser et al* [2015] and *DiGangi et al.* [2012] was attributed to the inadequate knowledge on the mechanism of CHOCHO and HCHO production from isoprene when the NO_x concentrations are lower. The dominant biogenic VOC source at the Phimai site is isoprene and during the wet season when the NO_2 concentrations are lower (~ 0.7 ppbv) compared to the dry seasons, higher R_{GF} values (~ 0.04) were observed. This contradicts with the results of Kaiser *et al.* [2015] which found lower R_{GF} values under dominant isoprene emissions and low NO_x conditions. However, Kaiser *et al.* [2015] did not completely eliminate the impact of diurnal variation of R_{GF} on the reported results. As seen in our analysis, during the wet season the R_{GF} showed strong diurnal variation. Despite such diurnal variations and under low NO_2 concentration, the R_{GF} values in Phimai were higher.

This indicates that under the similar VOC emission scenario, the R_{GF} response can be different due to the different climate conditions and VOC distributions. Moreover, the study by *Baidar et al*. [2012] using airborne MAX-DOAS systems found that the R_{GF} values at altitude higher than the boundary layer (~ 2.5 km) is higher compared to that of within the boundary layer. Therefore, the altitude dependence of R_{GF} can also be attributed as a potential reason of the discrepancies observed among different studies. More measurements and studies on R_{GF} are required focusing on different region and climatic conditions to address the overall discrepancy of the R_{GF} response among different studies.

References

Akagi, S. K., R. J. Yokelson, C. Wiedinmyer, M. J. Alvarado, J. S. Reid, T. Karl, J. D. Crounse, and P. O. Wennberg (2011), Emission factors for open and domestic biomass burning for use in atmospheric models, *Atmos. Chem. and Phys*, *11*(9), 4039-4072, doi:10.5194 /acp-11-4039-2011.

Akagi, S. K., J. S. Craven, J. W. Taylor, G. R. McMeeking, R. J. Yokelson, I. R. Burling, S. P. Urbanski, C. E. Wold, J. H. Seinfeld, and H. Coe (2012), Evolution of trace gases and particles emitted by a chaparral fire in California, *Atmos. Chem. and Phys.*, *12*(3), 1397-1421, doi:10.5194 /acp-12-1397-2012.

Alvarado, L. M. A., A. Richter, M. Vrekoussis, F. Wittrock, A. Hilboll, S. F. Schreier, and J. P. Burrows (2014), An improved glyoxal retrieval from OMI measurements, *Atmos. Meas. Tech*, *7*(12), 4133, doi:10.5194 /amt - 7 - 4 133 - 2014.

Andreae, M. O., and P. Merlet (2001), Emission of trace gases and aerosols from biomass burning, *Global Biogeochem. Cy.*, *15*(4), 955-966, doi:10.1029/2000GB001382.

Arlander, D., D. Brüning, U. Schmidt, and D. Ehhalt (1995), The tropospheric distribution of formaldehyde during TROPOZ II, *J. Atmos. Chem.*, *22*(3), 251-269, doi:10.1007/BF00696637.

Baidar, S., H. Oetjen, S. Coburn, B. Dix, I. Ortega, R. Sinreich, and R. Volkamer (2013), The CU Airborne MAX-DOAS instrument: Vertical profiling of aerosol extinction and trace gases, *Atmos. Meas. Tech.*, *6*(3), 719-739, doi:10.5194/amt-6-719-2013.

Bogumil, K., J. Orphal, T. Homann, S. Voigt, P. Spietz, O. Fleischmann, A. Vogel, M. Hartmann, H. Kromminga, and H. Bovensmann (2003), Measurements of molecular absorption spectra with the SCIAMACHY pre-flight model: instrument characterization and reference data for atmospheric remote-sensing in the 230–2380 nm region, *J. Photocho. Photobio. A*, *157*(2), 167-184, doi:10.1016/S1010-6030(03)00062-5.

Bosilovich, M. G., S. Akella, L. Coy, R. Cullather, C. Draper, R. Gelaro, R. Kovach, Q. Liu, A. Molod, P. Norris, K. Wargan, W. Chao, R. Reiche, L. Takacs, Y. Vikhliaev, S. Bloom, A. Collow, S. Firth, G.Labow, G. Partyka, S. Pawson, O. Reale, S.D. Schubert, and M. Suarez (2015), MERRA-2: Initial evaluation of the climate, *NASA Tech. Rep. Series on Global Modelling an.d Data Assimilation NASA/TM-2015-104606.*

Burkert, J., M. D. Andrés-Hernández, D. Stöbener, J. P. Burrows, M. Weissenmayer, and A. Kraus (2001), Peroxy radical and related trace gas measurements in the boundary layer above the Atlantic Ocean, *J. Geophys. Res.*, *106*(D6), 5457-5477, doi:10.1029/2000JD900613.

Cadeddu, M. P., J. C. Liljegren, and D. D. Turner (2013), The Atmospheric radiation measurement (ARM) program network of microwave radiometers: instrumentation, data, and retrievals, in *Atmos. Meas. Tech.*, edited, pp. 2359-2372, Copernicus Publications, doi:10.5194/amt-6-2359-2013.

Campbell, J. R., J. S. Reid, D. L. Westphal, J. Zhang, J. L. Tackett, B. N. Chew, E. J. Welton, A. Shimizu, N. Sugimoto, and K. Aoki (2013), Characterizing the vertical profile of aerosol particle extinction and linear depolarization over Southeast Asia and the Maritime Continent: The 2007–2009 view from CALIOP, *Atmos. Res.*, *122,* 520-543, doi:10.1016/j.atmosres.2012.05.007.

Cerqueira, M., C. Pio, P. Gomes, J. Matos, and T. Nunes (2003), Volatile organic compounds in rural atmospheres of central Portugal, *Sci. Total Environ.,313*(1), 49-60, doi:10.1016/S0048-9697(03)00250-X.

Chance, K. V., and R. J. D. Spurr (1997), Ring effect studies: Rayleigh scattering, including molecular parameters for rotational Raman scattering, and the Fraunhofer spectrum, *Appl. Opt.* , *36*(21), 5224-5230, doi:10.1364/AO.36.005224.

Choi, W., I. Faloona, N. Bouvier-Brown, M. McKay, A. Goldstein, J. Mao, W. Brune, B. LaFranchi, R. Cohen, and G. Wolfe (2010), Observations of elevated formaldehyde over a forest canopy suggest missing sources from rapid oxidation of arboreal hydrocarbons, *Atmos. Chem. Phys.*, *10*(18), 8761-8781, doi:10.5194/acp-10-8761-2010.

Clémer, K., M. Van Roozendael, C. Fayt, F. Hendrick, C. Hermans, G. Pinardi, R. Spurr, P. Wang, and M. De Mazière (2010), Multiple wavelength retrieval of tropospheric aerosol optical properties from MAXDOAS measurements in Beijing, *Atmos. Meas. Tech.*, *3*, 863-878, doi:10.5194/amt-3-863-2010.

Cohen, W. B., T. K. Maiersperger, S. T. Gower, and D. P. Turner (2003), An improved strategy for regression of biophysical variables and Landsat ETM+ data, *Remote Sens. Environ.*, *84*(4), 561-571, doi:10.1016/S0034 - 4257 (02) 00173 - 6.

Crutzen, P. J., L. E. Heidt, J. P. Krasnec, W. H. Pollock, and W. Seiler (1979), Biomass burning as a source of atmospheric gases CO, H_2, N_2O, NO, CH_3Cl and COS, *Nature*, *282*(5736), 253-256, doi: 10.1038/282253a0.

Davi, H., K. Soudani, T. Deckx, E. Dufrene, V. Le Dantec, and C. Francois (2006), Estimation of forest leaf area index from SPOT imagery using NDVI distribution over forest stands, *Int. J. Remote Sens.*, *27*(05), 885-902, doi:10.1080/01431160500227896.

DiGangi, J., S. Henry, A. Kammrath, E. Boyle, L. Kaser, R. Schnitzhofer, M. Graus, A. Turnipseed, J. Park, R. Weber, R. S. Hornbrook, C. A. Cantrell, R. L. Maudlin III, S. Kim, Y. Nakashima, G. M. Wolfe, Y. Kajii, E. Apel, A. H. Goldstein, A. Guenther, T. Karl, A. Hansel, and F. N. Keutsch (2012), Observations of glyoxal and formaldehyde as metrics for the anthropogenic impact on rural photochemistry, *Atmos. Chem. Phys.*, *12*(20), 9529-9543, doi:10.5194/acp-12-9529-2012.

Fleischmann, O. C., M. Hartmann, J. P. Burrows, and J. Orphal (2004), New ultraviolet absorption cross-sections of BrO at atmospheric temperatures measured by time-windowing Fourier transform spectroscopy, *J. Photocho. Photobio. A*, *168*(1), 117-132, doi:10.1016/j.jphotochem.2004.03.026.

Franco, B., F. Hendrick, M. Van Roozendael, J. -F. Müller, T. Stavrakou, E. Marais, B. Bovy, W. Bader, C. Fayt, and C. Hermans (2015), Retrievals of formaldehyde from ground-based FTIR and MAX-DOAS observations at the Jungfraujoch station and comparisons with GEOS-Chem and IMAGES model simulations, *Atmos. Meas. Tech*, *8*, 1733-1756, doi:10.5194/amt-8-1733-2015.

Frieß, U., P. Monks, J. Remedios, A. Rozanov, R. Sinreich, T. Wagner, and U. Platt (2006), MAX-DOAS O_4 measurements: A new technique to derive information on atmospheric aerosols: 2. Modeling studies, *J. Geophys. Res.*, *111*, D14203, doi:10.1029/2005JD006618.

Frieß, U., H. K. Baltink, S. Beirle, K. Clémer, F. Hendrick, B. Henzing, H. Irie, G. de Leeuw, A. Li, and M. Moerman (2016), Intercomparison of aerosol extinction profiles retrieved from MAX-DOAS measurements, *Atmos. Meas. Tech.*, *9*(7), 3205-3222, doi:10.5194/amt-9-3205-2016.

Fu, T., D. Jacob, F. Wittrock, J. P. Burrows, M. Vrekoussis, and D. Henze (2008), Global budgets of atmospheric glyoxal and methylglyoxal, and implications for formation of secondary organic aerosols, *J. Geophys. Res.*, *113*, D15303, doi:10.1029/2007JD009505.

Galanter, M., H. Levy, and G. R. Carmichael (2000), Impacts of biomass burning on tropospheric CO, NO_x, and O_3, *J. Geophys. Res.*, *105*(D5), 6633-6653, doi:10.1029/1999JD901113.

Gao, X., A. R. Huete, W. Ni, and T. Miura (2000), Optical–biophysical relationships of vegetation spectra without background contamination, *Remote Sens. Environ.*, *74*(3), 609-620, doi:10.1016/S 003-4-4257 (00) 00150-4.

Geiger, H., J. Kleffmann, and P. Wiesen (2002), Smog chamber studies on the influence of diesel exhaust on photosmog formation, *Atmos. Environ.*, *36*(11), 1737-1747, doi:10.1016/S1352-2310(02)00175-9.

Grosjean, E., D. Grosjean, M. P. Fraser, and G. R. Cass (1996), Air quality model evaluation data for organics. 2. C1− C14 carbonyls in Los Angeles air, *Environ. Sci. Technol.*, *30*(9), 2687-2703, doi:10.1021/es950758w.

Hendrick, F., J. -F. Müller, K. Clémer, P. Wang, M. D. Mazière, C. Fayt, C. Gielen, C. Hermans, J. Ma, and G. Pinardi (2014), Four years of ground-based MAX-DOAS observations of HONO and NO_2 in the Beijing area, *Atmos. Chem. Phys.*, *14*(2), 765-781, doi:10.5194/acp-14-765-2014.

Hermans, C., A. Vandaele, S. Fally, M. Carleer, R. Colin, B. Coquart, A. Jenouvrier, and M. -F. Merienne (2003), Absorption cross-section of the collision-induced bands of oxygen from the UV to the NIR, in *Weakly interacting molecular pairs: unconventional absorbers of radiation in the atmosphere*, edited, pp. 193-202, Springer.

Holzinger, R., C. Warneke, A. Hansel, A. Jordan, W. Lindinger, D. H. Scharffe, G. Schade, and P. J. Crutzen (1999), Biomass burning as a source of formaldehyde, acetaldehyde, methanol, acetone, acetonitrile, and hydrogen cyanide, *Geophys. Res. Lett.*, *26*(8), 1161-1164, doi:10.1029/1999GL900156.

Hönninger, G., C. v. Friedeburg, and U. Platt (2004), Multi axis differential optical absorption spectroscopy (MAX-DOAS), *Atmos. Chem. Phys.*, *4*(1), 231-254, doi:10.5194/acp-4-231-2004.

Houweling, S., F. Dentener, and J. Lelieveld (1998), The impact of nonmethane hydrocarbon compounds on tropospheric photochemistry, *J. Geophys. Res.*, *103*(D9), 10673-10696, doi:10.1029/97JD03582.

Irie, H., Y. Kanaya, H. Akimoto, H. Iwabuchi, A. Shimizu, and K. Aoki (2008a), First retrieval of tropospheric aerosol profiles using MAX-DOAS and comparison with lidar and sky radiometer measurements, *Atmos. Chem. Phys.*, *8*(2), 341-350, doi:10.5194/acp-8-341-2008.

Irie, H., Y. Kanaya, H. Akimoto, H. Tanimoto, Z. Wang, J. Gleason, and E. Bucsela (2008b), Validation of OMI tropospheric NO_2 column data using MAX-DOAS measurements deep inside the North China Plain in June 2006: Mount Tai Experiment 2006, *Atmos. Chem. Phys.*, *8*(22), 6577-6586, doi:10.5194/acp-8-6577-2008.

Irie, H., Y. Kanaya, H. Akimoto, H. Iwabuchi, A. Shimizu, and K. Aoki (2009), Dual-wavelength aerosol vertical profile measurements by MAX-DOAS at Tsukuba, Japan, *Atmos. Chem. Phys.*, *9*(8), 2741-2749, doi:10.5194/acp-9-2741-2009.

Irie, H., H. Takashima, Y. Kanaya, K. Boersma, L. Gast, F. Wittrock, D. Brunner, Y. Zhou, and M. V. Roozendael (2011), Eight-component retrievals from ground-based MAX-DOAS observations, *Atmos. Meas. Tech.*, *4*(6), 1027-1044, doi:10.5194/amt-4-1027-2011.

Irie, H., T. Nakayama, A. Shimizu, A. Yamazaki, T. Nagai, A. Uchiyama, Y. Zaizen, S. Kagamitani, and Y. Matsumi (2015), Evaluation of MAX-DOAS aerosol retrievals by coincident observations using CRDS, lidar, and sky radiometer in Tsukuba, Japan, *Atmos. Meas. Tech.*, *8*(7), 2775-2788, doi:10.5194/amt-8-2775-2015.

Iwabuchi, H. (2006), Efficient Monte Carlo methods for radiative transfer modeling, *J. Atmos. Sci.*, *63*(9), 2324-2339, doi:10.1175/JAS3755.1.

Justice, C. O., E. Vermote, J. R. G. Townshend, R. Defries, D. P. Roy, D. K. Hall, V. V. Salomonson, J. L. Privette, G. Riggs, A. Strahler, W. Lucht, R. B. Myneni, Y. Knyazikhin, S. W. Running, R. R. Nemani, Z. Wan, A. R. Huete, W. van Leeuwen, R. E. Wolfe, L. Gigilo, J. P. Muller, P. Lewis, and M. J. Barnsley (1998), The Moderate Resolution Imaging Spectroradiometer (MODIS): Land remote sensing for global change research, *IEEE. T. Geosci. Remote.*, *36*(4), 1228-1249, doi:10.1109/36.701075.

Justice, C. O., L. Giglio, S. Korontzi, J. Owens, J. T. Morisette, D. P. Roy, J. Descloitres, S. Alleaume, F. Petitcolin, and Y. Kaufman (2002), The MODIS fire products, *Remote Sens. Environ.*, *83*(1), 244-262, doi:10.1016/S0034 - 4257 (02) 00076 - 7.

Kaiser, J., G. M. Wolfe, K. E. Min, S. S.Brown , C. C. Miller, D. J. Jacob, J. A. deGouw, M. Graus, T. F. Hanisco, J. Holloway, J. Peischl, I. B. Pollack, T. B. Ryerson, C. Warkene, R. A. Washenfelder, and F. N. Keutsch (2015), Reassessing the ratio of glyoxal to formaldehyde as an indicator of hydrocarbon precursor speciation, *Atmos. Chem. and Phys.*, *15*(13), 7571-7583, doi:10.5194/acp-15-7571-2015.

Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K. C. Mo, C. Ropelewski, J. Wang, A. Leetma, R. Reynolds, R. Jenne, and D. Joseph (1996), The NCEP/NCAR reanalysis 40-year project, *Bull. Am. Meteorol. Soc*, *77*(3), 437-471.

Kaufman, Y. J., C. O. Justice, L. P. Flynn, J. D. Kendall, E. M. Prins, L. Giglio, D. E. Ward, W. P. Menzel, and A. W. Setzer (1998), Potential global fire monitoring from EOS-MODIS, *J. Geophys. Res.*, *103*(D24), 32215-32238, doi:10.1029/98JD01644.

Kurucz, R., I. Furenlid, J. Brault, and L. Testerman (1984), Solar Flux Atlas from 296 to 1300 nm, *Natl. Sol. Obs., Sunspot, New Mexico*, 240.

Lee, Y. N., X. Zhou, and K. Hallock (1995), Atmospheric carbonyl compounds at a rural southeastern United States site, *J Geophys. Res.*, *100*(D12), 25933-25944, doi:10.1029/95jd02605.

Li, C., S. -C. Tsay, N. C. Hsu, J. Y. Kim, S. G. Howell, B. J. Huebert, Q. Ji, M. -J. Jeong, S. -H. Wang, and R. A. Hansell (2013a), Characteristics and composition of atmospheric aerosols in Phimai, central Thailand during BASE-ASIA, *Atmos. Environ.*, *78*, 60-71, doi:10.1016/j.atmosenv.2012.04.003.

Li, X., T. Brauers, A. Hofzumahaus, K. Lu, Y. Li, M. Shao, T. Wagner, and A. Wahner (2013b), MAX-DOAS measurements of NO_2, HCHO and CHOCHO at a rural site in Southern China, *Atmos. Chem. Phys.*, *13*(4), 2133-2151, doi:10.519/acp-13-2133-2013.

Liakakou, E., M. Vrekoussis, B. Bonsang, C. Donousis, M. Kanakidou, and N. Mihalopoulos (2007), Isoprene above the Eastern Mediterranean: Seasonal variation and contribution to the oxidation capacity of the atmosphere, *Atmos. Environ.*, *41*(5), 1002-1010, doi:10.1016/j.atmosenv.2006.09.034.

Logan, J. A., M. J. Prather, S. C. Wofsy, and M. B. McElroy (1981), Tropospheric chemistry: A global perspective, *J. Geophys. Res.*, *86*(C8), 7210-7254, doi:10.1029/JC086iC08p07210.

McMillan, W. W., C. Barnet, L. Strow, M. T. Chahine, M. L. McCourt, J. X. Warner, P. C. Novelli, S. Korontzi, E. S. Maddy, and S. Datta (2005), Daily global maps of carbon monoxide from NASA's Atmospheric Infrared Sounder, *Geophys. Res. Lett.*, *32*(11), doi:10.1029/2004GL021821.

Meller, R., and G. K. Moortgat (2000), Temperature dependence of the absorption cross sections of formaldehyde between 223 and 323 K in the wavelength range 225–375 nm, *J. Geophys. Res.*, *105*(D6), 7089-7101, doi:10.1029/1999JD901074.

Miller, C., G. Gonzalez Abad, H. Wang, X. Liu, T. Kurosu, D. J. Jacob, and K. V. Chance (2014), Glyoxal retrieval from the ozone monitoring instrument, *Atmos. Meas. Tech.* , doi:10.5194/amt-7-3891-2014.

Monks, P. S. (2005), Gas-phase radical chemistry in the troposphere, in *Chem. Soc. Rev.*, 376-395, doi:10.1039/B307982C.

Moortgat, G., D. Grossmann, A. Boddenberg, G. Dallmann, A. Ligon, W. Turner, S. Gäb, F. Slemr, W. Wieprecht, and K. Acker (2002), Hydrogen peroxide, organic peroxides and higher carbonyl compounds determined during the BERLIOZ campaign, *J. Atmos. Chem.*, *42*, 443-463, doi:10.1023/A:1015743013107.

Munger, J. W., D. J. Jacob, B. C. Daube, L. Horowitz, W. Keene, and B. Heikes (1995), Formaldehyde, glyoxal, and methylglyoxal in air and cloudwater at a rural mountain site in central Virginia, *J. Geophys. Res.*, *100*(D5), 9325-9333, doi:10.1029/95JD00508.

Myriokefalitakis, S., M. Vrekoussis, K. Tsigaridis, F. Wittrock, A. Richter, C. Brühl, R. Volkamer, J. P. Burrows, and M. Kanakidou (2008), The influence of natural and anthropogenic secondary sources on the glyoxal global distribution, *Atmos. Chem. Phys.*, *8*(16), 4965-4981, doi:10.5194/acp-8-4965-2008.

Ortega, I., T. Koenig, R. Sinreich, D. Thomson, and R. Volkamer (2015), The CU 2-D-MAX-DOAS instrument–Part 1: Retrieval of 3-D distributions of NO 2 and azimuth-dependent OVOC ratios, *Atmos. Meas. Tech.*, *8*(6), 2371-2395, doi:10.5194/amt - 8 - 2371 - 2015.

Pérez-Ramírez, D., D. N. Whiteman, A. Smirnov, H. Lyamani, B. N. Holben, R. Pinker, M. Andrade, and L. Alados-Arboledas (2014), Evaluation of AERONET precipitable water vapor versus microwave radiometry, GPS, and radiosondes at ARM sites, *J. Geophys. Res.*, *119*(15), 9596-9613, doi:10.1002/2014JD021730.

Platt, U. (1994), Differential optical absorption spectroscopy (DOAS), in *Chemical Analysis Series*, edited, pp. 27-84, Wiley & Sons. Inc.

Platt, U., and J. Stutz (2008), *Differential Optical Absorption Spectroscopy*, Springer.

Poisson, N., M. Kanakidou, and P. J. Crutzen (2000), Impact of non-methane hydrocarbons on tropospheric chemistry and the oxidizing power of the global troposphere: 3-dimensional modelling results, *J. Atmos. Chem.*, *36*(2), 157-230, doi:10.1023/A:1006300616544.

Possanzini, M., G. Tagliacozzo, and A. Cecinato (2007), Ambient levels and sources of lower carbonyls at Montelibretti, Rome (Italy), *Water Air Soil Poll.* , *183*(1-4), 447-454, doi:10.1007/s11270-007-9393-1.

Rodgers, C. D. (2000), *Inverse methods for atmospheric sounding: theory and practice*, World scientific Singapore.

Rothman, L. S., I. E. Gordon, A. Barbe, D. C. Benner, P. F. Bernath, M. Birk, V. Boudon, L. R. Brown, A. Campargue, J. -P. Champion, K. Chance, L. H. Coudert, V. Dana, V. M. Devi, S. Faly, J. -M. Flaud, R. R. Gamache, A. Goldman, D. Jacquemart, I. Kleiner, N. Lacome, W. J. Lafferty, J. -Y. Mandin, S. T. Massaie, S. N. Mikhailenko, C. E. Miller, N. Moazzen-Ahmadi, O. V. Naumenko, A. V. Nikitin, J.

Orphal, V. I. Perevalov, A. Perrin, A. Predoi-Cross, C. P. Rinsland, M. Rogter, M. Simeckova, M. A. H. Smith, K. Sung, S. A. Tashkun, J. Tennyson, R. A. Toth, A. C. Vandaele, and J. Vander Auwera (2009), The HITRAN 2008 molecular spectroscopic database, *J. Quant. Spectrosc. Ra.*, *110*(9-10), 533-572, doi:10.1016 / j.jsrt.2009.02.013.

Running, S. W., C. O. Justice, V. Salomonson, D. Hall, J. Barker, Y. J. Kaufmann, A. H. Strahler, A. R. Huete, J. -P. Muller, V. Vanderbilt, Z. M. Wan, P. Teillet, and D. Carneggie (1994), Terrestrial remote sensing science and algorithms planned for EOS/MODIS, *Int. J. Remote Sens.*, *15*(17), 3587-3620, doi:10.1080/01431169408954346.

Schreier, S. F., A. Richter, J. W. Kaiser, and J. P. Burrows (2014), The empirical relationship between satellite-derived tropospheric NO 2 and fire radiative power and possible implications for fire emission rates of NO x, *Atmos. Chem. and Phys.*, *14*(5), 2447-2466, doi:10.5194/acp-14-2447-2014.

Seco, R., J. Peñuelas, and L. Filella (2007), Short-chain oxygenated VOCs: Emission and uptake by plants and atmospheric sources, sinks, and concentrations, *Atmos. Environ.*, *41*(12), 2477-2499, doi:10.1016/j.atmosenv.2006.11.029.

Singh, H. B., Y. Chen, A. Staudt, D. Jacob, D. Blake, B. Heikes, and J. Snow (2001), Evidence from the Pacific troposphere for large global sources of oxygenated organic compounds, *Nature*, *410*(6832), 1078-1081, doi:10.1038/35074067.

Singh, H. B., L. J. Salas, R. B. Chatfield, E. Czech, A. Fried, J. Walega, M. J. Evans, B. D. Field, D. J. Jacob, and D. Blake (2004), Analysis of the atmospheric distribution, sources, and sinks of oxygenated volatile organic chemicals based on measurements over the Pacific during TRACE-P, *J. Geophys. Res.*, *109*, D15S07, doi:10.1029/2003JD003883.

Sinreich, R., U. Frieß, T. Wagner, and U. Platt (2005), Multi axis differential optical absorption spectroscopy (MAX-DOAS) of gas and aerosol distributions, *Faraday Discuss.*, *130*, 153-164, doi:10.1039/B419274P.

Son, N. T., C. F. Chen, C. R. Chen, L. Y. Chang, H. N. Duc, and L. D. Nguyen (2013), Prediction of rice crop yield using MODIS EVI− LAI data in the Mekong Delta, Vietnam, *Int. J. Remote Sens.*, *34*(20), 7275-7292, doi:10.1080/01431161.2013.818258.

Stavrakou, T., J. -F. Müller, M. Bauwens, I. De Smedt, M. Van Roozendael, A. Guenther, M. Wild, and X. Xia (2014), Isoprene emissions over Asia 1979–2012: impact of climate and land-use changes, *Atmos. Chem. and Phys.*, *14*(9), 4587-4605, doi:10.5194/acp-14-4587-2014.

Stein, A. F., R. R. Draxler, G. D. Rolph, B. J. B. Stunder, M. D. Cohen, and F. Ngan (2015), NOAA’s HYSPLIT atmospheric transport and dispersion modeling system, *Bull. Amer. Meteor. Soc.*, *96*(12), 2059-2077, doi:10.1175/BAMS-D-14-00110.1.

Stockwell, C. E., P. R. Veres, J. Williams, and R. J. Yokelson (2015), Characterization of biomass burning emissions from cooking fires, peat, crop residue, and other fuels with high-resolution proton-transfer-reaction time-of-flight mass spectrometry, *Atmos. Chem. and Phys.*, *15*(2), 845-865, doi:10.5194/acp-15-845-2015.

Sugimoto, N., A. Shimizu, T. Nishizawa, I. Matsui, Y. Jin, P. Khatri, H. Irie, T. Takamura, K. Aoki, and B. Thana (2015), Aerosol characteristics in Phimai, Thailand determined by continuous observation with a polarization sensitive Mie–Raman lidar and a sky radiometer, *Environ. Res. Lett.* , *10*(6), 065003, doi:10.1088/1748-9326/10/6/065003.

Takashima, H., H. Irie, Y. Kanaya, and F. Syamsudin (2012), NO2 observations over the western Pacific and Indian Ocean by MAX-DOAS on Kaiyo, a Japanese research vessel, *Atmos. Meas. Tech.* , *5*(10), 2351, doi:10.5194/amt-5-2351-2012.

Thompson, A. M., K. E. Pickering, R. R. Dickerson, W. G. Ellis, D. J. Jacob, J. R. Scala, W. Tao, D. P. McNamara, and J. Simpson (1994), Convective transport over the central United States and its role in regional CO and ozone budgets, *J. Geophys. Res.*, *99*(D9), 18703-18711, doi:10.1029/94JD01244.

Vandaele, A., C. Hermans, P. Simon, M. Van Roozendael, J. Guilmot, M. Carleer, and R. Colin (1996), Fourier transform measurement of NO_2 absorption cross-section in the visible range at room temperature, *J. Atmos. Chem.*, *25*(3), 289-305, doi:10.1007/BF00053797.

Volkamer, R., P. Spietz, J. P. Burrows, and U. Platt (2005), High-resolution absorption cross-section of glyoxal in the UV–vis and IR spectral ranges, *J. Photocho. Photobio. A*, *172*(1), 35-46, doi:10.1016/ j.jphotochem.2004.11.011.

Volkamer, R., J. L. Jimenez, F. San Martini, K. Dzepina, Q. Zhang, D. Salcedo, L. T. Molina, D. R. Worsnop, and M. J. Molina (2006), Secondary organic aerosol formation from anthropogenic air pollution: Rapid and higher than expected, *Geophys. Res. Lett.*, *33,* L17811, doi:10.1029/2006GL026899.

Volkamer, R., F. San Martini, L. T. Molina, D. Salcedo, J. L. Jimenez, and M. J. Molina (2007), A missing sink for gas-phase glyoxal in Mexico City: Formation of secondary organic aerosol, *Geophys. Res. Lett.*, *34*(19), doi:10.1029 / 2007GL 030752.

Vrekoussis, M., N. Mihalopoulos, E. Gerasopoulos, M. Kanakidou, P. Crutzen, and J. Lelieveld (2007), Two-years of NO_3 radical observations in the boundary layer over the Eastern Mediterranean, *Atmos. Chem. Phys.*, *7*(2), 315-327, doi:10.5194/acp-7-315-2007.

Vrekoussis, M., F. Wittrock, A. Richter, and J. P. Burrows (2009), Temporal and spatial variability of glyoxal as observed from space, *Atmos. Chem. Phys.*, *9*(13), 4485-4504, doi:10.5194/acp-9-4485-2009.

Vrekoussis, M., F. Wittrock, A. Richter, and J. P. Burrows (2010), GOME-2 observations of oxygenated VOCs: what can we learn from the ratio glyoxal to formaldehyde on a global scale?, *Atmos. Chem. Phys.* , *10*(21), 10145-10160, doi:10.5194/acp-10-10145-2010.

Wagner, T., B. v. Dix, C. v. Friedeburg, U. Frieß, S. Sanghavi, R. Sinreich, and U. Platt (2004), MAX-DOAS O_4 measurements: A new technique to derive information on atmospheric aerosols—Principles and information content, *J. Geophys. Res.*, *109*, D22205, doi:10.1029/2004JD004904.

Wagner, T., J. P. Burrows, T. Deutschmann, B. Dix, C. v. Friedeburg, U. Frieß, F. Hendrick, K. -P. Heue, H. Irie, and H. Iwabuchi (2007), Comparison of box-air-mass-factors and radiances for Multiple-Axis Differential Optical Absorption Spectroscopy (MAX-DOAS) geometries calculated from different UV/visible radiative transfer models, *Atmos. Chem. Phys.*, *7*(7), 1809-1833, doi:10.5194/acp-7-1809-2007.

Wang, Q., S. Adiku, J. Tenhunen, and A. Granier (2005), On the relationship of NDVI with leaf area index in a deciduous forest site, *Remote Sens. Environ.*, *94*(2), 244-255, doi:10.1016 / j.rse. 2004.10.006.

Wang, T., F. Hendrick, P. Wang, G. Tang, K. Clémer, H. Yu, C. Fayt, C. Hermans, C. Gielen, and J. F. Müller (2014), Evaluation of tropospheric SO_2 retrieved from MAX-DOAS measurements in Xianghe, China, *Atmos. Chem. Phys.*, *14*(20), 11149-11164, doi:10.5194/acp-14-11149-2014.

Warneke, C., J. A. de Gouw, A. Stohl, O. R. Cooper, P. D. Goldan, W. C. Kuster, J. S. Holloway, E. J. Williams, B. M. Lerner, S. A. McKeen, M. Trainer, F. C. Fehsenfeld, E. L. Atlas, S. G. Donnelly, V. Stroud, A. Lueb, and S. Kato (2006), Biomass burning and anthropogenic sources of CO over New England in the summer 2004, *J. Geophys. Res.*, *111*(D23), doi:10.1029/2005JD006878.

Weller, R., O. Schrems, A. Boddenberg, S. Gäb, and M. Gautrois (2000), Meridional distribution of hydroperoxides and formaldehyde in the marine boundary layer of the Atlantic (48° N-35° S) measured during the Albatross campaign, *J. Geophys. Res.* , *105*(D11), 14401-14412, doi:10.1029/1999JD901145.

Westwater, E. R. (1993), Ground-based microwave remote sensing of meteorological variables, in *Atmospheric remote sensing by microwave radiometry*, edited, p. 145, Wiley, New York

Wittrock, F., H. Oetjen, A. Richter, S. Fietkau, T. Medeke, A. Rozanov, and J. P.Burrows (2004), MAX-DOAS measurements of atmospheric trace gases in Ny-Ålesund-Radiative transfer studies and their application, *Atmos. Chem. Phys.*, *4*(4), 955-966, doi:10.5194/acp-4-955-2004

Yokelson, R. J., J. G. Goode, D. E. Ward, R. A. Susott, R. E. Babbitt, D. D. Wade, I. Bertschi, D. W. Griffith, and W. M. Hao (1999), Emissions of formaldehyde, acetic acid, methanol, and other trace gases from biomass fires in North Carolina measured by airborne Fourier transform infrared spectroscopy, in *J. Geophys. Res.*, *104*(D23), 30109-30125, doi:10.1029/1999JD900817.

Yokelson, R. J., J. D. Crounse, P. F. DeCarlo, T. Karl, S. P. Urbanski, E. Atlas, T. Campos, Y. Shinozuka, V. Kasputin, and A. D. Clarke (2009), Emissions from biomass burning in the Yucatan, *Atmos. Chem. and Phys.*, *9*(15), 5785-5812, doi:/10.5194/acp-9-5785-2009.

Yurganov, L. N., W. W. McMillan, A. V. Dzhola, E. I. Grechko, N. B. Jones, and G. R. van der Werf (2008), Global AIRS and MOPITT CO measurements: Validation, comparison, and links to biomass burning variations and carbon cycle, *J. Geophys. Res.*, *113*(D9), doi:10.1029/2007JD009229.

Zarzana, K. J., K. Min, R. A. Washenfelder, J. Kaiser, M. Krawiec-Thayer, J. Peischl, J. A. Neuman, J. B. Nowak, N. L. Wagner, W. P. Dubè, J. M. Clair, G. M. Wolfe, T. H. Hanisco, F. N. Keutsch, T. B. Ryerson, and S. S. Brown (2017), Emissions of Glyoxal and Other Carbonyl Compounds from Agricultural Biomass Burning Plumes Sampled by Aircraft, *Environ. Sci. Technol.*, *51*(20), 11761-11770, doi:10.1021/acs.est.7b03517.

Chapter 4: MAX-DOAS observations of formaldehyde and glyoxal in Pantnagar, India

This chapter is primarily based on the research article of *Hoque et al* [2018b].

4.1 Introduction

As a part of the current study focusing on the VOC measurements in the South Asian region, the first continuous simultaneous MAX-DOAS observations of glyoxal and formaldehyde were performed at Pantnagar (15.18 °N, 102.56 °E). Pantnagar is situated in India and is a semi-urban site. One of the important features of Pantnagar is that it is situated in the Indo-Gangetic Plain (IGP) and is also a site of the SKYNET observational network. The MAX-DOAS instrument in Pantanagar started operation in January, 2017. Similar to the measurements and analysis of the Phimai data, the HCHO, CHOCHO, and NO_2 concentrations in Pantnagar were retrieved for the 0-1 km layer and analyzed for the observations from January 2017 to November 2017. The ratio of CHOCHO to HCHO concentrations (R_{GF}) was estimated. Similar to the dry season in Phimai, during spring and autumn season fire events occurred in Pantnagar. Thus, the R_{GF} under biomass burning was compared between both sites.

4.2 Observation Site and Methodology

Encompassing India, Nepal, and few parts of Bangladesh, the Indo-Gangetic Plain (IGP) is the most fertile region in the world. Such fertile lands and corresponding massive agricultural activities has turned IGP into one of the most densely inhabited regions in the world. The Pantnagar site is situated in the Indian part of the IGP region. The distance of Pantnagar from New Delhi (capital of India) is around 225 km and it is a semi-rural site. (Fig. 4.1). Towards the east and north direction of Pantnagar, the Himalayan mountain ranges are lying. The lower altitude land (i.e., flat topography) are situated in the south and west direction of the site. The main road which connects Pantnagar to the adjacent cities is situated in the close vicinity of the measurement site and thus, the site is expected to have impact from vehicular emissions. Besides, a small local airport is situated approximately 3 km away from the site. Rudrapur and Haldwani are two nearby cities situated ~12 km southwest and ~25 km northeast, respectively, from Pantnagar, which host some small scale industries. There are four seasons in Pantnagar: (1) winter (December to February), (2) spring (March to May), (3) summer monsoon (June to August), and (4) autumn (September to November).

Figure 4.1. Pantnagar site is situated ~225 km northeast of New Delhi, the capital of India.

Figure 4.2. MAX-DOAS instrument employed at the Pantnagar site

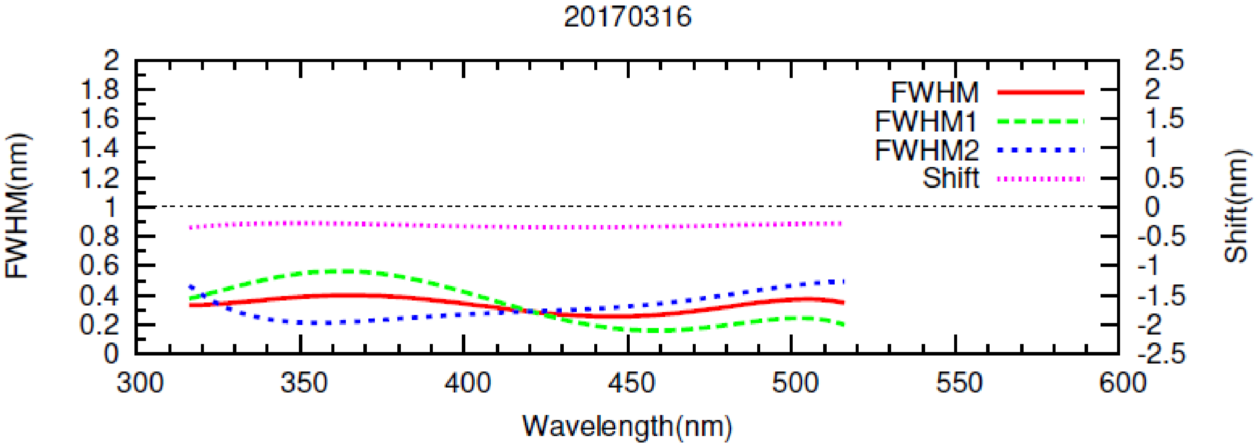


Figure 4.3. Example of wavelength calibration on March 16, 2017

The MAX-DOAS instrument used for observations in Pantnagar was the similar instrument used for observations in Phimai. The instrument was installed in a building named "Biotech Bhavan", which is situated in the G.B. Pant University of Agriculture and Technology campus in Pantnagar. The instrument started its operation on 17 January 2017. The elevation angle (EL) sequence of 3°, 4°, 5°, 6°, 8°, and 70° was used and the same sequence of ELs was repeated every 15 min. The lowest EL used in Pantanagar was 3°, whereas the lowest EL used for measurements in Phimai was 2°. Because, in Pantanagar, there were some obstruction in the field of view of the instrument below 3°. A question might arise whether the lowest EL of 3° can account for sufficient surface information or not. However, *Irie et al* [2008a, 2015] used the lowest EL of 3° and found good agreement with LIDAR measurements. Thus, EL of 3° can account for sufficient surface information.

Similar to measurements in Phimai, the off- axis ELs used in Pantnagar were less than 10°, which was based on the study of *Irie et al.* [2015] to minimize the fitting errors of oxygen collision complex (O_4). The spectral range of the measured spectra were 310 – 515 nm. For the wavelength calibration, the solar spectrum reported by *Kurucz et al*. [1984] was used, similar to the procedures adopted for the measurements in Phimai. The spectral resolution at 357 and 476 nm was ~ 0.3 nm. Similar to analysis of the Phimai data, the JM2 algorithm [*Irie et al.,2008a*, 2008b, 2011, 2015] was used to retrieve the trace gas and aerosol profile information. The analysis procedure was similar to that explained in details in chapter 3. Thus, only a brief description is provided here. In the first step, the spectra covering the spectral range of 310-515 nm were analyzed using the DOAS method [*Platt*, 1994], which yields the differential slant column density (ΔSCD) of the absorbers of interest. The HCHO and CHOCHO retrieval fitting windows were similar to that of used for the observations in Phimai. O_4 ΔSCD values were retrieved for the strong O_4 absorptions around 360 and 477 nm [*Irie et al.* 2009, 2011]. The HCHO, CHOCHO, and other interfering species absorption cross section data were similar to that used for the observations in Phimai (Table 3.1 of chapter 3 and *Hoque et al* [2018a]).

The obtained O_4 ΔSCD values from the DOAS fittings were input to the aerosol retrieval algorithm which yielded the aerosol optical depth (AOD) and vertical profiles of aerosol extinction coefficient (AEC). The aerosol retrieval algorithm is based on optimal estimation [*Irie et al*. 2008a; *Rodgers* 2000]. The radiative transfer calculations to create the lookup table (LUT) for the box air mass factor (A_{box}) were performed using the JACOSPAR radiative transfer model [*Irie et al*. 2015]. Then the A_{box} LUT was used to determine the AOD and AEC. Similarly, the obtained HCHO and CHOCHO ΔSCD from the DOAS analysis were input into the trace gas retrieval algorithm which yielded the vertical column densities (VCDs) and vertical profiles of HCHO and CHOCHO, respectively. The trace gas wavelengths were slightly different from the wavelengths used (357 and 476 nm) for the A_{box} calculations in the LUT. To overcome such difference in the representative wavelengths, the AOD at the wavelengths of HCHO and CHOCHO was determined from the AOD at the wavelength closer to the HCHO and CHOCHO wavelengths, respectively. For such wavelength conversion of AOD, Ångström exponent value of 1.00 was used. The AOD values recalculated at the representative wavelength of the trace gas retrievals were used to match the respective A_{box} values in the LUT.

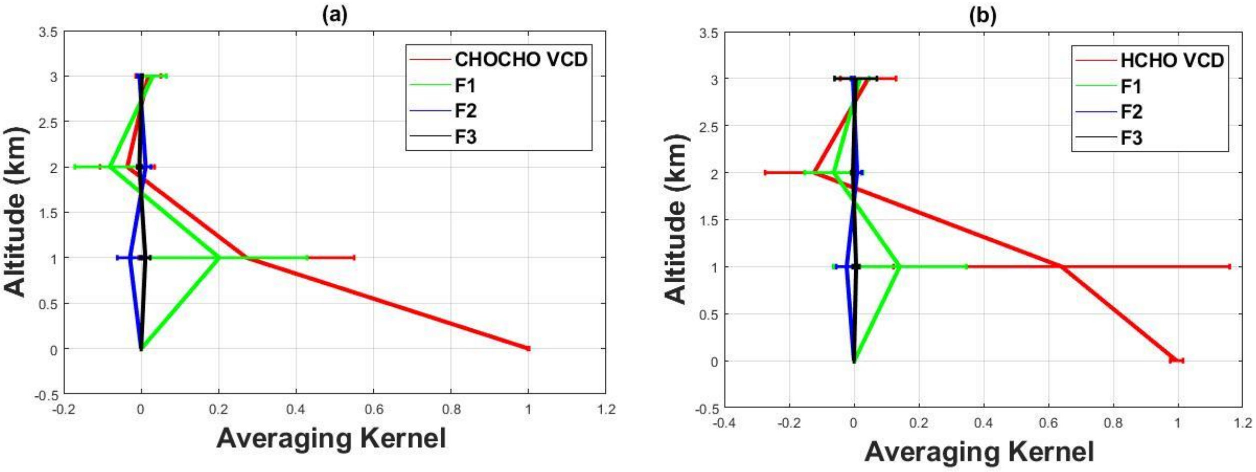


Figure 4.4. Mean averaging kernel for CHOCHO (left panel) and HCHO (right panel) retrieval for the whole observation period in Pantnagar

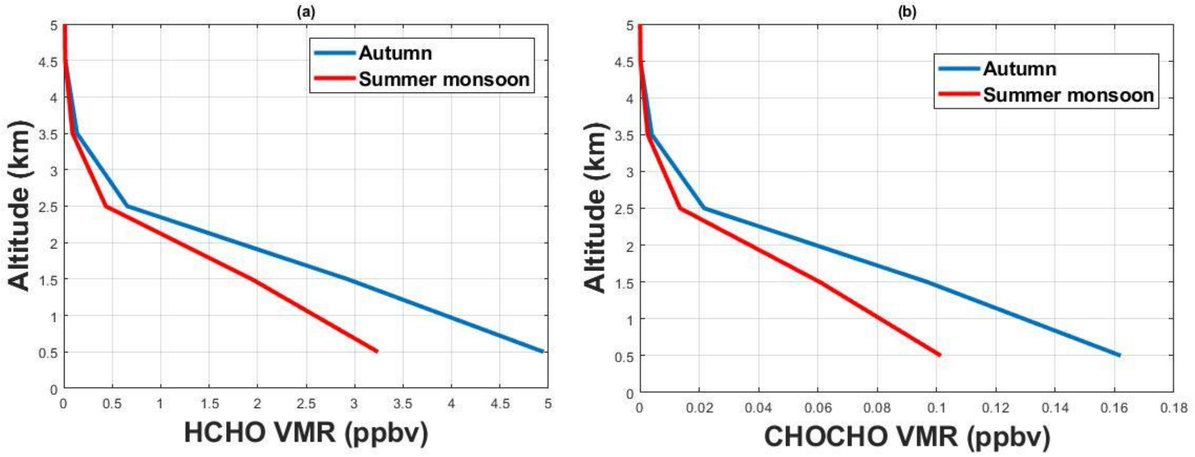


Figure 4.5. Mean HCHO (left panel) and CHOCHO (right panel) profiles during the autumn and summer monsoon in Pantnagar.

Figure 4.4 shows the mean averaging kernel of the CHOCHO and HCHO retrievals for the whole period in Pantnagar. The highest values were observed for AOD, followed by f_1 and f_2, indicating that, the retrrieval was done scaling the a priori information similar to that observed for the measurements in Phimai. The mean degree of freedom of the signal of the CHOCHO and HCHO retrievals for the whole observation period were 1.21 and 1.15, respectively.

Figure 4.5 shows the mean HCHO and CHOCHO profiles for the autumn and summer moonson seasons. The higher concentration of CHOCHO and HCHO during the autumn could potentially be due to the influence of biomass burning.

The random and systematic error were calculated similarly using a method explained in chapter 3. The random error was estimated from the residuals of the HCHO and CHOCHO ΔSCD. The systematic error was estimated by varying the AOD by additional ± 30 and ±50% for the CHOCHO and HCHO retrieval respectively. These statistics were based on the uncertainty in the aerosol retrieval in the UV and visible region [*Irie et al*, 2011]. The estimated errors for the observations in Pantnagar are summarized in Table 4.1.

Table 4.1. Estimated errors (%) for the retrieved CHOCHO and HCHO VMRs for the 0-1-km layer.

Component	Total error	Random error	Systematic error	Error due to bias in EL
CHOCHO	25	14	20	5
HCHO	37	20	31	5

Data screening criterion to take into account the influence of clouds was similar to that used for the measurements in Phimai. The similar two-step screening was used. In the first step, retrievals for retrieved AOD > 3, which is the largest value in the LUT are removed which is expected to be mainly for optically thick clouds. Further data screening was based on residuals of the O_4 ΔSCD (< 10%), CHOCHO ΔSCD (< 50%), HCHO ΔSCD (< 50%), and the degree of freedom for the signal greater than 1.02.

From the simultaneous measurements of HCHO and CHOCHO at both sites, the R_{GF} was estimated and compared under the influence of biomass burning. Although similar instruments were used at both sites, yet the instrumental performance can impact the comparison of the R_{GF}. To estimate such systematic error, a test was conducted with two MAX-DOAS instrument directing at the same direction in Chiba. This test was mainly conducted to investigate the consistency between MAX-DOAS systems for potential application of spatial distribution studies.

Simultaneous observations by two MAX-DOAS systems (hereinafter denoted as #1 and #2, respectively) at the same viewing azimuth direction (north) were carried out at Chiba University from October 18 to November 11, 2013. The complete sequence of elevation angle scan was repeated every 30 min for both systems. For each AOD 476 and AEC 476 (0-1 km) data, correlation plots of data from simultaneous observations by two north-viewing MAX-DOAS systems are shown in Fig 4.6. Very tight correlations, with R^2 values of 0.83 and 0.97 for AOD476 and AEC476, respectively, were readily seen in Fig. 4.6. For AOD476 and AEC476, the mean differences (MAX-DOAS #2 minus #1) were calculated to be as small as 0.00±0.11 and -0.01±0.02 km^{-1}, respectively.

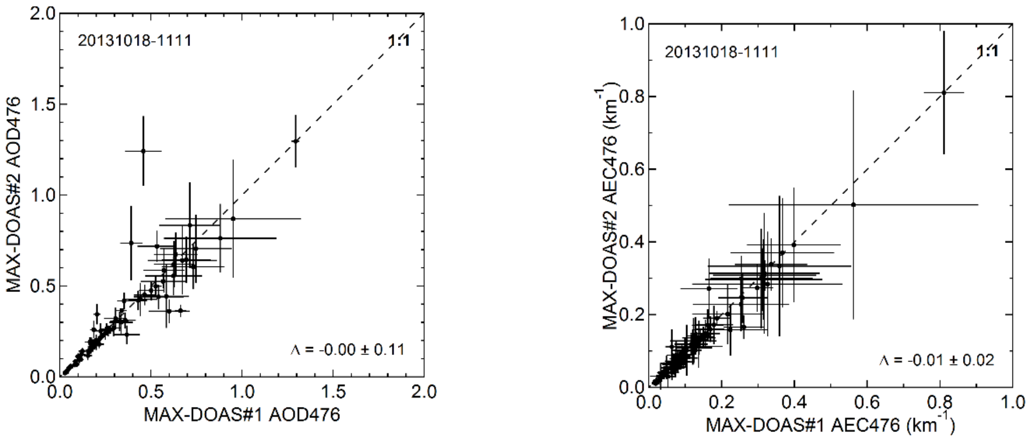


Figure 4.6. Correlation plots of (left) AOD476 and (right) AEC476 (0-1 km) data from simultaneous observations by two north-viewing MAX-DOAS systems. The observations were made at Chiba from October 18 to November 11, 2013. Uncertainties estimated in the retrieval procedures are represented by error bars.

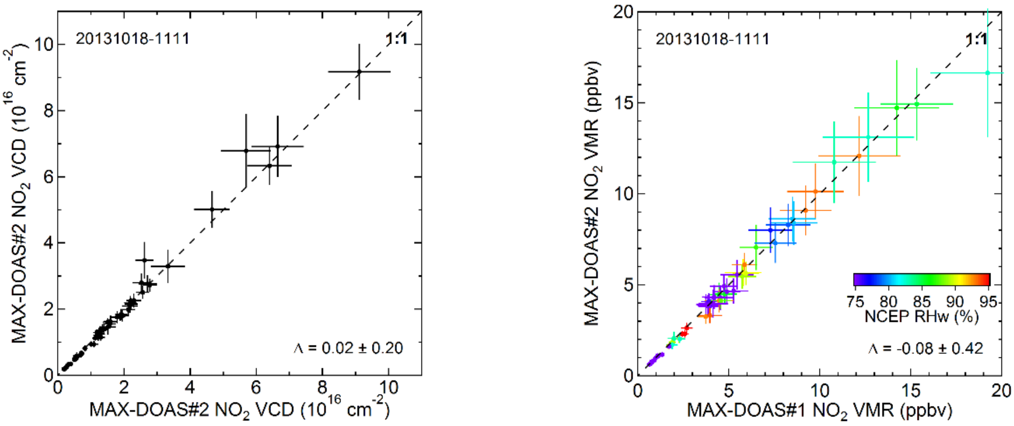


Figure 4.7. Correlation plots of (left) NO_2(vis) VCD and (right) VMR (0-1 km) data from simultaneous observations by two north-viewing MAX-DOAS systems. The observations were made at Chiba from October 18 to November 11, 2013. Uncertainties estimated in the retrieval procedures are represented by error bars. The mean difference (Δ) and its standard deviation are given in each plot. The unit is the same as x and y axis. The mean VCD and VMR values for MAX-DOAS #1 data were 1.7×10^{16} molecules cm^{-2} and 4.8 ppbv, respectively. The determination coefficients (R^2) were 0.99 and 0.99, respectively. Colors indicate the surface RHw values from NCEP.

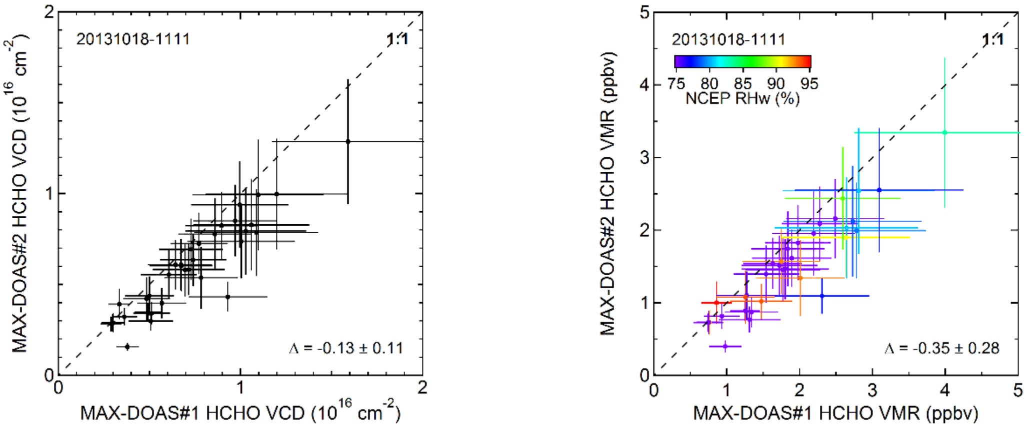


Figure 4.8. Correlation plots of (left) HCHO VCD and (right) VMR (0-1 km) data from simultaneous observations by two north-viewing MAX-DOAS systems. The observations were made at Chiba from October 18 to November 11, 2013. Uncertainties estimated in the retrieval procedures are represented by error bars. The mean difference (Δ) and its standard deviation are given in each plot. The unit is the same as x and y axises. The mean VCD and VMR values for MAX-DOAS #1 data were 0.74×10^{16} molecules cm^{-2} and 1.9 ppbv, respectively. The determination coefficients (R^2) were 0.82 and 0.81, respectively. Colors indicate the surface RHw values from NCEP.

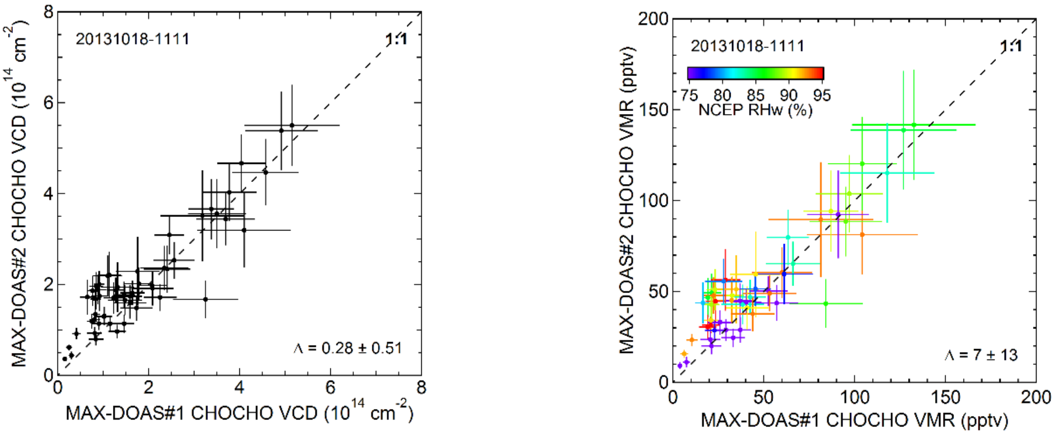


Figure 4.9. Correlation plots of (left) CHOCHO VCD and (right) VMR (0-1 km) data from simultaneous observations by two north-viewing MAX-DOAS systems. The observations were made at Chiba from October 18 to November 11, 2013. Uncertainties estimated in the retrieval procedures are represented by error bars. The mean difference (Δ) and its standard deviation are given in each plot. The unit is the same as x and y axises. The mean VCD and VMR values for MAX-DOAS #1 data were 1.8×10^{14} molecules cm^{-2} and 47 pptv, respectively. The determination coefficients (R^2) were 0.83 and 0.84, respectively. Colors indicate the surface RHw values from NCEP.

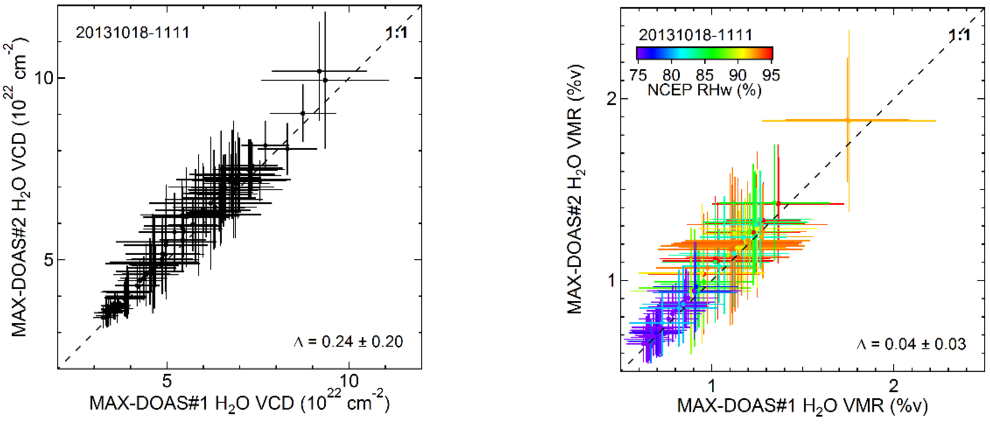


Figure 4.10. Correlation plots of (left) H_2O VCD and (right) VMR (0-1 km) data from simultaneous observations by two north-viewing MAX-DOAS systems. The observations were made at Chiba from October 18 to November 11, 2013. Uncertainties estimated in the retrieval procedures are represented by error bars. The mean difference (Δ) and its standard deviation are given in each plot. The unit is the same as x and y axises. The mean VCD and VMR values for MAX-DOAS #1 data were 5.3×10^{22} molecules cm^{-2} and 1.0 %v, respectively. The determination coefficients (R^2) were 0.83 and 0.84, respectively. Colors indicate the surface RHw values from NCEP. The determination coefficients (R^2) were 0.99 and 0.99, respectively.

For NO_2 and H_2O, the mean differences were very small at 1-5%. Relatively large mean differences were seen for HCHO and CHOCHO but were within ±18% and were on the same order of magnitude of the uncertainty estimated in the retrieval (given with error bars in Figs 4.6 - 4.10). Furthermore, for all species of NO_2, HCHO, CHOCHO, and H_2O, very tight correlations with high R^2 (ranging from 0.84 to 0.99) were obtained (Tables 4.2 & 4.3). These results strongly support the idea that spatial distribution can be studied by multiple different-azimuth-viewing MAX-DOAS instruments. The similar systematic differences to those in comparisons for HCHO and CHOCHO were seen in corresponding ΔSCD data. This suggests that an unresolved element yielding systematic errors might remain in the DOAS fitting algorithm of JM2 for relatively weaker absorbers such as HCHO and CHOCHO. According to the results from the comparisons made here, the systematic errors were estimated to be ±18%, which is within the estimated error for HCHO and CHOCHO for both Phimai and Pantnagar.

Table 4.2 Differences (MAX-DOAS #2 - MAX-DOAS #1) seen in VCD data.

Parameter	Absolute Difference ($molec.cm^{-2}$)	Relative Difference (%)	R^2
NO_2 VCD	$(0.02\pm0.20)\times10^{16}$	1±12	0.99
HCHO VCD	$(-0.13\pm0.11)\times10^{16}$	-18±15	0.87
CHOCHO VCD	$(0.28\pm0.51)\times10^{14}$	16±28	0.83
H_2O VCD	$(0.24\pm0.20)\times10^{22}$	5±4	0.99

Table 4.3 Differences (MAX-DOAS #2 - MAX-DOAS #1) seen in VMR data.

Parameter	Absolute Difference	Relative Difference (%)	R^2
NO_2 VMR	-0.08±0.42 ppbv	2±9	0.99
HCHO VMR	-0.35±0.28 ppbv	-18±15	0.87
CHOCHO VMR	7±13 pptv	15±28	0.84
H_2O VMR	0.04±0.03 %v	4±3	0.99

4.3. Results and discussion

4.3.1 Seasonal variation of HCHO and CHOCHO

Figure 4.11 shows the time series of monthly mean HCHO, CHOCHO, and NO_2 concentrations in Pantnagar. The concentrations are shown for the 0-1 km layer. The concentrations of all the three trace gases in Pantnagar showed similar seasonal variations, with high concentrations during spring and autumn. All three trace gas concentrations peaked in May during spring and around October or November during autumn. Such enhancement in the seasonal variation is likely due to the fire activities around Pantnagar. In general, biomass burning activities vary with geographical locations within India but the dominant biomass burning events occur in spring [*Venkataraman et al.* 2006].

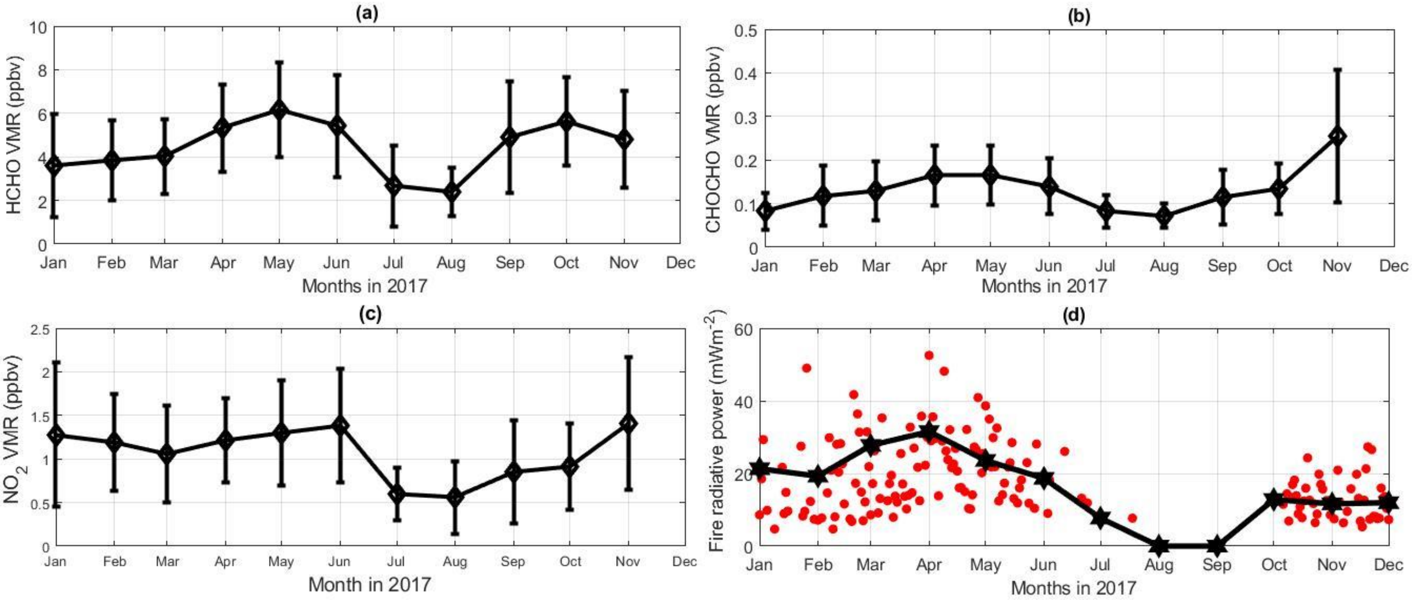


Figure 4.11. The seasonal variation of (a) HCHO and (b) CHOCHO, (c) NO_2 concentration in Pantnagar retrieved from the MAX-DOAS observations. The trace gas concentrations are plotted for the 0-1 km layer. The error bars in the monthly means indicate the standard deviation of the means. (d) Daily and monthly mean fire radiative power (FRP) values covering the period of the MAX-DOAS observations in Pantnagar. FRP values were averaged within an area of 100 km from the Phimai site. The red dots in (d) are the daily mean FRP values whereas the black solid line indicates the monthly mean FRP values.

Previous studies based on Pantnagar [i.e., *Kumar et al*. 2011; *Bhardwaj et al*. 2016] have found biomass burning influencing the trace gas concentration at the site. Figure 4.11(d) shows the plot of fire radiative power (FRP) data plotted within an area of 100 km centering at the Pantnagar site. As mentioned in chapter 3, FRP observations from the MODIS satellite can be used as tracer for fire locations around the globe [*Justice et al*. 2002]. The FRP data were obtained from the similar source mentioned in chapter 3. The majority of the fire were detected during the spring and autumn seasons in Pantnagar, which also matches with enhancements of HCHO, CHOCHO, and NO_2 concentrations. The fire counts were higher during the spring, indicating stronger biomass burning influence compared to autumn. According to *Venkataraman et al* [2006], majority of the biomass burned in the western Indo-Gangetic Plain are crop residues and smaller amounts are in form of forest fires. This is also consistent with *Streets et al* [2003] and *Bhardwaj et al* [2016], reporting crop residues as the dominant form of biomass over South Asia.

However, Pantnagar can also be influenced by anthropogenic emissions from vehicular sources and industrial emissions from Haldwani and Rudrapur. Figure 4.12 shows some example trajectories in Pantnagar. On June 16 and February 14, 2017 the air masses travelled over Haldwani and Rudrapur before arriving at Pantnagar. On August 22 and February 28 air masses arrived at Pantnagar from the southeast and northwest direction, respectively. The HCHO and CHOCHO concentrations on June 16 and February 14 were higher compared to the other two days, which could potentially be due to the influence from the emissions from the nearby cities. Thus, both pyrogenic and anthropogenic emissions can influence the trace gas concentrations in Pantnagar.

Similar to the measurement in Phimai, the temperature dependences of HCHO and CHOCHO were also investigated. Figure 4.13 shows the seasonal variation of HCHO, CHOCHO, and temperature. Higher temperatures were observed during spring (MAM) and lower temperatures during winter (JF). The highest and lowest temperatures were observed during June (i.e., summer monsoon) and January (i.e., winter), respectively. Low correlations among HCHO, CHOCHO, and temperature were observed in Pantnagar. *Guenther et al* [2006] estimated global isoprene emission using the MEGAN (Model of Emissions of Gases and Aerosols from Nature) model and evaluated the change of isoprene emissions with respect to change of temperature and leaf area index (LAI). Isoprene emissions showed almost no change with respect to LAI in the IGP region indicative of strong vegetation throughout the year. Significant changes in the isoprene emissions with respect to change in temperature were observed in Central and southern part of India. However, over the IGP region, temperature difference had minimal impact on the isoprene emissions. This also agrees with the lower correlations observed among temperature, HCHO, and CHOCHO in Pantnagar. Moreover, during autumn the temperature was low, yet enhanced HCHO and CHOCHO concentrations were observed in Pantnagar.

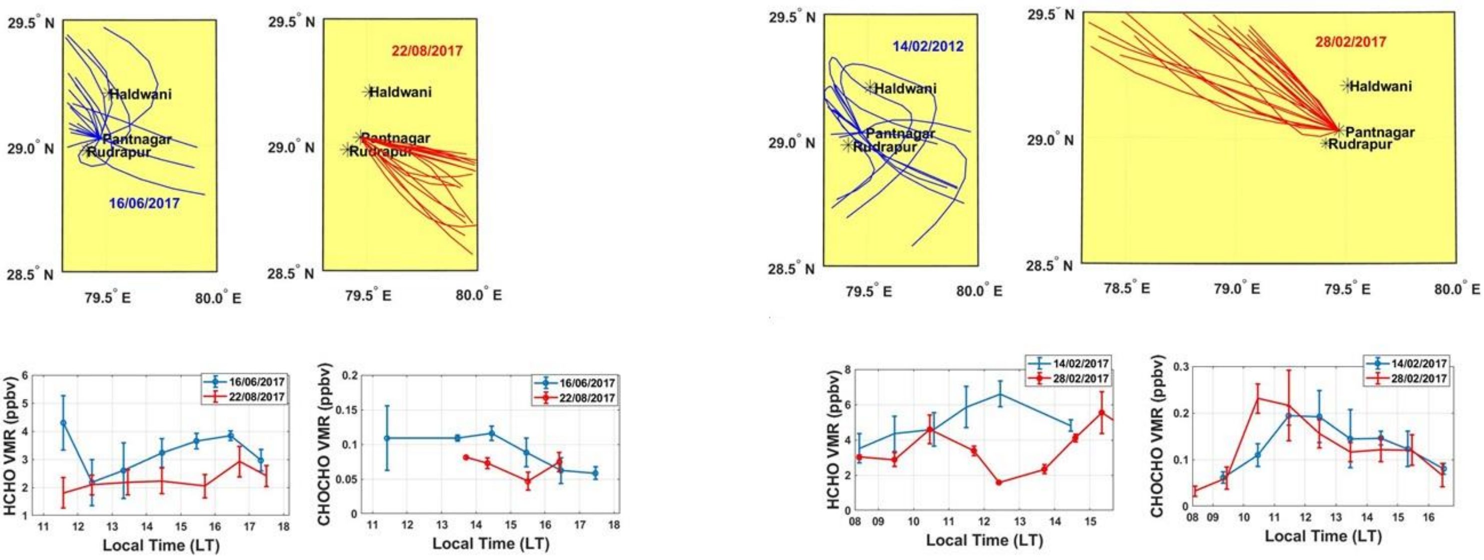


Figure 4.12. (top panel) The backward trajectories initiated at Pantnagar site on June 16, August 22, February 14, and February 28, 2017. (lower panel) Hourly concentration of HCHO and CHOCHO on the respective dates.

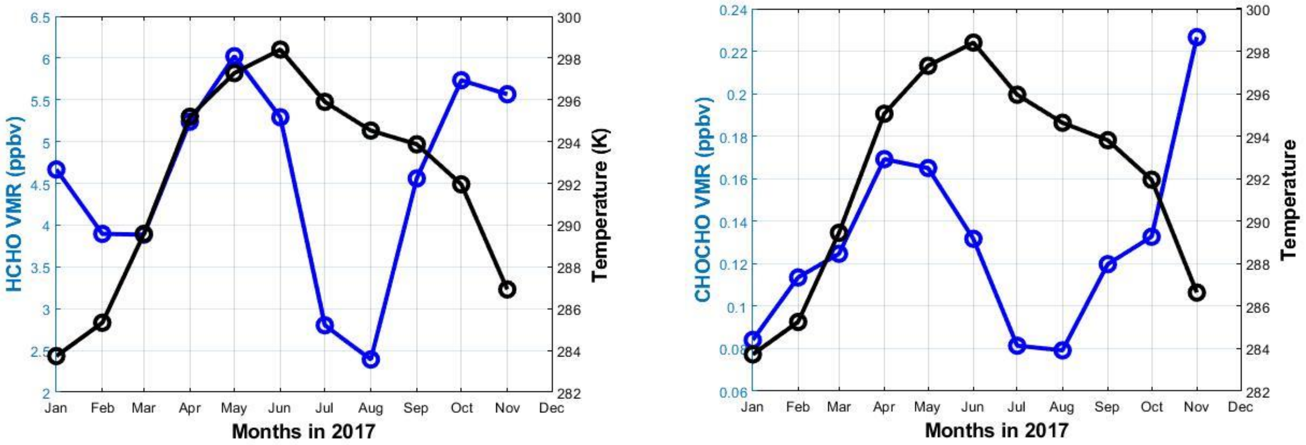


Figure 4.13. (left panel) Monthly variation of temperature and HCHO VMR in Pantnagar. (b) Seasonal variation of temperature and CHOCHO VMR in Pantnagar.

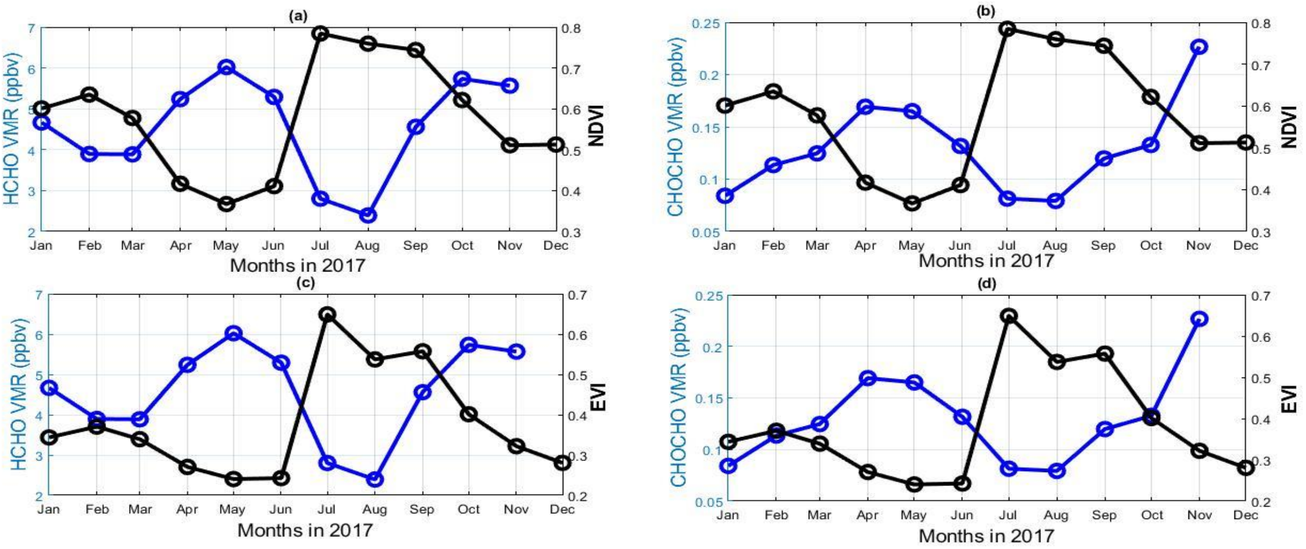


Figure 4.14. Seasonal variations in (a) HCHO and NDVI, (b) CHOCHO and NDVI, (c) HCHO and EVI, and (d) CHOCHO and EVI, in Pantnagar. NDVI and EVI over Pantnagar for the period of January – November. The spatial resolutions of the NDVI and EVI measured by the MODIS satellite were 0.05° × 0.05°.

Figure 4.14 shows the seasonal variation of normalized difference vegetation index (NDVI) and enhanced vegetation index (EVI) in Pantnagar. The NDVI and EVI showed different seasonal variation in Phimai, which was attributed to different response of NDVI and EVI to the phenological stages of the plants. However, in Pantnagar, both NDVI and EVI showed similar seasonal response with higher values during the summer monsoon (JJA). During spring and autumn, when HCHO and CHOCHO concentrations were higher, both NDVI and EVI values were lower compared to the summer monsoon period. This indicates a poor correlation among HCHO, CHOCHO, NDVI, and EVI. Thus, the contribution of biogenic sources to the observed OVOC levels in Pantnagar is likely to be less significant compared to the anthropogenic sources and biomass burning.

4.3.2 Diurnal variation of HCHO, CHOCHO, and NO_2

Figure 4.15 shows the hourly mean concentrations of CHOCHO, HCHO and NO_2 during the different seasons in Pantnagar. The CHOCHO diurnal variations during all seasons were similar, with higher values during the daytime and lower values during the late afternoon. The peak concentration in the CHOCHO diurnal variation were observed around noon. Higher CHOCHO concentrations during spring and autumn are indicative of stronger CHOCHO emission sources which potentially might be due to fire activities. During early hours of winter season almost similar CHOCHO concentrations were observed. This could be potentially due to the fog-induced suppression of photochemical activities. Because, during winter, especially in January, the region is generally covered with thick fog [*Ojha et al*, 2012].

Except during autumn, HCHO diurnal variations were similar with higher values in the early morning and late afternoon, and lower values during mid days. Similar diurnal variations in HCHO concentrations were reported from MAX-DOAS observations in Beijing [*DeSmelt et* *al* 2015, *Vlemix et al* 2015]. However, different diurnal variations in HCHO have been reported in literature [*Park et al*. 2018, *Pang et al* 2010, *Rappengluck et al* 2010] and have been attributed to the differences in the HCHO sources and loss mechanism. The higher HCHO concentrations in the early morning could be potentially due to the night time buildup of HCHO from vehicular emissions. The lower HCHO concentration during the day time is likely indicative of stronger photochemical removal by photolysis reactions. In contrast, the higher concentrations in the late afternoon could be potentially due to weaker photolysis sink of HCHO. Nighttime removal of HCHO occurs by reaction with the nitrate radicals [*Finlayson-Pitts and Pitts, Jr.*, 2000, *Senfields and Pandis* 1998]. However, the loss mechanism by reaction with nitrate radicals during nighttime is not considered important in urban environments, where sufficient emissions of nitrogen oxide is available which can suppress the concentrations of nitrate radicals [*Tanner et al* 1988, *Anderson et al* 1996]. However, the nighttime sinks of HCHO in Pantnagar requires further investigation and supporting information.

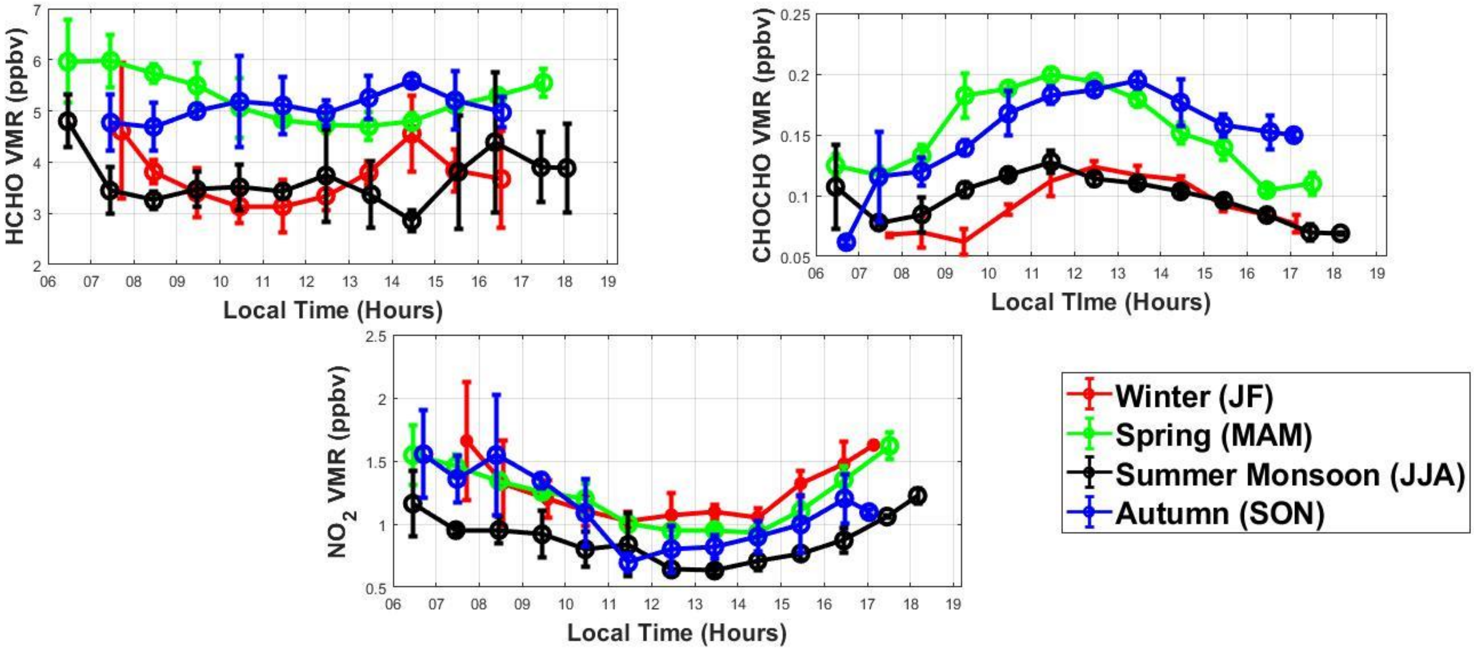


Figure 4.15. Diurnal variation in HCHO, CHOCHO, and NO_2 in Pantnagar. The observation data covering the analysis period i.e., from January to November 2017 were used to calculate the diurnal means of the trace gases.

Figure 4.15 also shows the diurnal variation of NO_2 in Pantnagar. Higher NO_2 concentrations were observed during morning and late afternoon, whereas midday NO_2 concentrations were lower. Such diurnal variations of NO_2 are common in rural environments. The NO_2 fate in the atmosphere is determined by its photolysis disassociation to NO, reaction with hydroxyl radicals, and the reactions among NO, ozone and hydroxyl radicals. The photolysis of NO_2 forms NO, which leads to formation of ozone. However, O_3 is oxidized by NO to form NO_2 again. Thus for a net O_3 production, other mechanism is required to form NO_2 from NO. The interaction of NO with hydroperoxyl radicals provides such mechanism to form NO_2 from NO without oxidizing O_3 leading to a net production of ozone.

4.3.3 Estimation of R_{GF} in Pantnagar

The mean R_{GF} estimated for the Pantnagar site was 0.029 ± 0.006. The observation data from January – November 2017 were used to calculate the mean R_{GF} value. The uncertainty was derived from the uncertainties in individual HCHO and CHOCHO retrievals. The consideration of the contribution of systematic and random errors in the HCHO and CHOCHO retrievals to the ratio were made similarly to the Phimai site. The total error in the mean ratio was estimated to ~21%. The estimated mean R_{GF} was compared with the literature values and studies and was found to be consistent with the range of the literature values [*Munger et al.* 1995; *Grosjean et al* 1996; *Possanzini et al.* 2007; *Choi et al.* 2010; *Irie et al.* 2011; *DiGangi et al*. 2012; *Kaiser et al.* 2015]. Similar to the measurements in Phimai, the seasonal mean R_{GF} was estimated for the Pantnagar site (Fig. 4.16). Monthly mean R_{GF} values at Pantnagar varied between ~0.020 and ~0.039. Figure 4.16(b) shows the diurnal variation in R_{GF} at Pantnagar. Comparing the diurnal R_{GF} values among different season, no significant differences in the R_{GF} values were observed and the diurnal R_{GF} were found to be mostly less than 0.04. The R_{GF} estimated during different season in Pantnagar were also almost similar and was ~ 0.030 (Table 4.4).

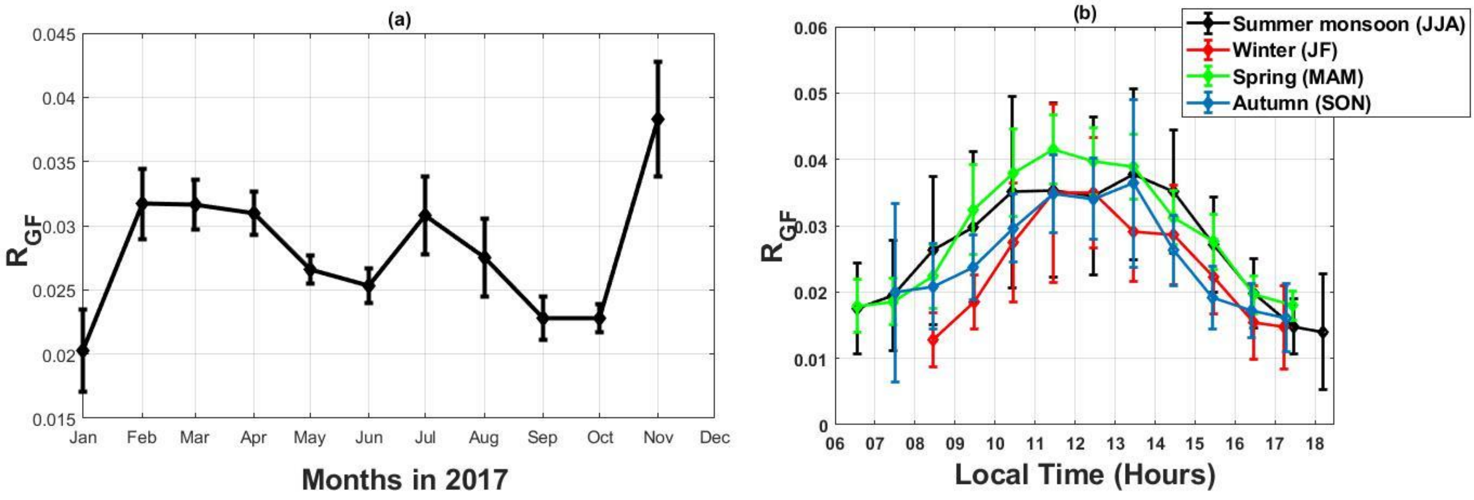


Figure 4.16. (a) Monthly mean R_{GF} estimated for individulas months of the during different seasons in Pantnagar. The observations for the whole period were used to estimate the monthly mean R_{GF}. (b) Hourly averaged diurnal variations of R_{GF} during different seasons in Pantnagar.The error bars in (a) and (b) represent the 2σ standard error.

Table 4.4. Estimated R_{GF} during seasons influenced by biomass burning in Phimai and Pantnagar. The errors represent the 2σ standard error.

Season and site	R_{GF}
Winter (JF), Pantnagar	0.029 ± 0.003
Spring 2017 (MAM), Pantnagar	0.030 ± 0.002
Summer Monsoon (JJA), Pantnagar	0.030 ± 0.004
Autumn 2017 (SON), Pantnagar	0.026 ± 0.002
Dry Season 2015 (JFMA), Phimai	0.030 ± 0.001
Dry Season 2016 (JFMA), Phimai	0.028 ± 0.002

The R_{GF} study by *Vrekoussis et al*. [2010] using global observations of HCHO and CHOCHO, found the range of R_{GF} between 0.04 – 0.06 when the dominant VOC sources were biogenic whereas R_{GF} values tend to be less than 0.03 under dominant anthropogenic VOC emissions (also high NO_2 concentrations). Almost the similar range of R_{GF} values (0.04 -0.07) under dominant biogenic VOC emissions were also reported by *Kaiser et al*. [2015]. During spring and autumn, significant fire events occured around the Pantnagar site. Biomass burning during the spring time in India has a significant impact on the annual HCHO variation in India [*Fu et al* 2007]. However, from our current analysis of the data in Pantnagar, it is hard to judge the contribution of biomass burning and anthropogenic emissions to the trace gas levels during spring and autumn. Thus, on a quantitative sense, the enhancement of HCHO and CHOCHO and the corresponding lower R_{GF} observed during spring and autumn is likely due to the influence of the both anthropogenic and pyrogenic emissions. In Fig. 4.13, the plots show low correlation coefficients among temperature and the trace gas (HCHO and CHOCHO) concentrations. However, satellite data analysis by *Mahajan et al* [2015] found higher concentration of HCHO over the IGP throughout year and attributed it to high vegetation. Our findings of lower correlation among temperature, HCHO, and CHOCHO concentrations were further supported by the low correlation among the vegetation indexes, HCHO and CHOCHO (Fig. 4.14). The lower correlations observed in Figs. 4.13 and 4.14 are likely an indication of biogenic VOC emission having less influence on the VOC levels in Pantnagar. Thus, stronger impact of biomass burning and anthropogenic emissions are mainly responsible for the R_{GF} values being mostly less than 0.04 in Pantnagar.

4.3.4 Comparison of R_{GF} under the influence of anthropogenic and pyrogenic emissions

Pronounced biomass burning occurred during the dry season in Phimai [*Sugimoto et al* 2015, *Hoque et al* 2018a]. The Pantnagar site was also likely influenced by anthropogenic and pyrogenic emissions. Thus, the R_{GF} values under such conditions at the two sites can be compared. The diurnal R_{GF} values during different seasons in Pantanagar and during dry seasons in Phimai are shown in Fig. 4.17. During dry seasons in Phimai R_{GF} values were found to be < 0.03. Similarly, the R_{GF} values in Pantnagar was mostly lower than ~0.04. Moreover, the estimated mean R_{GF} values during the dry seasons in Phimai and different seasons in Pantnagar were mostly ~ 0.03 (Table 4.4). Figure 4.17 also shows the diurnal variation of R_{GF} estimated for the Chiba, an urban site in Japan. Chiba is one of the super sites of the SKYNET network. The MAX-DOAS instrumentation and retrieval of trace gases were similar to that of described for the Phimai and Pantnagar sites. The measurements from November 2014 to November 2016 were used to estimate the R_{GF} values. Due to an urban environment, the Chiba site is expected to be influenced by anthropogenic emissions. Similar to the Pantnagar site, the R_{GF} values in Chiba were mostly lower than 0.04 and no significant differences were observed among the seasons. Although *DiGangi et al.* [2012] found R_{GF} deviated towards higher values due to biomass burning event, the R_{GF} value was ~0.03. Moreover, study of R_{GF} under biomass burning by *Zarzana et al.* [2017] found that R_{GF} ranges from 0.008 to 0.11 with the mean values of approximately 0.04. Under the influence of anthropogenic VOC emissions, mean R_{GF} values less than 0.04 were also found by *Ortega et al.* [2015] and *Miller et al.* [2014]. Thus, when biomass burning and anthropogenic emissions are the dominant VOC sources, then the mean R_{GF} value shows a tendency to be <0.04.

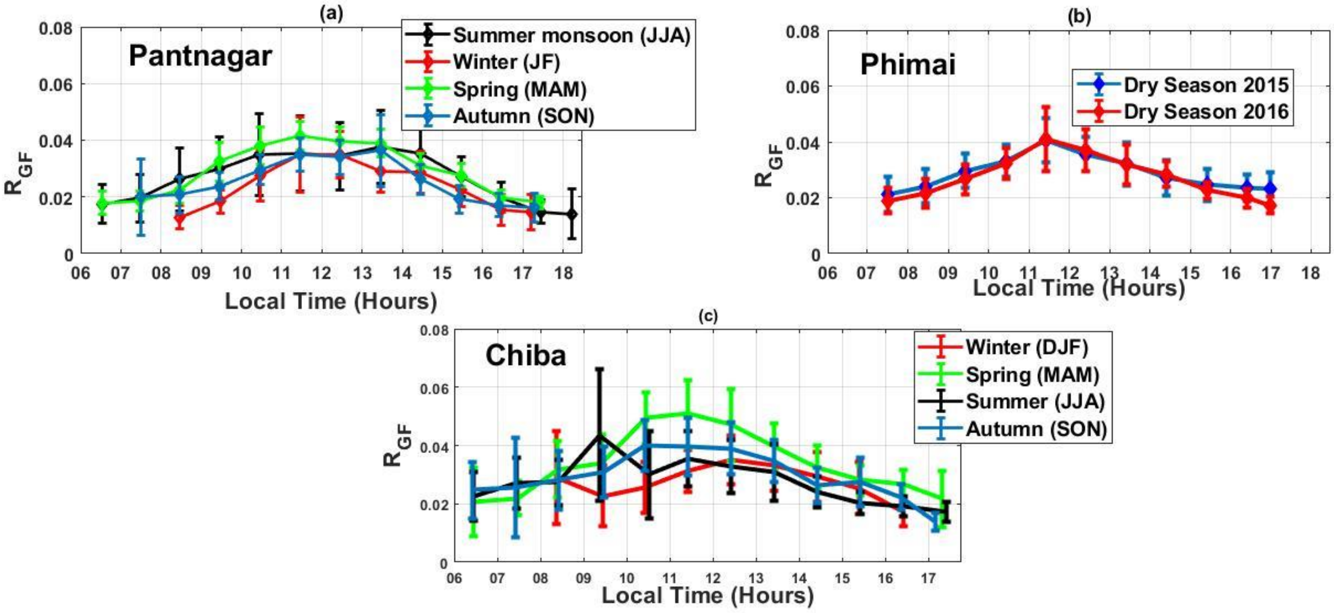


Figure 4.17. The diurnal variations of R_{GF} during different seasons in (a)Pantnagar, (b) Phimai, and (c) Chiba.The error bars in the plots are indicating the 2σ standard error of the mean values.

However, as different biomass fuels produce different levels of HCHO and CHOCHO, this will also impact the R_{GF} values observed during fire events [*Akagi et al.* 2013; *Stcokwell et al.* 2014]. At both regions (Phimai and Pantnagar), burning the remaining waste after the crop harvest is a common practice. Due to the different yields of HCHO and CHOCHO from different crop residue types, R_{GF} values can also be different. If we assume that biomass burning has the most significant impact on the VOC levels during spring and autumn, then the R_{GF} values less than

< 0.04 could potentially also be representative of crop residues as biomass fuel. From the R_{GF} values estimated for all the three SKYNET sites, we found almost consistent R_{GF} values (i.e., < 0.04) under the influence of anthropogenic and pyrogenic emissions.

4.3.5 The impact of NO_2 on R_{GF}

Figure 4.18 shows the seasonal variations of NO_2 and R_{GF} in Phimai and Pantnagar. The seasonal change of NO_2 concentration was unlikely correlated with that of the R_{GF} at Pantnagar, whereas significant correlation was observed at Phimai. The correlation among HCHO, CHOCHO, and NO_2 concentrations at Phimai and Pantnagar is depicted in Fig. 4.19. The significant source of NO_2 in Phimai was biomass burning during the dry season. In contrast, Pantnagar site had impact from local, regional, vehicular emissions, and biomass burning.

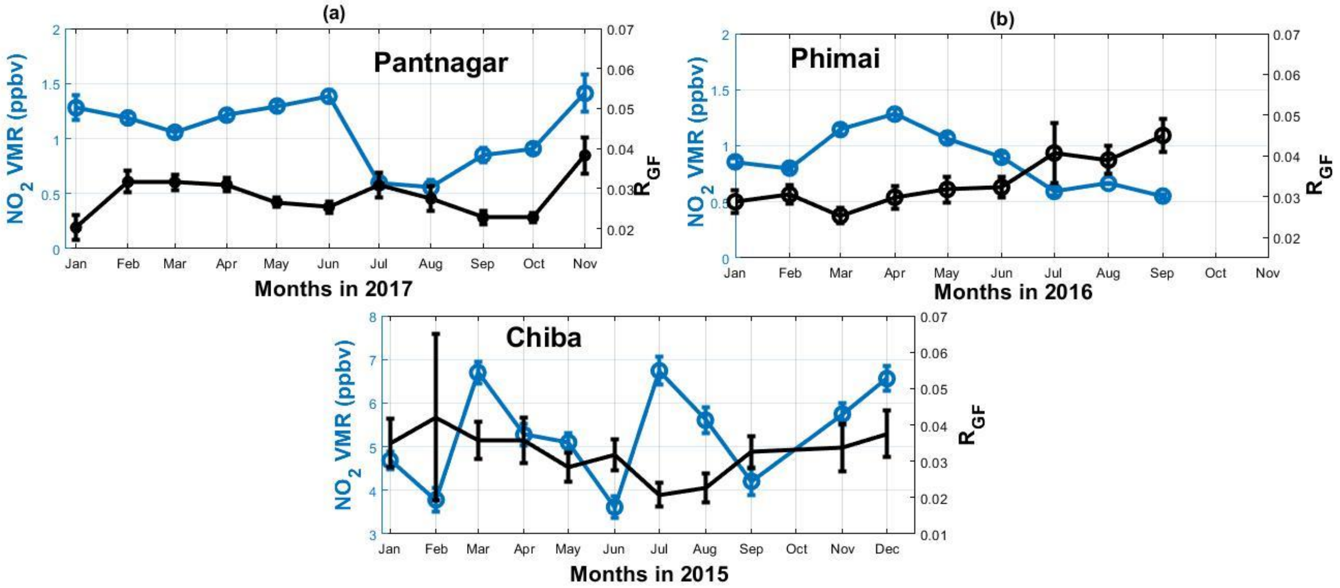


Figure 4.18. Time series of monthly mean NO_2 VMRs for the 0-1 km layer and the estimated monthly mean R_{GF} values in (a)Pantnagar, (b) Phimai, and (c) Chiba. The error bars in (a), (b),and (c) show the 2σ standard error.

The differences in the impact of NO_2 at both sites can lead to the different correlations among the trace gases. Utilizing models, *Li et al.* [2014] reported higher sensitivity of R_{GF} to lower NO_2 concentration. During the flight campaign over southern US, *Kaiser et al.* [2015] also observed no significant impact of NO_x on R_{GF} over urban areas. The Phimai site is characterized by low NO_2 (< 1 ppbv) environment, whereas the NO_2 concentrations are higher in Pantnagar. Moreover, the overall NO_x emissions in IGP are higher even within India [*Ghude et al.* 2008; *Ojha et al.* 2012]. Thus, the higher NO_2 concentrations in Pantanagar could be a probable reason of the R_{GF} response becoming less sensitive to NO_2. Figure 4.18 also shows the seasonal variation in NO_2 concentrations and R_{GF} in Chiba in 2015. As an urban site, Chiba is expected to be influenced by anthropogenic emissions. The monthly mean NO_2 concentrations in Chiba were mostly above ~4 ppbv and the monthly mean R_{GF} values are mostly < 0.04. No significant impact of NO_2 concentrations on the R_{GF} values was observed in Chiba.

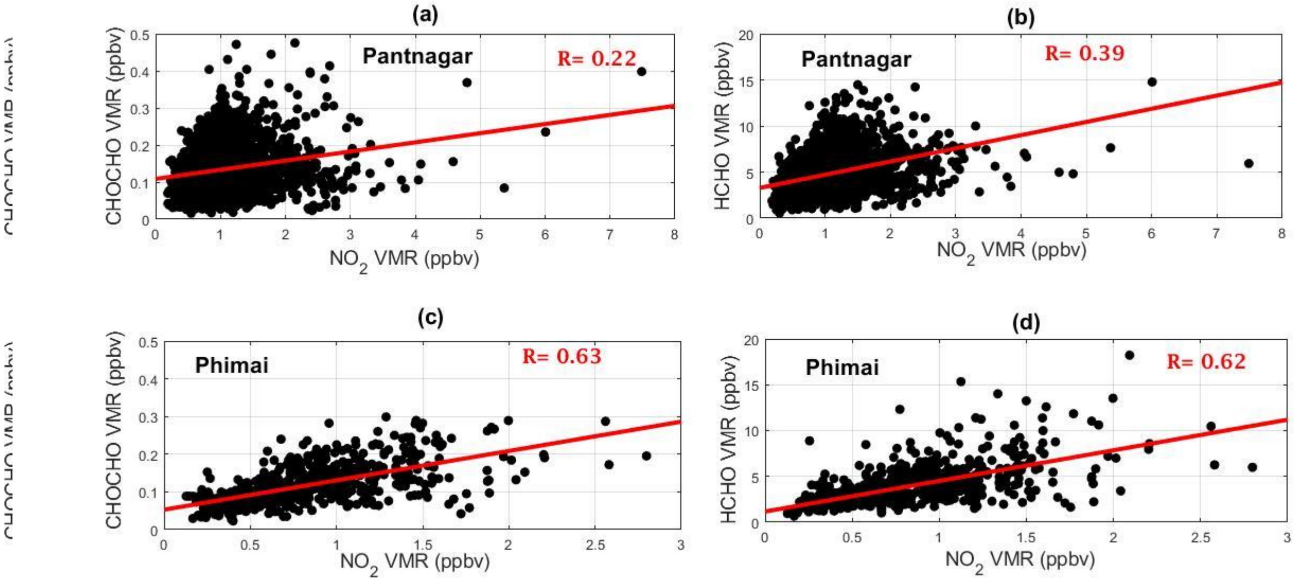


Figure 4.19. Correlation among NO_2, HCHO, and CHOCHO at Pantnagar ((a) and (b)) and Phimai ((c) and (d)). The daily mean values are used for the correlation plots. Data for the whole observation period at both sites are included in the plots.

Kaiser et al [2015] also assessed the impact of NO_x on R_{GF}, where the NO_x concentrations were mostly ~ 1 ppbv and the highest concentrations were ~ 3ppbv and found no significant impact of NO_x on R_{GF}.

The mean NO_2 concertation in Pantnagar was > 1 ppbv and no significant impact on R_{GF} was observed. Only at Phimai, where the mean NO_2 is < 1 ppbv, significant impact of NO_2 on R_{GF} was observed. From the observations at the three SKYNET sites and findings of *Kaiser et al* [2015], it is seen that, mean concentration of NO_2 less than 1 ppbv can have significant impact on the R_{GF} values. Recent studies by *Miller et al* [2017] also found significant dependence of R_{GF} on low NO_x (< 0.5 ppbv). This was attributed to relatively faster production of CHOCHO than HCHO in the early stage of isoprene productions under low NO_x conditions. This also supports our findings of R_{GF} becoming less sensitive to NO_2 in Pantnagar and Chiba.

References

Akagi, S. K., R. J. Yokelson, C. Wiedinmyer, M. J. Alvarado, J. S. Reid, T. Karl, J. D. Crounse, and P. O. Wennberg (2011), Emission factors for open and domestic biomass burning for use in atmospheric models, *Atmos. Chem. and Phys*, ***11*****(9)**, 4039-4072.

Alvarado, L. M. A., A. Richter, M. Vrekoussis, F. Wittrock, A. Hilboll, S. F. Schreier, and J. P. Burrows (2014), An improved glyoxal retrieval from OMI measurements, *Atmos. Meas. Tech*, **7(12)**, 4133.

Anderson, L. G., J. A. Lanning, R. Barrell, J. Miyagishima, R. H. Jones, and P. Wolfe (1996), Sources and sinks of formaldehyde and acetaldehyde: An analysis of Denver's ambient concentration data, *Atmos. Environ.*, *30*(12), 2113-2123, doi:10.1016/1352-2310(95)00175-1.

Andreae, M. O., and P. Merlet (2001), Emission of trace gases and aerosols from biomass burning, *Global Biogeochem. Cy.*, **15(4)**, 955-966.

Arlander, D., D. Brüning, U. Schmidt, and D. Ehhalt (1995), The tropospheric distribution of formaldehyde during TROPOZ II, *J. Atmos. Chem.*, **22(3)**, 251-269.

Bhardwaj, P., M. Naja, R. Kumar, and H.C. Chandola (2016). Seasonal, interannual and long-term variabilities in biomass burning activity over South *Asia. Environ. Sci. Pollut. Res.,* **23**, 4397-4410.

Choi, W., I. Faloona, N. Bouvier-Brown, M. McKay, A. Goldstein, J. Mao, W. Brune, B. LaFranchi, R. Cohen, and G. Wolfe (2010), Observations of elevated formaldehyde over a forest canopy suggest missing sources from rapid oxidation of arboreal hydrocarbons, *Atmos. Chem. Phys.*, **10(18)**, 8761-8781.

DiGangi, J., S. Henry, A. Kammrath, E. Boyle, L. Kaser, R. Schnitzhofer, M. Graus, A. Turnipseed, J. Park, and R. Weber, R.S.Hornbrook, C.A. Cantrell, R.L. Maudlin III, S. Kim, Y. Nakashima, G.M. Wolfe, Y. Kajii, E. Apel, A.H. Goldstein, A. Guenther, T. Karl, A. Hansel, and F.N. Keutsch (2012), Observations of glyoxal and formaldehyde as metrics for the anthropogenic impact on rural photochemistry, *Atmos. Chem. Phys.*, **12(20)**, 9529-9543.

De Smedt, I., T. Stavrakou, F. Hendrick, T. Danckaert, T. Vlemmix, G. Pinardi, N. Theys, C. Lerot, C. Gielen, and C. Vigouroux (2015), Diurnal, seasonal and long-term variations of global formaldehyde columns inferred from combined OMI and GOME-2 observations, *Atmos. Chem. and Phys.* , *15*(8), doi:10.5194 / acp - 15 - 12519 - 2015.

Ervens, B. and Volkamer, R.: Glyoxal processing by aerosol multiphase chemistry: towards a kinetic modeling framework of secondary organic aerosol formation in aqueous particles, *Atmos. Chem. Phys*., **10**, 8219-8244.

Fu, T.-M., D. J. Jacob, P. I. Palmer, K. Chance, Y. X. Wang, B. Barletta, D. R. Blake, J. C. Stanton, and M. 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.*, **112**, D06312

Fu, T., D. Jacob, F. Wittrock, J. P. Burrows, M. Vrekoussis, and D. Henze (2008), Global budgets of atmospheric glyoxal and methylglyoxal, and implications for formation of secondary organic aerosols, *J. Geophys. Res.,* **113**, D15303

Galloway, M. M., Chhabra, P. S., Chan, A. W. H., Surratt, J. D., Flagan, R. C., Seinfeld, J. H., and Keutsch, F. N.: Glyoxal uptake on ammonium sulphate seed aerosol: reaction products and reversibility of uptake under dark and irradiated conditions, *Atmos. Chem. Phys*., **9,** 3331-3345.

Geiger, H., J. Kleffmann, and P. Wiesen (2002), Smog chamber studies on the influence of diesel exhaust on photosmog formation, *Atmos. Environ.*, **36(11)**, 1737-1747.

Ghude, S. D., S. Fadnavis, G. Beig, S. D. Polade, and R. J. van der A (2008), Detection of surface emission hot spots, trends, and seasonal cycle from satellite-retrieved NO_2 over India, *J. Geophys. Res.*, **113(D20)**.

Ghude, S. D., S. H. Kulkarni, C. Jena, G. G. Pfister, G. Beig, S. Fadnavis, and R. J. van der A (2013), Application of satellite observations for identifying regions of dominant sources of nitrogen oxides over the Indian Subcontinent, *J. Geophys. Res.*, **118(2)**, 1075-1089.

Grosjean, E., D. Grosjean, M. P. Fraser, and G. R. Cass (1996), Air quality model evaluation data for organics. 2. C1− C14 carbonyls in Los Angeles air, *Environ. Sci. Technol.*, **30(9)**, 2687-2703.

Holzinger, R., C. Warneke, A. Hansel, A. Jordan, W. Lindinger, D. H. Scharffe, G. Schade, and P. J. Crutzen (1999), Biomass burning as a source of formaldehyde, acetaldehyde, methanol, acetone, acetonitrile, and hydrogen cyanide, *Geophys. Res. Lett.*, **26(8)**, 1161-1164.

Hoque, H. M. S., H. Irie, and A. Damiani (2018a), First MAX-DOAS obseravtions of formaldehyde and glyoxal in Phimai, Thailand, *J. Geophys. Res.*

Hoque, H.M.S., H. Irie, A. Damiani, P. Rawat, and M. Naja (2018b), First simultaneous observations of formaldehyde and glyoxal by MAX-DOAS in the Indo-Gangetic Plain region, *SOLA*, doi:10.2151/sola.2018-028.

Houweling, S., F. Dentener, and J. Lelieveld (1998), The impact of nonmethane hydrocarbon compounds on tropospheric photochemistry, *J. Geophys. Res.*, **103(D9)**, 10673-10696.

Irie, H., Y. Kanaya, H. Akimoto, H. Iwabuchi, A. Shimizu, and K. Aoki (2008a), First retrieval of tropospheric aerosol profiles using MAX-DOAS and comparison with lidar and sky radiometer measurements, *Atmos. Chem. Phys.*, **8(2)**, 341-350.

Irie, H., Y. Kanaya, H. Akimoto, H. Tanimoto, Z. Wang, J. Gleason, and E. Bucsela (2008b), Validation of OMI tropospheric NO_2 column data using MAX-DOAS measurements deep inside the North China Plain in June 2006: Mount Tai Experiment 2006, *Atmos. Chem. Phys.*, **8(22)**, 6577-6586.

Irie, H., Y. Kanaya, H. Akimoto, H. Iwabuchi, A. Shimizu, and K. Aoki (2009), Dual-wavelength aerosol vertical profile measurements by MAX-DOAS at Tsukuba, Japan, *Atmos. Chem. Phys.*, **9(8)**, 2741-2749.

Irie, H., H. Takashima, Y. Kanaya, K. Boersma, L. Gast, F. Wittrock, D. Brunner, Y. Zhou, and M. V. Roozendael (2011), Eight-component retrievals from ground-based MAX-DOAS observations, *Atmos. Meas. Tech.*, **4(6)**, 1027-1044.

Irie, H., T. Nakayama, A. Shimizu, A. Yamazaki, T. Nagai, A. Uchiyama, Y. Zaizen, S. Kagamitani, and Y. Matsumi (2015), Evaluation of MAX-DOAS aerosol retrievals by coincident observations using CRDS, lidar, and sky radiometer in Tsukuba, Japan, *Atmos. Meas. Tech.*, *8*(7), 2775-2788.

Iwabuchi, H. (2006), Efficient Monte Carlo methods for radiative transfer modeling, *J. Atmos. Sci.*, **63(9)**, 2324-2339.

Jenkin, M. E., S. M. Saunders, and M. J. Pilling (1997), The tropospheric degradation of volatile organic compounds: a protocol for mechanism development, *Atmos. Environ.*, ***31*****(1),** 81-104.

Justice, C. O., L. Giglio, S. Korontzi, J. Owens, J. T. Morisette, D. P. Roy, J. Descloitres, S. Alleaume, F. Petitcolin, and Y. Kaufman (2002), The MODIS fire products, *Remote Sens. Environ.*, **83(1)**, 244-262.

Kaiser, J., G.M. Wolfe, K.E. Min, S.S.Brown , C.C. Miller, D.J. Jacob, J.A. deGouw, M. Graus, T.F. Hanisco, J. Holloway, J. Peischl, I.B. Pollack, T.B. Ryerson, C. Warkene, R.A. Washenfelder, and F.N. Keutsch (2015), Reassessing the ratio of glyoxal to formaldehyde as an indicator of hydrocarbon precursor speciation, *Atmos. Chem. and Phys.*, **15(13)**, 7571-7583.

Kansal, A (2009), Sources and reactivity of NMHCs and VOCs in the atmosphere: a review, *J. Hazard. Mater.*, **166**, *17-26.*

Kroll, J. H., N. L. Ng, S. M. Murphy, V. Varutbangkul, R. C. Flagan, and J. H. Seinfeld (2005), Chamber studies of secondary organic aerosol growth by reactive uptake of simple carbonyl compounds, *J. Geophys. Res.*, **110,** D23207

Kumar, R., M. Naja, S. K. Satheesh, N. Ojha, H. Joshi, T. Sarangi, P. Pant, U. C. Dumka, P. Hegde, and S. Venkataramani (2011), Influences of the springtime northern Indian biomass burning over the central Himalayas, *J. Geophys. Res.*, **116(D19)**.

Kurucz, R., I. Furenlid, J. Brault, and L. Testerman (1984), Solar Flux Atlas from 296 to 1300 nm, *Natl. Sol. Obs.*, *Sunspot, New Mexico*, 240.

Kyle Zarzana, kyle.zarzana@noaa.gov.

Lee, D., I. Köhler, E. Grobler, F. Rohrer, R. Sausen, L. Gallardo-Klenner, J. Olivier, F. Dentener, and A. Bouwman (1997), Estimations of global NO, emissions and their uncertainties, *Atmos. Environ.*, **31(12)**, 1735-1749.

Li, X., F.Rohrer, T.Brauers, A. Hofzumahaus, K. Lu, M. Shao, Y. H. Zhang, and A. Wahner, Modelling of HCHO and CHOCHO at a semi-rural site in southern China during the PRIDE-PRD2006 campaign (2014), *Atmos. Chem. Phys.*, **14**, 12291-12305

Liakakou, E., M. Vrekoussis, B. Bonsang, C. Donousis, M. Kanakidou, and N. Mihalopoulos (2007), Isoprene above the Eastern Mediterranean: Seasonal variation and contribution to the oxidation capacity of the atmosphere, *Atmos. Environ.*, **41(5)**, 1002-1010.

MacDonald, S. M., H. Oetjen, A. S. Mahajan, L. K. Whalley, P. M. Edwards, D. E. Heard, C. E. Jones, and J. C. Plane (2012), DOAS measurements of formaldehyde and glyoxal above a south-east Asian tropical rainforest, *Atmos. Chem. and Phys*, **12(13)**, 5949-5962.

Mahajan, A. S., I. De Smedt, M. S. Biswas, S. D. Ghude, S. Fadnavis, C. Roy, and M. van Roozendael (2015), Inter-annual variations in satellite observations of nitrogen dioxide and formaldehyde over India, *Atmos. Environ.*, **116**, 194-201.

Miller, C., G. Gonzalez Abad, H. Wang, X. Liu, T. Kurosu, D. J. Jacob, and K. V. Chance (2014), Glyoxal retrieval from the ozone monitoring instrument, *Atmos. Meas. Tech.* doi:10.5194/amt-7-3891-2014

Miller, C. C., D. J. Jacob, E. A. Marais, K. Yu, K. R. Travis, P. S. Kim, J. A. Fisher, L. Zhu, G. M. Wolfe, and T. F. Hanisco et al (2017), Glyoxal yield from isoprene oxidation and relation to formaldehyde: chemical mechanism, constraints from SENEX aircraft observations, and interpretation of OMI satellite data, *Atmospheric Chemistry and Physics*, *17*(14), 8725-8738, doi:10.5194/acp-17-8725-2017.

Monks, P. S. (2005), Gas-phase radical chemistry in the troposphere, in *Chem. Soc. Rev.*, 376-395.

Munger, J. W., D. J. Jacob, B. C. Daube, L. Horowitz, W. Keene, and B. Heikes (1995), Formaldehyde, glyoxal, and methylglyoxal in air and cloudwater at a rural mountain site in central Virginia, *J. Geophys. Res.*, **100(D5)**, 9325-9333.

Myriokefalitakis, S., M. Vrekoussis, K. Tsigaridis, F. Wittrock, A. Richter, C. Brühl, R. Volkamer, J.P. Burrows, and M. Kanakidou (2008), The influence of natural and anthropogenic secondary sources on the glyoxal global distribution, *Atmos. Chem. Phys.*, **8(16)**, 4965-4981.

Ojha,N., M. Naja, K. P. Singh, T. Sarangi, R. Kumar, S. Lal, M. G. Lawrence, T. M. Butler and H. C. Chandola (2012), *J. Geophys. Res.*,**117 (D20)**

Ortega, I., T. Koenig, R. Sinreich, D. Thomson, and R. Volkamer (2015), The CU 2-D-MAX-DOAS instrument–Part 1: Retrieval of 3-D distributions of NO 2 and azimuth-dependent OVOC ratios, *Atmos. Meas. Tech.*, **8(6)**, 2371-2395.

Pang, X., Y. Mu, Y. Zhang, X. Lee, and J. Yuan (2009), Contribution of isoprene to formaldehyde and ozone formation based on its oxidation products measurement in Beijing, China, *Atmos. Environ.*, *43*(13), 2142-2147, doi:10.1016/j.atmosenv.2009.01.022.

Park, J., H. Lee, J. Kim, J. Herman, W. Kim, H. Hong, W. Choi, J. Yang, and D. Kim (2018), Retrieval Accuracy of HCHO Vertical Column Density from Ground-Based Direct-Sun Measurement and First HCHO Column Measurement Using Pandora, *Remote Sensing*, *10*(2), 173, doi:10.3390/rs10020173.

Platt, U. (1994), Differential optical absorption spectroscopy (DOAS), in *Chemical Analysis Series*, edited, pp. 27-84, Wiley & Sons. Inc.

Poisson, N., M. Kanakidou, and P. J. Crutzen (2000), Impact of non-methane hydrocarbons on tropospheric chemistry and the oxidizing power of the global troposphere: 3-dimensional modelling results, *J. Atmos. Chem.*, *36*(2), 157-230, doi:10.1023/A:1006300616544.

Possanzini, M., G. Tagliacozzo, and A. Cecinato (2007), Ambient levels and sources of lower carbonyls at Montelibretti, Rome (Italy), *Water Air Soil Poll.* , **183(1-4)**, 447-454.

Rappenglück, B., P. K. Dasgupta, M. Leuchner, Q. Li, and W. Luke (2010), Formaldehyde and its relation to CO, PAN, and SO 2 in the Houston-Galveston airshed, *Atmos. Chem. Phys.*, *10*(5), 2413-2424, doi:10.5194/acp-10-2413-2010.

Rodgers, C. D. (2000), *Inverse methods for atmospheric sounding: theory and practice*, World scientific Singapore.

Seco, R., J. Peñuelas, and L. Filella (2007), Short-chain oxygenated VOCs: Emission and uptake by plants and atmospheric sources, sinks, and concentrations, *Atmos. Environ.*, **41(12)**, 2477-2499.

Sugimoto, N., A. Shimizu, T. Nishizawa, I. Matsui, Y. Jin, P. Khatri, H. Irie, T. Takamura, K. Aoki, and B. Thana (2015), Aerosol characteristics in Phimai, Thailand determined by continuous observation with a polarization sensitive Mie–Raman lidar and a sky radiometer, *Environ. Res. Lett.* , **10(6)**, 065003.
Stockwell, C. E., P. R. Veres, J. Williams, and R. J. Yokelson (2015), Characterization of biomass burning emissions from cooking fires, peat, crop residue, and other fuels with high-resolution proton-transfer-reaction time-of-flight mass spectrometry, *Atmos. Chem. and Phys.*, **15(2),** 845-865.

Tanner, R. L., and Z. Meng (1984), Seasonal variations in ambient atmospheric levels of formaldehyde and acetaldehyde, *Environ. Sci. Technol.*, *18*(9), 723-726, doi:10.1021/es00127a017.

Venkataraman, C., G. Habib, D. Kadamba, M. Shrivastava, J. F. Leon, B. Crouzille, O. Boucher, and D. G. Streets (2006), Emissions from open biomass burning in India: Integrating the inventory approach with high-resolution Moderate Resolution Imaging Spectroradiometer (MODIS) active-fire and land cover data, *Global Biogeochem. Cycles*, **20(2)**.

Volkamer, R., J. L. Jimenez, F. San Martini, K. Dzepina, Q. Zhang, D. Salcedo, L. T. Molina, D. R. Worsnop, and M. J. Molina (2006), Secondary organic aerosol formation from anthropogenic air pollution: Rapid and higher than expected, *Geophys. Res. Lett.*, **33**, L17811.

Vrekoussis, M., F. Wittrock, A. Richter, and J. P. Burrows (2010), GOME-2 observations of oxygenated VOCs: what can we learn from the ratio glyoxal to formaldehyde on a global scale?, *Atmos. Chem. Phys.* , **10(21)**, 10145-10160.

Wagner, T., J. P.Burrows, T. Deutschmann, B. Dix, C. v. Friedeburg, U. Frieß, F. Hendrick, K.-P. Heue, H. Irie, and H. Iwabuchi (2007), Comparison of box-air-mass-factors and radiances for Multiple-Axis Differential Optical Absorption Spectroscopy (MAX-DOAS) geometries calculated from different UV/visible radiative transfer models, *Atmos. Chem. Phys.*, **7(7)**, 1809-1833.

Yokelson, R. J., J. G. Goode, D. E. Ward, R. A. Susott, R. E. Babbitt, D. D. Wade, I. Bertschi, D. W. Griffith, and W. M. Hao (1999), Emissions of formaldehyde, acetic acid, methanol, and other trace gases from biomass fires in North Carolina measured by airborne Fourier transform infrared spectroscopy, in *J. Geophys. Res.*, **104(D23)**, 30109-30125.

Zarzana, K. J., K. Min, R. A. Washenfelder, J. Kaiser, M. Krawiec-Thayer, J. Peischl, J. A. Neuman, J. B. Nowak, N. L. Wagner, W. P. Dubè, J.M. Clair, G.M. Wolfe, T.H. Hanisco, F.N. Keutsch, T.B. Ryerson, and S.S. Brown (2017), Emissions of Glyoxal and Other Carbonyl Compounds from Agricultural Biomass Burning Plumes Sampled by Aircraft, *Environ. Sci. Technol.*, **51(20)**, 11761-11770.

Chapter 5: Comparison with satellite measurements and model

In this chapter, ground-based formaldehyde (HCHO) and glyoxal (CHOCHO) measurements are compared with satellite observations. The purpose of the comparison is not to validate the satellite observations, but to compare the general seasonal and diurnal features of HCHO and CHOCHO observed from the ground-based and satellite observations. Only the HCHO measurements in Phimai are also compared with the model. The results presented in this chapter are preliminary results and require more investigation prior to drawing definite conclusions. The results have not been published in any peer-reviewed journal.

5.1 Satellite Instruments

The information provided in this section are based on *Levelt et al.* [2006], *De Smedt et al.* [2015], and *Callies et al.* [2000]

5.1.1 Ozone Monitoring Instrument (OMI)

Aura is a Sun-synchronous polar orbit satellite crossing the Equator around 13:30 local time (LT) in ascending mode. It was launched in July 2004 and is the third major component of the NASA earth observing system after Terra and Aqua, launched in 1999 and 2002, respectively. The Ozone Monitoring Instrument (OMI) is a nadir-viewing imaging spectrometer measuring the backscattered sunlight by the Earth's atmosphere's and surface over the wavelength from 270 to 500 nm with a spectral resolution of ~ 0.5 nm. A scrambler is used to depolarize the light entering the telescope and then split into UV (270 – 380 nm) and VIS channel (350 – 500 nm). The field of view of the telescope is 114° which corresponds to the swath width of 2600 km, enabling nearly daily global coverage. During the nominal global operation mode, the OMI ground pixel size varies from 13×24 km^2 at nadir to 28×150 km^2 at the edge of the swath [*DeSmedt et al* 2015].

5.1.2 Global Ozone Monitoring Experiment 2 (GOME-2)

The global ozone monitoring experiment 2A (GOME-2A) is a nadir-viewing scanning spectrometer aboard the MetOp-A satellite launched in October 2006, in a Sun-synchronous polar orbit with Equator-crossing time of 09:30 LT in descending node. A similar instrument named GOME-2B is aboard MetOp-B satellite launched in September 2012, with Equator-crossing time of 09:30 LT in the descending node. The earlier version of GOME-2 was GOME which flew on the ERS-2 satellite. GOME-2 has four main optical channels which cover the spectral range between 240 and 790 nm with spectral resolution between 0.26 and 0.51 nm. Two polarizing components at 30 broad-band channels covering the full spectral range are measured using polarization measurement devices. Using a diffuser plate, the direct sun spectrum is measured once per day. GOME-2 scan has a default swath width of 1920 km, which provides a global coverage with 1.5 – 3 days at the equator. The nominal ground pixel size of GOME-2 is 80×40 km^2.

5.2 Retrieval Products

There are different products of HCHO and CHOCHO retrieved by various research groups. For the comparison with GOME-2 HCHO columns, the product provided by the German Aerospace Center [*Hassinen et al* 2016] was used. For the comparison with OMI HCHO columns over Phimai, the products provided by the BIRA group [*De Smelt et al.* 2015] were used. The OMI HCHO products provided by the NASA group [*Gonzalez Abad et al.* 2015] were used for the comparison of HCHO columns over Pantnagar. The OMI CHOCHO column products were obtained from Christopher Lerot of BIRA group (personal communication).

5.3 Comparison of HCHO vertical columns in Phimai

Monthly averaged HCHO vertical column densities (VCDs) retrieved from GOME-2, OMI and MAX-DOAS measurements in Phimai are shown in Fig. 5.1 The days in common respectively between GOME-2 and morning MAX-DOAS data and OMI and afternoon MAX-DOA data were used to calculate the monthly averages. Both GOME-2 and OMI data have been averaged within 80km around the Phimai station. The selection of coincident criterion can vary according to locations and conditions. For example, *Wang et al* [2017] and *De Smedt et al* [2015] used a coincidence criterion of <50 km and within 100 km, respectively, for validation of satellite HCHO product using MAX-DOAS measurements. The numbers of correlative days were 476 for OMI and 188 for GOME-2. The error bars for the satellite measurements and MAX-DOAS measurements represent the 2 sigma. The first conclusion that can be drawn from Fig. 5.1 is that the seasonal variation in HCHO has been well produced by both satellite measurements. The enhancements during the dry seasons and the peaks during March were clearly observed in the satellite measurements. The correlations of respectively 0.89 between OMI and MAX-DOAS, and 0.82 between GOME-2 and MAX-DOAS were found. Although the seasonal variations are in good agreement, satellite measurements underestimate HCHO columns by 20% compared to the MAX-DOAS observation during the dry season. In contrast, an overestimation of the satellite columns by 20% compared to the MAX-DOAS observations are found during the wet seasons. Both *De Smedt et al* [2015] and *Wang et al* [2017] reported underestimation of satellite HCHO columns compared to MAX-DOAS observations during the summer time maximum. *De Smedt et al* [2015] and *Wang et al* [2017] used different coincidence criterion for the comparison, but found similar results (i.e., underestimation of satellite HCHO during summertime maximum). So, the coincidence criterion used in this comparison is expected to have little impact on the results. The probable reasons of this underestimation are different vertical sensitivity of satellite and MAX-DOAS measurements, impact of a priori columns on satellite retrieved vertical columns, uncertainties in MAX-DOAS profiles, and effect of coarse resolution on the global chemistry transport model on the modelled profiles used for satellite retrievals [*De Smedt et al* 2015]. Because the Phimai site is affected by biomass burning during the dry season, MAX-DOAS measurements can potentially be strongly affected by local sources. This can also be a potential reason for the difference observed between satellite and MAX-DOAS measurements during the dry season.

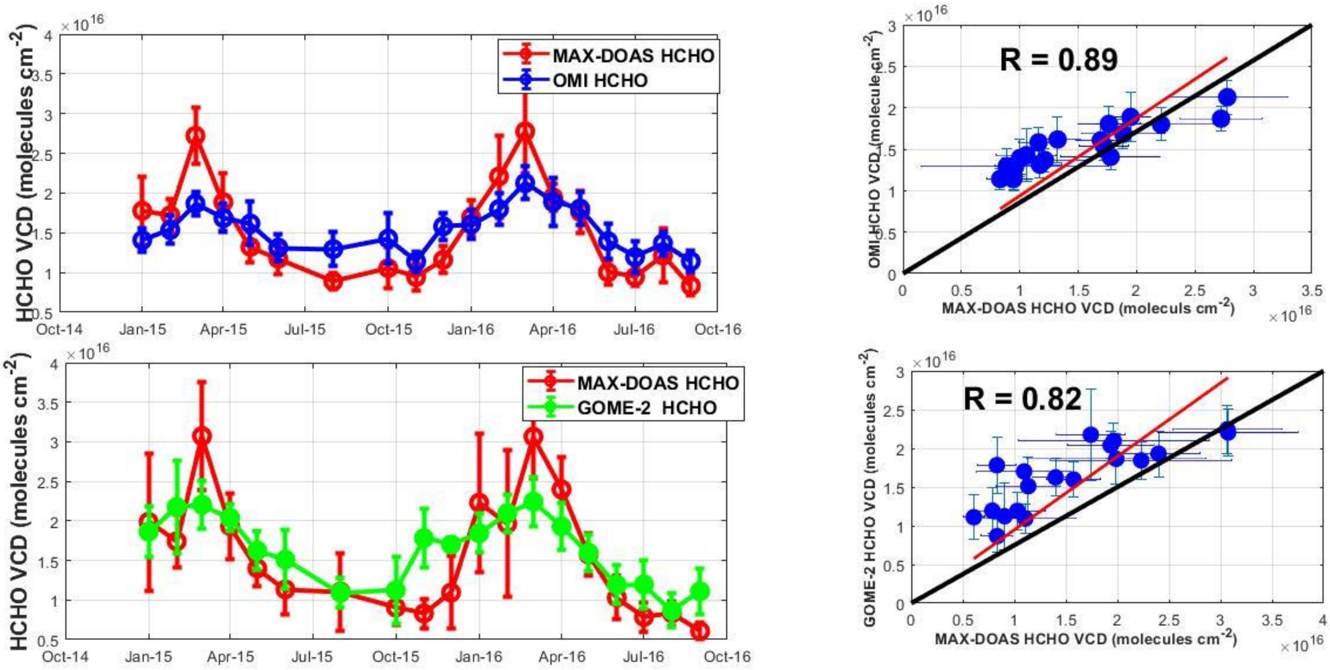


Figure 5.1 Comparisons among MAX-DOAS, GOME-2, and OMI HCHO VCDs over Phimai. The respective correlation coefficient is shown beside each comparison.

Table 5.1 Seasonal mean HCHO VCDs ((molecules cm^{-2}) calculated from MAX-DOAS, GOME-2, and OMI observations in Phimai.

Season	MAX-DOAS VCD (Morning)	GOME-2 VCD	MAX-DOAS VCD (afternoon)	OMI
Dry season 2015	2.18×10^{16}	2.07×10^{16}	$2..02\times10^{16}$	1.62×10^{16}
Dry season 2016	2.41×10^{16}	2.02×10^{16}	2.15×10^{16}	1.85×10^{16}
Wet season 2015	1.12×10^{16}	1.30×10^{16}	1.03×10^{16}	1.30×10^{16}
Wet season 2016	8.13×10^{15}	1.09×10^{16}	9.97×10^{15}	1.27×10^{16}

Because of the morning and afternoon overpass time of GOME-2 and OMI, respectively, the diurnal variation in HCHO columns can also be characterized [*De Smedt et al* 2015, *Wang et al* 2017]. Figure 5.2 shows the diurnal variations observed from MAX-DOAS and satellite measurements during the dry and wet seasons. HCHO concentrations did not show any significant diurnal variation during both seasons, as inferred from the MAX-DOAS observations. Similarly, the OMI and GOME-2 values during the dry and wet seasons were also almost similar, indicating almost a flat diurnal variability in HCHO columns in Phimai. However, overestimation of the satellite columns during the wet season were observed in the diurnal variations.

As seen from above, the enhancement in the HCHO columns during the dry season agrees well between the satellite and MAX-DOAS measurements. However, the peaks during the dry seasons were sharper in the OMI observations compared to the GOME-2A observations. *Surl et al.* [2018] stated a hypothesis that owing to the early afternoon overpass time, OMI can better capture the role of biogenic emissions and biomass burning in the observed column variations of HCHO. Because biogenic emissions and to a lesser extent biomass burning peaks in the early afternoon, whereas anthropogenic VOCs related to vehicular emissions peak in the early morning and late afternoon. *Surl et al.* [2018] tested this hypothesis analyzing HCHO column variations over India. This could be the potential cause of the sharper peak observed in the OMI HCHO columns in Phimai during the dry season compared to the GOME-2A observations. However, the diurnal variations of HCHO columns inferred from both MAX-DOAS and satellite measurements didn't show any significant differences between the morning and afternoon values. Thus, the hypothesis by *Surl et al.* [2018] may not be valid for this case. The difference seen in the peaks in OMI and GOME-2 might be related to the number of coincidence days with the MAX-DOAS measurements. Addressing this difference requires further investigation.

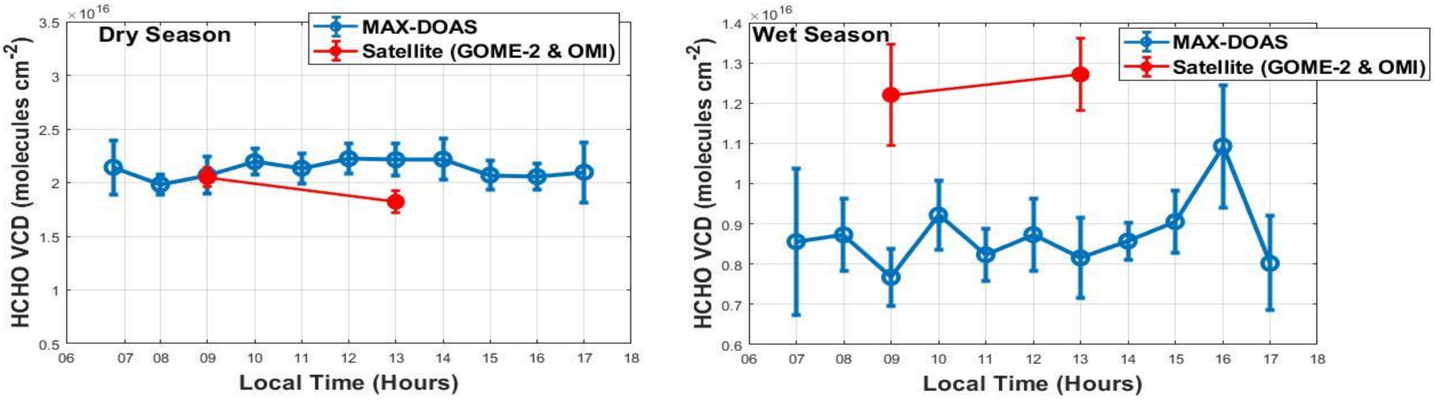


Figure 5.2 Diurnal variations of HCHO VCDs during the dry and wet season in Phimai inferred from the MAX-DOAS and satellite observations. The days in common in all the three datasets were used to calculate the hourly VCDs.

5.4 Comparison of CHOCHO vertical columns in Phimai

Figure 5.3 shows the monthly averaged CHOCHO columns retrieved from OMI and MAX-DOAS measurements in Phimai. The days in common between OMI and afternoon MAX-DOAS data (1200-1500 LT) were used to calculate the monthly means. The number of correlative days among MAX-DOAS and OMI were 467. The OMI CHOCHO columns were averaged within 100 km around the Phimai station. The error bars for both OMI and MAX-DOAS represent the 2-sigma standard deviation. A general overestimation of the satellite CHOCHO column by 0 to 60% was observed. Higher CHOCHO concentrations during the dry season were also observed in the satellite measurements. Low correlation (R = 0.30) was observed between OMI and MAX-DOAS observations. Considering the measurements in 2016, when MAX-DOAS observations were available for the whole dry and wet seasons, good agreement (*R* = 0.71) was found between OMI and MAX-DOAS observations. The satellite observations show higher concentrations before the strongest impact of biomass during March and April, as observed in the MAX-DOAS observations and other studies. This also probably led to the lower correlation between OMI and MAX-DOAS observations. To check whether the difference in the timing of the CHOCHO column enhancements is due to the coincidence criterion used, smaller (<80 km) and higher (>100) coincidence criteria were tested, and the similar results were obtained. Whether such difference in the seasonal behavior is observed in other places or not, similar comparison between OMI and MAX-DOAS CHOCHO columns was performed for Chiba and Pantnagar (Figure 5.4).

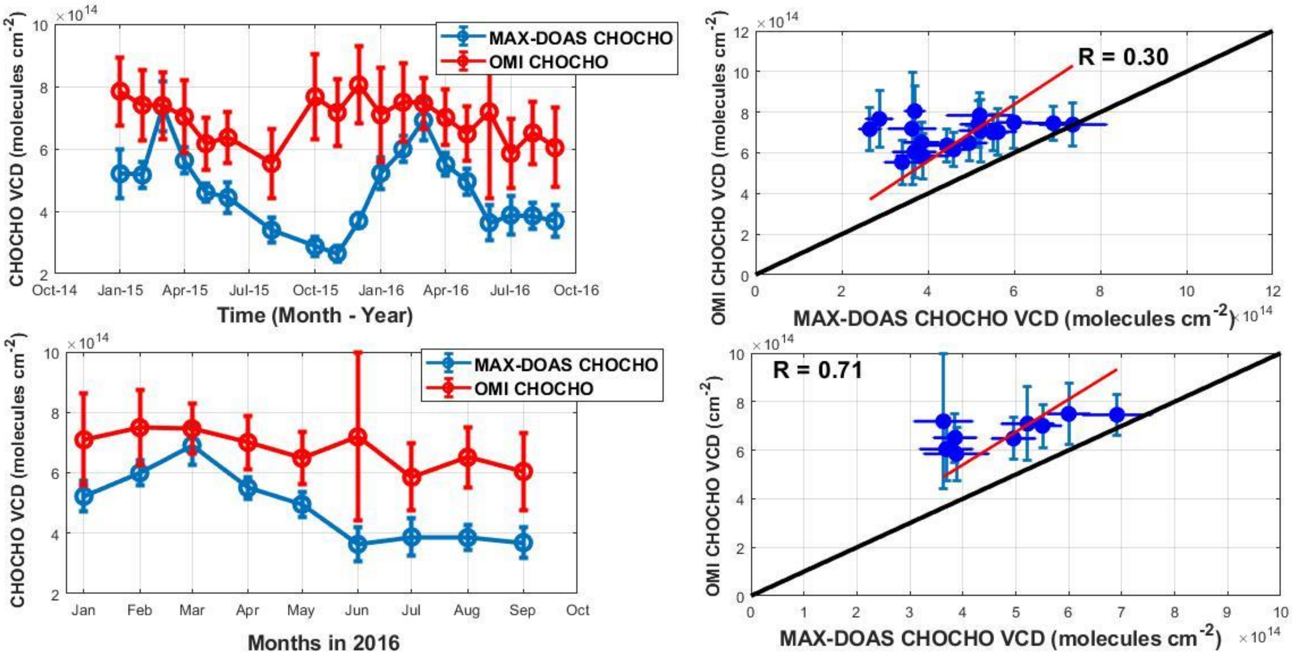


Figure 5.3 Comparison among MAX-DOAS and OMI CHOCHO VCDs over Phimai. The respective correlation coefficient is shown beside each comparison. Observation for both 2015 and 2016 are used for the comparison in the upper panel. The lower panel shows the comparison only for 2016.

For both the sites, the coincidence criterion for the satellite measurement was < 100km and only afternoon MAX-DOAS data (1200 -1500 LT) were used. Figure 5.4 shows the monthly averaged CHOCHO columns retrieved from OMI and MAX-DOAS observations in Chiba and Pantnagar. Similar to Phimai, an overestimation of the satellite CHOCHO columns compared to the MAX-DOAS measurements was observed. Higher correlations between OMI and MAX-DOAS in Pantnagar and Chiba were found compared to Phimai. No significant seasonal pattern in Chiba was observed. The seasonal variations in Pantanagar are in good agreement, with the enhancements during the spring and autumn. Similar to the comparison in Phimai, the difference in the occurrence of the seasonal peaks in the MAX-DOAS and satellite CHOCHO columns were observed. The reason of such difference in the seasonal maxima is not clear. Comparing GOME-2 CHOCHO columns and model simulations, *Lerot et al* [2010] reported overestimation and underestimation of satellite columns. This was attributed to the requirement of improving the representation of glyoxal sources and sinks in the chemistry transport models. Overall, a good agreement in the seasonal variation between the satellite and MAX-DOAS CHOCHO columns in all three SKYNET sites was observed.

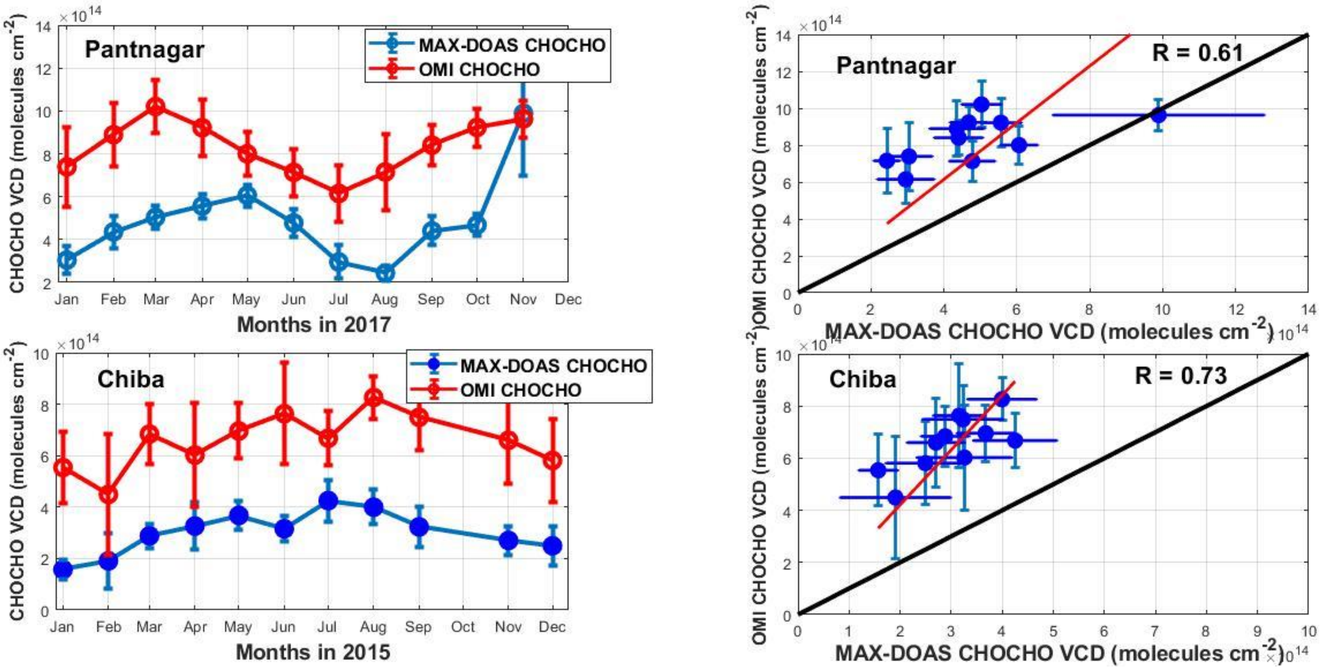


Figure 5.4 Comparison among MAX-DOAS and OMI CHOCHO VCDs over Pantnagar (upper panel) and Chiba (lower panel). The respective correlation coefficient is shown beside each comparison. Observations for 2015 and 2017 in Chiba and Pantnagar, respectively were used for the comparison.

5.5 Comparison of HCHO vertical columns in Pantnagar

Figure 5.5 shows the monthly averaged HCHO columns retrieved from GOME-2, OMI and MAX-DOAS measurements in Pantnagar. Similar to the comparison in Phimai, the days in common respectively between GOME-2 and morning MAX-DOAS data, and OMI and afternoon MAX-DOAS data were used to calculate the monthly averages. GOME- 2 and OMI data were averaged within 50 and 80 km around the Pantnagar station, respectively. The number of correlative days was 196 for OMI and 56 for GOME-2. Two different coincidence criteria has been used for OMI and GOME-2. The coincident criteria have been optimized in terms of the number of comparisons and spatiotemporal inhomogeneity, according to our additional tests for different coincidence criteria. The error bars for the satellite measurements and MAX-DOAS measurements represent the 2 sigma. The seasonal variation in HCHO has been well produced by both satellite measurements, with the enhancements during spring and autumn. The correlation of 0.69 between OMI and MAX-DOAS, and 0.71 between GOME-2 and MAX-DOAS were found. MAX-DOAS and GOME-2 observations agree well during all seasons. However, MAX-DOAS overestimates HCHO column by ~ 30% compared to OMI, although both observations show almost similar values during the summer monsoon. Overall, the seasonal variation in HCHO in Pantnagar inferred from MAX-DOAS, GOME-2, and OMI shows good agreement.

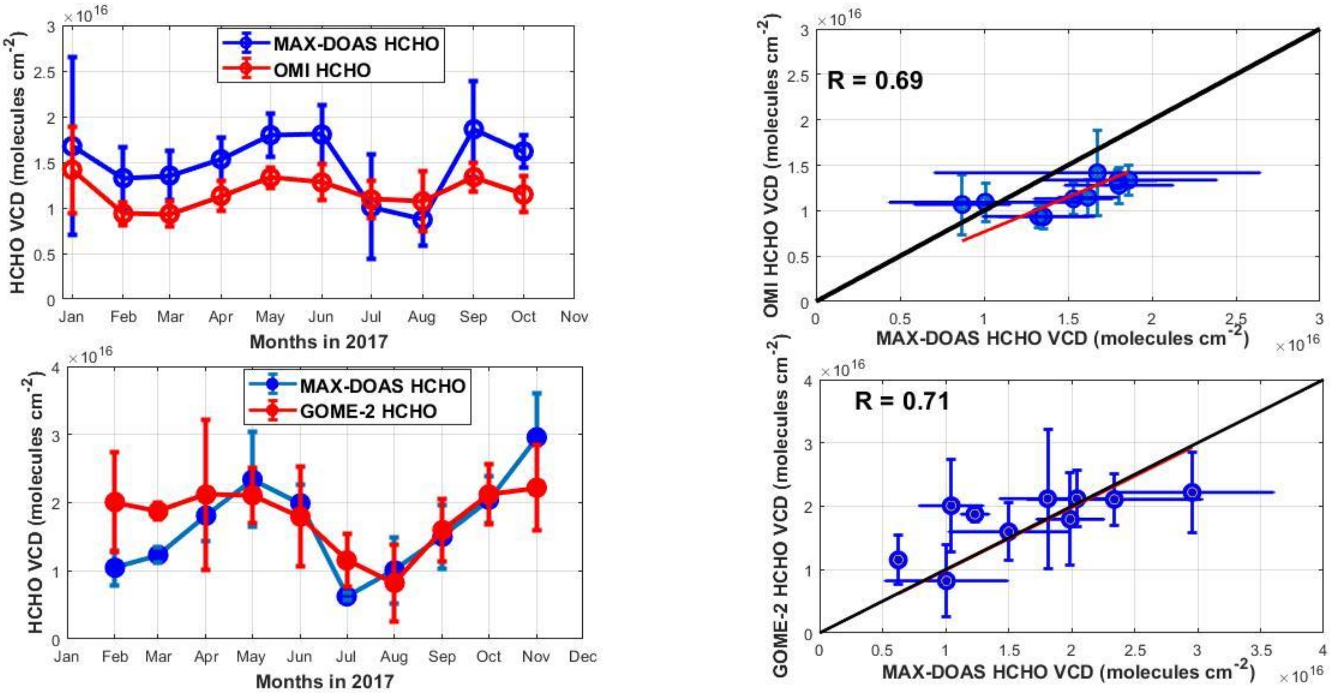


Figure 5.5 Comparison among MAX-DOAS, GOME-2, and OMI HCHO VCDs over Pantnagar. The respective correlation coefficient is shown beside each comparison.

Table 5.2 Seasonal mean HCHO VCDs (molecules cm^{-2}) calculated from MAX-DOAS, GOME-2, and OMI observations in Pantnagar.

Season	MAX-DOAS VCD (Morning)	GOME-2 VCD	MAX-DOAS VCD (afternoon)	OMI
Winter (JF)	----------	-----------	1.50×10^{16}	1.17×10^{16}
Spring (MAM)	1.80×10^{16}	2.03×10^{16}	1.56×10^{16}	1.13×10^{16}
Summer Monsoon (JJA)	1.20×10^{16}	1.25×10^{16}	1.23×10^{16}	1.14×10^{16}
Autumn (SO)	1.77×10^{16}	1.85×10^{16}	1.73×10^{16}	1.24×10^{16}

Similar to the comparison in Phimai, utilizing the morning and afternoon overpass time of GOME-2 and OMI, respectively, the diurnal variations in HCHO columns were investigated. Figure 5.6 shows the diurnal variations observed from MAX-DOAS and satellite measurements during spring, summer monsoon, and autumn in Pantnagar. GOME-2 measurements were averaged within 80 km centered at Pantnagar, slightly different from coincidence criterion used for the seasonal variation. Similar to the seasonal variation, good agreement is also observed between GOME-2 and MAX-DOAS diurnal values during all seasons. Except during summer monsoon, underestimation of OMI values was observed. The satellite observations show higher morning and lower afternoon values during spring and autumn.

Surl et al. [2018] also analyzed the HCHO observations by OMI over India. They also reported discrepancies between model and satellite observations over the IGP region. However, *Surl et al.* [2018] also found better agreement between model and OMI observations during the monsoon period (similar to our analysis). This has been attributed to the relationship between leaf phenology and HCHO variation. During the monsoon period, vegetation takes advantage of the decreased temperature and higher precipitation rates to regulate leaf flushing. Thus, a strong reduction is observed in the leaf area index (LAI) and in the HCHO concentration. The leaf flushing was first explained by *Barkley et al.* [2009], to explain the variability of HCHO columns observed over the amazon.

Although GOME-2 shows good agreement with MAX-DOAS observations, yet based on few correlative days, a definitive conclusion can't be drawn. One of the probable reason of discrepancies observed among ground-based measurements, satellite observations, and models over the IGP region could be the representation of the source of HCHO. As seen in the discussion in chapter 4, HCHO was not strongly correlated with temperature and vegetation indexes in Pantnagar. There was contribution from biogenic, pyrogenic, and anthropogenic emissions to the VOC variability. As the relative contribution of the sources to the observed VOC variability is not very clear, this is likely to lead to the differences observed among different platforms. However, despite the differences in the absolute values, the overall seasonal variations agree well among different observation platforms.

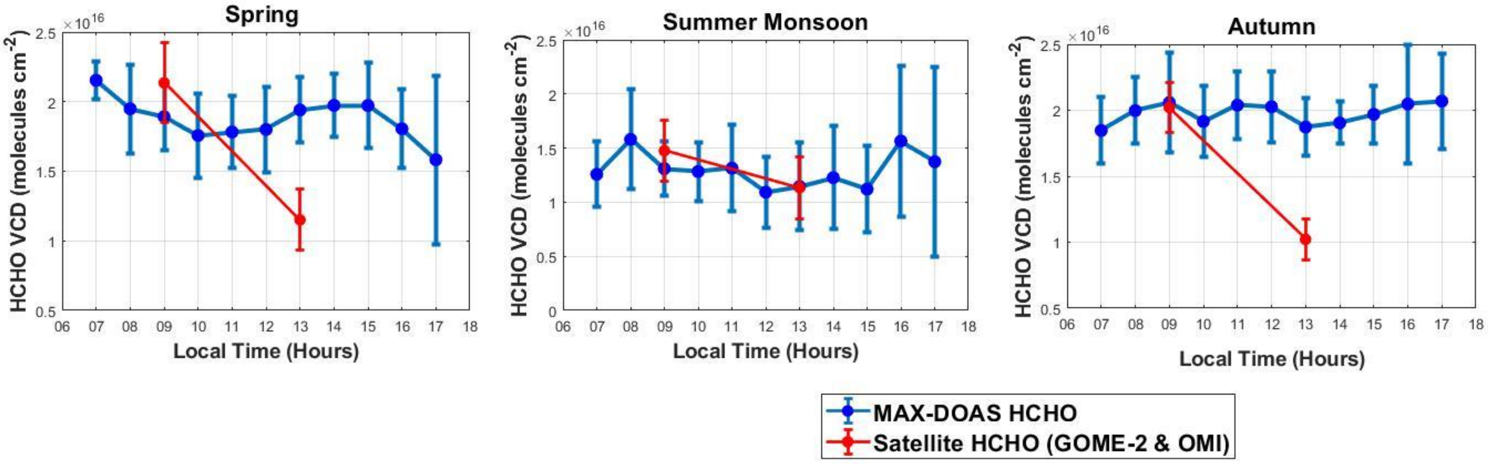


Figure 5.6. Diurnal variations of HCHO VCDs during spring, summer monsoon, and autumn in Pantnagar inferred from the MAX-DOAS and satellite observations. The days in common in all the three dataset were used to calculate the hourly VCDs.

5.6 Comparison of R_{GF} estimated from satellite and MAX-DOAS observations

The R_{GF} response to different VOC emission sources inferred from satellite measurements and field studies is still under discussion. Using satellite measurements, *Vrekoussiss et al.* [2010] found higher R_{GF} for strong biogenic emission scenarios, and lower R_{GF} (< 0.030) for regions with dominant anthropogenic VOC sources. In contrast, *Miller et al.* [2014] reported higher R_{GF} values for regions associated with biomass burning and monoterpene emissions. Figure 5.7 shows the comparison of R_{GF} values estimated from OMI and MAX-DOAS observations in Phimai. For the comparison, the satellite HCHO and CHOCHO columns were averaged with 80 km centered at the Phimai station and only the afternoon MAX-DOAS (1200-1500 LT) data were used. The comparison has been shown only for 2016, because in 2016 MAX-DOAS measurements covers the whole dry and wet seasons. As seen from Fig. 5.7, satellite R_{GF} values are higher compared to the MAX-DOAS observations. The satellite mean R_{GF} (0.050) is ~30% higher than the mean R_{GF} estimated from MAX-DOAS observations. *Kaiser et al.* [2015] reported similar higher satellite averaged R_{GF} compared to flight-based observations. The higher R_{GF} values can be potentially due to the overall overestimation of the satellite CHOCHO columns compared to the MAX-DOAS columns. Moreover, such difference can also be related to the altitude dependence of R_{GF} due to the difference in the vertical sensitivity of the satellite and MAX-DOAS observations. *Baidar et al* [2013] reported higher R_{GF} at an altitude ~2.5 km compared to the boundary layer.

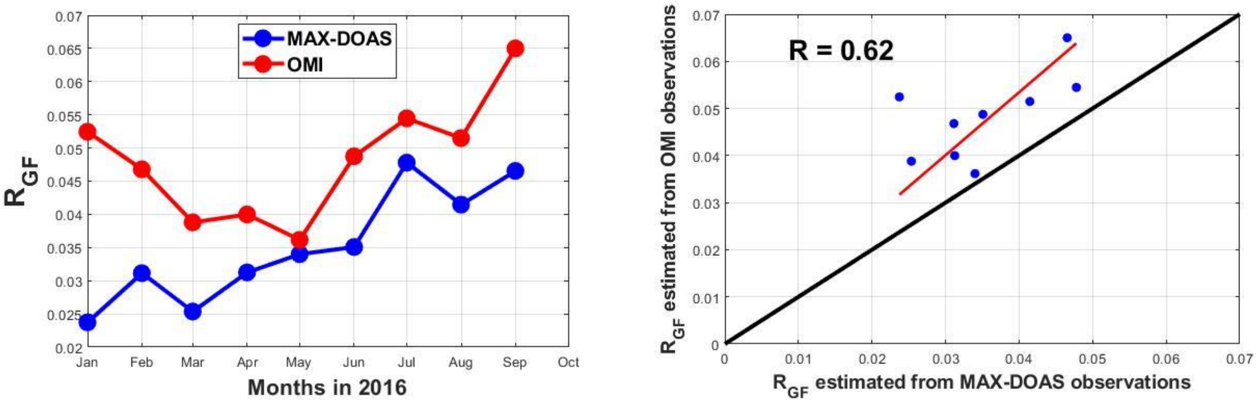


Figure 5.7 R_{GF} estimated for different months from MAX-DOAS and OMI observations in Phimai.

However, a good positive correlation ($R = 0.62$) between the R_{GF} estimated from satellite and MAX-DOAS observations were observed. Although the satellite retrievals have relative higher uncertainties [*Miller et al* 2014, *Gonzalez Abad et al* 2015, *DeSmedt et al* 2015], the seasonal variations are well reflected. Thus, the seasonal change in satellite mean R_{GF} values are also observed (Table 5.3). Similar to the ground-based observations, higher R_{GF} (> 0.05) values during the wet season were also seen in the satellite observations, when VOC emissions are predominantly biogenic. There were significant differences in the R_{GF} values estimated for the dry and wet seasons from the satellite observations. Thus, despite the difference in the absolute mean R_{GF} values, similar response of R_{GF} to different VOC emission sources were observed in Phimai.

Table 5.3 R_{GF} estimated from MAX-DOAS and OMI observations in Phimai. Only the afternoon (1200 – 1500 LT) MAX-DOAS data and days in common between OMI and MAX-DOAS observations were used for calculating the R_{GF}.

Seasons	**R_{GF} (MAX-DOAS)**	**R_{GF} (OMI)**
Over all	0.032±0.003	0.050±0.004
Dry Season 2015	0.029±0.002	0.049±0.008
Wet Season 2015	0.034±0.005	0.050±0.002
Dry Season 2016	0.028±0.004	0.045±0.006
Wet Season 2016	0.042±0.005	0.055±0.007

To check the consistency of the R_{GF} response between the satellite and ground-based observations, R_{GF} values estimated for the Chiba site were also compared with satellite R_{GF} values. The coincidence criterion for the satellite HCHO and CHOCHO columns were similar to that of Phimai.

The R_{GF} values from the satellite observations over Pantnagar were not estimated in this study. Because although the OMI CHOCHO data provided by the BIRA group were available for 2017, but the OMI HCHO product of the BIRA group was not available for the similar period. As the BIRA group products were available during the MAX-DOAS observation of HCHO and CHOCHO in Chiba, the R_{GF} values estimated from the satellite measurements were estimated for Chiba which is plotted in Fig. 5.8

The satellite R_{GF} values in Chiba were mostly around 0.05 or less. The mean R_{GF} for the observation in 2015 estimated from satellite and MAX-DOAS observation were 0.053 and 0.031, respectively. From the satellite measurements over Phimai and Chiba, it has been observed that under the influence of biomass burning and anthropogenic emissions, the R_{GF} tends to be ~0.05 or less. Similarly, ground-based R_{GF} response under anthropogenic and pyrogenic influences were consistent for all the three sites, with $R_{GF} < 0.04$.

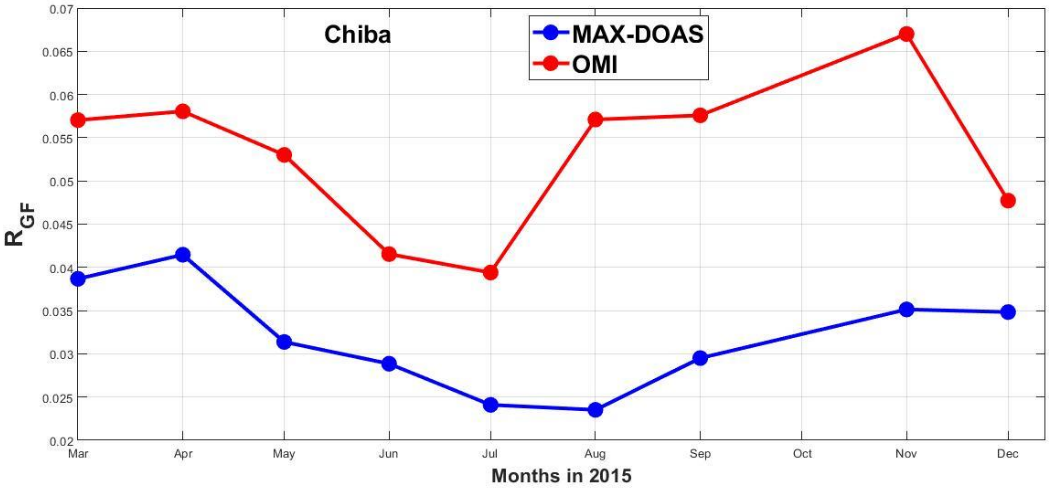


Figure 5.8 R_{GF} estimated for different months from MAX-DOAS and OMI observations in Chiba.

5.7 Model Simulations

The MAX-DOAS HCHO observations at Phimai were compared with the results from the global three-dimensional chemical model named chemical atmospheric general circulation model for study of atmospheric environment and radiative forcing (CHASER) [*Sudo et al.* 2002]. At first we provide a short description of the model and then present the comparison result. The description of the model is mostly based on the research article by *Sudo et al.* [2002].

The primary objective of the CHASER model is to study the global distribution and budgets of ozone and its precursors. CHASER can also be employed for assessing the global impacts of changes in the atmospheric composition on climate change. CHASER is based on the atmospheric global circulation model (AGCM) developed by the Center for Climate System Research (CCSR), the University of Tokyo and National Institute for Environmental Studies (NIES). The basic features of the CCSR/NIES have been described by *Numaguti et al.* [1993,1995]. The climatology of the meteorological fields is reproduced by the AGCM. CHASER uses the climatological data of sea surface temperature (SST) as an input to the AGCM. To simulate a specific time period, in addition to the SST, analyzed data of wind velocities, temperature, and specific humidity from the European Center for Medium Range Weather Forecasts are used as constrains. Because, for a certain period, only climatological output from the model might be difficult to validate with observational data.

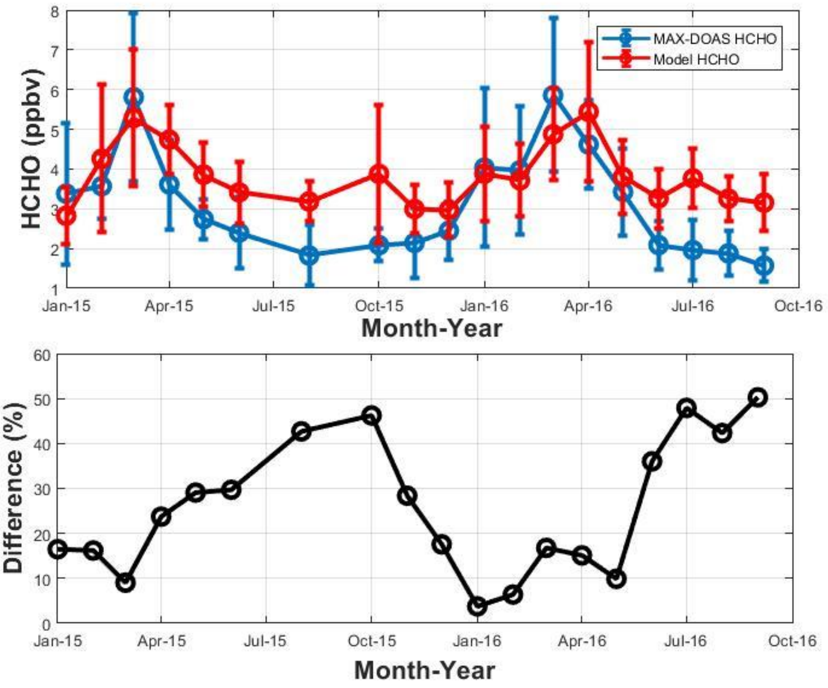
The dynamical processes in CHASER, such as tracer transport, vertical distribution, vertical diffusion, surface emissions, and deposition are simulated in the flow of the AGCM. The chemical component of CHASER uses the AGCM variables such as temperature, pressure, humidity, etc. to calculate the chemical transformations. The radiation component calculates the radiative transfer and photolysis rates bases on the concentrations of chemical species calculated in the chemistry component. The physical and dynamical components of the CHASER model are evaluated at a time resolution of 30 min and a chemical time step of 10 min. In this comparison, the horizontal model resolution was 1.4°×1.4° with around 32 vertical layers from the surface up to ~40 km (3 hPa). The vertical layers are based on the σ coordinate system. The centers of the 32 layers are approximately at 995, 980, 950, 900, 830, 745, 657, 576, 501, 436, 380, 331, 288, 250, 2018, 190, 165, 144, 125, 109, 95, 82, 72, 62, 54, 47, 40, 34, 27, 19, 11, and 3 hPa, which correspond to a vertical resolution of 1 km in the free troposphere and the much of the lower stratosphere.

At present, CHASER calculates the concentrations of 44 chemical species, which include 34 tracers (transported) and 16 non tracers (radical species and members of family tracers). The concentration is calculated from the surface to an altitude of 20 km. More detailed information about the CHASER model can be obtained from http://atmos.ccsr.u-tokyo.ac.jp/~kengo/chaser .

5.7.1 Seasonal Variation

Figure 5.9 shows the comparison between the MAX-DOAS HCHO and HCHO concentration simulated by the CHASER model for Phimai. As mentioned above, the model resolution of the simulation was 1.4°×1.4° and the altitude was roughly 1 km. The model output was calculated for every two hours. For the comparison with the model, MAX-DOAS HCHO data for the 0-2 km were used and the observations were averaged for every two hours. The days in common between the CHASER simulation and MAX-DOAS observations were used to calculate the monthly mean HCHO. The model simulation for the day time (0700-1700 LT) were only used for the monthly means. The seasonal variation in HCHO has been well produced by the model. The enhancements during the dry seasons and the peaks during March were clearly observed in model simulation. The correlations of 0.79 between CHASER and MAX-DOAS were observed. Although the seasonal variations are in good agreement, differences in the HCHO concentration up to 50% between the CHASER model and MAX-DOAS were observed. During the wet seasons, CHASER overestimated HCHO concentrations by up to 50%

compared to the MAX-DOAS observation. The differences in the HCHO concentrations during the dry seasons were much lower, mostly less than 20%, which yields a better agreement between the model and observations during the dry season. Table 5.4 lists the mean HCHO concentration calculated from the model simulation and MAX-DOAS observations for different seasons. Although the overall mean HCHO concentrations are almost similar (~3 ppbv), larger differences in the CHASER and MAX-DOAS HCHO concentrations are seen during the wet season. The better agreement during the dry season and overestimation by the model during the wet season were similar to those observed during the comparison of the MAX-DOAS HCHO columns with satellite observations. The different horizontal resolution of the model (~100 km) and MAX-DOAS (~10 km) could be a potential reason of the difference in the observed HCHO concentration. However, with better horizontal resolution, satellite observations overestimated HCHO columns during the wet seasons. Thus, the horizontal resolution is expected not to have a significant impact on the observed differences between the model and MAX-DOAS HCHO concentrations. The HCHO sources in Phimai during the dry and wet seasons were mainly biomass burning and biogenic emissions, respectively.



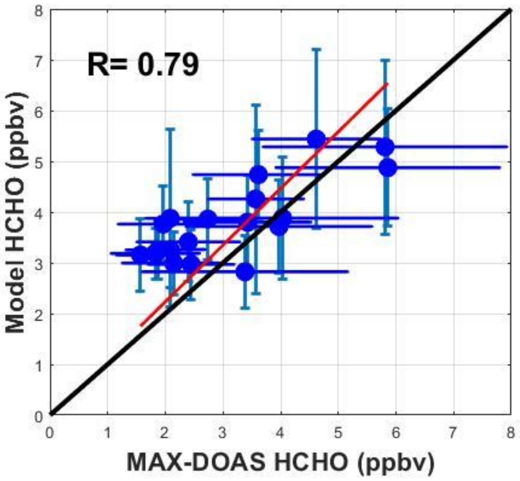


Figure 5.9 (upper panel) Monthly time series of HCHO concentration in Phimai, inferred from MAX-DOAS observations and CHASER simulation. The days in common between the observation and model were used to calculate the monthly means. (lower panel) Difference between the monthly mean HCHO concentration of the observations and model.

Table 5.4 HCHO concentrations in Phimai during the dry and wet seasons calculated from the MAX-DOAS observations and model simulations.

Seasons	**HCHO (ppbv) MAX-DOAS**	**HCHO (ppbv) Model**
Over all	3.12 ± 1.29	3.81 ± 0.78
Dry Season 2015	4.09 ± 1.14	4.27 ±1.05
Wet Season 2015	2.11 ± 0.40	3.29 ± 0.16
Dry Season 2016	4.61 ± 0.87	4.42 ± 0.82
Wet Season 2016	1.87 ± 0.22	3.35 ± 0.27

The smaller difference in the HCHO concentrations during the dry season are likely indicative of well presentation of HCHO from pyrogenic sources in the model. The higher differences during the wet season might be indicative for better representation of the biogenic emissions in the model. However, no definitive conclusion can be drawn from a single comparison. Overall, the seasonal variation of HCHO in Phimai agrees well between the model and observations.

5.7.2 Diurnal variation

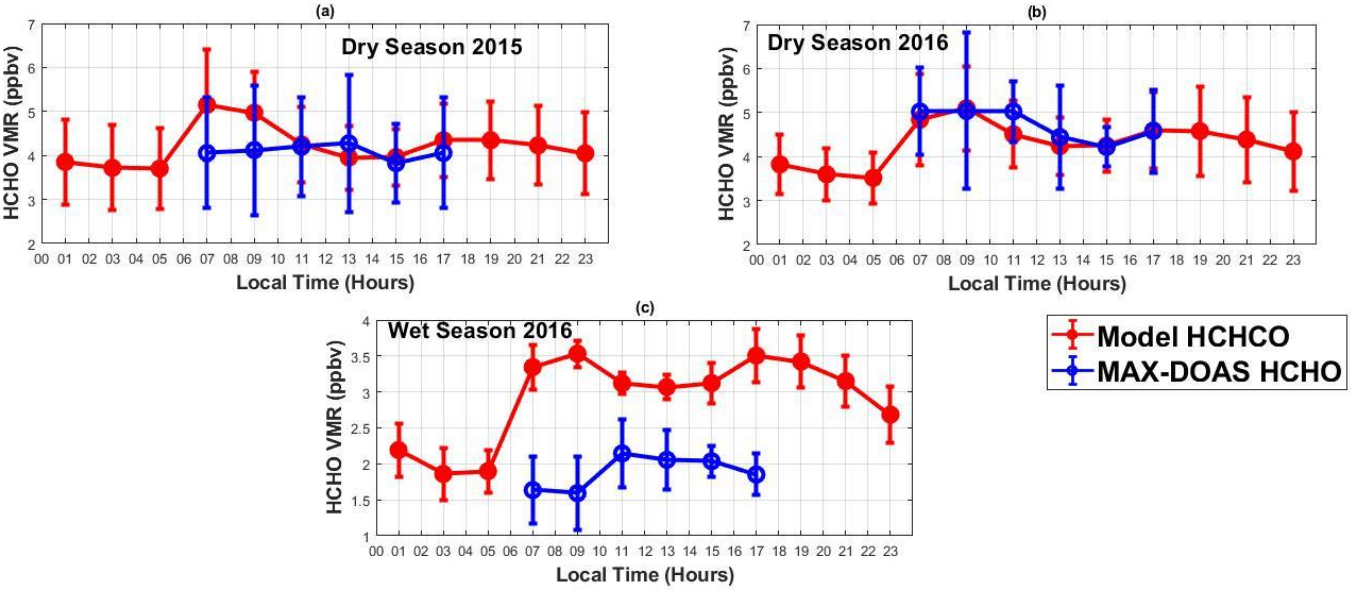


Figure 5.10 Diurnal variation of HCHO seen in the MAX-DOAS observations and model simulation in Phimai during (a) dry season 2015, (b) dry season 2016, and (c) wet season 2016.

Figure 5.10 shows the diurnal variation of HCHO inferred from the CHASER model and MAX-DOAS observations. Similar to the seasonal variation, MAX-DOAS observations were averaged for every two hours to match the model output. Days in common between the model and observations were used to calculate the hourly mean HCHO concentration. Only the day time (0700 – 1700 LT) HCHO concentrations were compared with the model.

Similar to the seasonal variation, the hourly variation in HCHO during the dry season inferred from the observations and model were in good agreement. The model showed the daytime (0700 – 1700 LT) HCHO concentration varied between 4 – 5 ppbv during both dry seasons. MAX-DOAS observations showed almost a flat diurnal variation ~4 ppbv during the dry season in 2015 and slightly higher (~5 ppbv) in the morning (0700 – 0800 LT) values during the dry season in 2016. The diurnal variation during the 2016 dry seasons shows better agreement with the model compared to the 2015 dry season. However, during both dry seasons, the model and observations showed no significant differences in the morning and afternoon HCHO concentrations. The comparison among MAX-DOAS HCHO columns and satellite columns also showed similar results, where GOME-2 (morning overpass) and OMI (afternoon over pass) showed almost the similar HCHO columns.

During the wet season, MAX-DOAS showed lower HCHO concentrations compared to the model. This was similar to the overestimation by the model during the wet season seen in the seasonal variation. The model shows the daytime (0700 – 1700 LT) HCHO during the wet season varied between 3 – 3.5 ppbv whereas MAX-DOAS observation shows the concentration between 1.5 – 2 ppbv. However, despite the differences in the absolute values, the diurnal variations in HCHO inferred from the model and observations show good agreement during both seasons.

References

Baidar, S., H. Oetjen, S. Coburn, B. Dix, I. Ortega, R. Sinreich, and R. Volkamer (2013), The CU Airborne MAX-DOAS instrument: Vertical profiling of aerosol extinction and trace gases, *Atmos. Meas. Tech.*, *6*(3), 719-739, doi:10.5194/amt-6-719-2013.

Barkley, M. P., Palmer, P. I., De Smedt, I., Karl, T., Guenther, A., and Van Roozendael, M.: Regulated large-scale annual shutdown of Amazonian isoprene emissions, *Geophys. Res. Lett.*, 36, L04803, https://doi.org/10.1029/2008GL036843, 2009

Callies, J., E. Corpaccioli, M. Eisinger, A. Hahne, and A. Lefebvre (2000), GOME-2-Metop's second-generation sensor for operational ozone monitoring, *ESA Bull.* , *102*, 28-36.

De Smedt, I., T. Stavrakou, F. Hendrick, T. Danckaert, T. Vlemmix, G. Pinardi, N. Theys, C. Lerot, C. Gielen, and C. Vigouroux (2015), Diurnal, seasonal and long-term variations of global formaldehyde columns inferred from combined OMI and GOME-2 observations, *Atmos. Chem. and Phys.* , *15*(8), doi:10.5194 / acp - 15 - 12519 - 2015.

González Abad, G., X. Liu, K. Chance, H. Wang, T. P. Kurosu, and R. Suleiman (2015), Updated Smithsonian Astrophysical Observatory Ozone Monitoring Instrument (SAO OMI) formaldehyde retrieval, *Atmos. Meas. Tech.*, *8*(1), 19-32, doi:10.5194/amt-8-19-2015.

Hassinen, S., D. Balis, H. Bauer, M. Begoin, A. Delcloo, K. Eleftheratos, S. Gimeno Garcia, J. Granville, M. Grossi, and N. Hao (2016), Overview of the O3M SAF GOME-2 operational atmospheric composition and UV radiation data products and data availability, *Atmos. Meas. Tech*, *9*(2), 383.

Kaiser, J., G. M. Wolfe, K. E. Min, S. S.Brown , C. C. Miller, D. J. Jacob, J. A. deGouw, M. Graus, T. F. Hanisco, J. Holloway, J. Peischl, I. B. Pollack, T. B. Ryerson, C. Warkene, R. A. Washenfelder, and F. N. Keutsch (2015), Reassessing the ratio of glyoxal to formaldehyde as an indicator of hydrocarbon precursor speciation, *Atmos. Chem. and Phys.*, *15*(13), 7571-7583, doi:10.5194/acp-15-7571-2015

Lerot, C., T. Stavrakou, I. DeSmedt, J.-F. Müller, and M. Van Roozendael (2010), Glyoxal vertical columns from GOME-2 backscattered light measurements and comparisons with a global model, *Atmos. Chem. and Phys.*, *10*(24), 12059-12072, doi:10.5194/acp-10-12059-2010.

Levelt, P. F., G. H. J. van den Oord, M. R. Dobber, A. Malkki, H. Visser, J. de Vries, P. Stammes, J. O. V. Lundell, and H. Saari (2006), The ozone monitoring instrument, *IEEE Trans. Geosci. Res. Sens.*, *44*(5), 1093-1101, doi:10.1109/TGRS.2006.872333.

Miller, C., G. Gonzalez Abad, H. Wang, X. Liu, T. Kurosu, D. J. Jacob, and K. V. Chance (2014), Glyoxal retrieval from the ozone monitoring instrument, *Atmos. Meas. Tech.* , doi:10.5194/amt-7-3891-2014.

Numaguti, A., Dynamics and energy balance of the hadley circulation and the tropical precipitation zones: Significance of the distribution of evaporation, *J. Atmos. Sci*., 50, 1874 – 1887, 1993.

Numaguti, A., M. Takahashi, T. Nakajima, and A. Sumi, Development of an atmospheric general circulation model, in Reports of a New Program for Creative Basic Research Studies, Studies of Global Environment Change With Special Reference to Asia and Pacific Regions, Rep. I-3, pp. 1 – 27, CCSR, Tokyo, 1995.

Sudo, K., M. Takahashi, J. Kurokawa, and H. Akimoto (2002), CHASER: A global chemical model of the troposphere 1. Model description, *J. Geophys. Res.*, *107*(D17), ACH 7-1-ACH 7-20, doi:10.1029/2001JD001113.

Surl, L., P. I. Palmer, and G. G. Abad (2018), Which processes drive observed variations of HCHO columns over India?, *Atmos. Chem. and Phys.*, *18*(7), 4549, doi:10.5194 / acp 18-4549-2018.

Vrekoussis, M., F. Wittrock, A. Richter, and J. P. Burrows (2010), GOME-2 observations of oxygenated VOCs: what can we learn from the ratio glyoxal to formaldehyde on a global scale?, *Atmos. Chem. Phys.* , *10*(21), 10145-10160, doi:10.5194/acp-10-10145-2010.

Wang, Y., et al. (2017), Validation of OMI, GOME-2A and GOME-2B tropospheric NO 2, SO 2 and HCHO products using MAX-DOAS observations from 2011 to 2014 in Wuxi, China: investigation of the effects of priori profiles and aerosols on the satellite products, *Atmos. Chem. and Phys*(8), 5007, doi:10.5194/acp-17-5007-2017.

Concluding Remarks

The first simultaneous observations of HCHO and CHOCHO using MAX-DOAS systems were conducted at Phimai (Thailand) and Pantnagar (India). Utilizing the observations, CHOCHO to HCHO concentration ratio (R_{GF}) at both sites were estimated and the response of R_{GF} to VOC emission sources were assessed. The Phimai site was a rural site and during the dry season (January – April) the site was affected by biomass burning. Corresponding to the biomass burning, enhanced HCHO, CHOCHO, and NO_2 concentrations were observed during the dry seasons. The influence of biomass burning was confirmed using backward trajectories and fire radiative power (FRP) values. Moreover, the leaf area index showed almost no variation throughout the year in Phimai, indicating similar strength of biogenic VOC sources in Phimai. Thus, the enhancement in trace gas concentrations during the dry season were likely due to the influence of biomass burning. The annual mean R_{GF} for Phimai was 0.032 ± 0.005. Lower R_{GF} values were observed during the dry season in correspondence to the influence of biomass burning. The VOC sources were dominantly biogenic during the wet season as higher R_{GF} values were observed. The diurnal variability of R_{GF} was similar to that of CHOCHO with higher values during the daytime and lower in the afternoon. However, the diurnal variability had no significant impact on the R_{GF} response to different VOC emission sources in Phimai. Moreover, the wet removal process during the wet season were also found to have minimal impact on the R_{GF} values. The response of R_{GF} to different VOC emissions sources in Phimai was consistent with few studies based on satellite observations and field measurements. However, discrepancies were found with few reported studies. The probable reasons of such discrepancies were different response of R_{GF} to similar VOC emission sources under different climate conditions, the altitude dependence of R_{GF} and poor understanding of VOC chemistry under low NO_x conditions.

The Pantnagar site is a semi-rural site in the IGP region in India. The site was influenced by emissions from vehicles and the nearby cities. Additionally, twice a year during spring and autumn the site was also influenced by biomass burning. Corresponding to the biomass burning, enhanced HCHO, CHOCHO, and NO_2 concentrations were observed during spring and autumn in Pantnagar. The mean R_{GF} in Pantnagar estimated from the whole observation period was 0.029 ± 0.006. The mean R_{GF} was mostly $< \sim 0.032$ throughout all seasons. The VOC emissions sources in Pantanagar were mostly dominated by anthropogenic emissions and biomass burning. The lower correlation of HCHO and CHOCHO with the vegetation indexes indicated that the biogenic VOC sources were less significant compared to other two sources. Thus the observed R_{GF} values in Pantnagar were likely mostly dominated by anthropogenic emissions and biomass burning. The similar R_{GF} values were also observed in Phimai during the dry season. Thus under the influence of anthropogenic emissions and biomass burning, the R_{GF} response (lower R_{GF} value) was similar at both site despite the different climatology. The NO_2 concentration in Pantnagar (~1.5 ppbv) was higher compared to that of Phimai (< ~1ppbv). The NO_2 concentrations in Pantnagar had no significant impact on the R_{GF} values, whereas in Phimai, NO_2 had a significant impact on the R_{GF}. This was consistent with studies reporting less sensitivity of R_{GF} to higher NO_2 concentration.

The comparison of HCHO vertical column density data from GOME-2 and OMI in Phimai and Pantnagar showed good seasonal agreement. The enhancement of HCHO due to the biomass burning at both sites was clearly seen in the satellite observations. The CHOCHO columns at both sites also showed good agreement with OMI observations. The HCHO concentrations in

Phimai also showed good agreement with the CHASER HCHO concentrations. The seasonal and diurnal variabilities simulated by the model during the dry season matches very well with the observations. However, the overestimation of HCHO during the wet season by the model, as well as by satellite observations requires further detailed investigation.

In conclusion, this study finds that, (1) the R_{GF} values are higher under dominant biogenic VOC emissions and (2) under the influence of anthropogenic VOC emissions and biomass burning the mean R_{GF} tends to be < ~0.04. This study also suggests the comparison of R_{GF} on a monthly or seasonal scale to assess the response of R_{GF} to different VOC emission sources.

Publications and Presentations

Publications

(1) Hoque, H. M. S., H. Irie, and A. Damiani (2018), First MAX-DOAS obseravtions of formaldehyde and glyoxal in Phimai, Thailand, *Journal of Geophysical Research,* 9957-9975, doi.org/10.1029/2018JD028480.

(2) Hoque, H. M. S., H. Irie, A. Damiani, P. Rawat, and M. Naja (2018), First simultaneous observations of formaldehyde and glyoxal by MAX-DOAS in the Indo-Gangetic Plain region, *Scientific Online Letters on the Atmosphere*, doi:10.2151/sola.2018-028.

Selected Presentations

(1) First MAX-DOAS observations of glyoxal and formaldehyde in Phimai, Thailand, Oral Presentration, Japan Society for Atmospheric Chemistry, Hokkaido, 2016.

(2) First MAX-DOAS observation of glyoxal and formaldehyde in Phimai, Thailand, Oral Presentation, JpGU – AGU joint seminar, Chiba, 2017.

(3) First MAX-DOAS observations of formaldehyde and glyoxal in Southeast and South Asia, Oral Presentation, DOAS workshop, Yokohama, 2017.

(4) First MAX-DOAS observations of formaldehyde and glyoxal in Phimai, Thailand, Poster Presentation, AGU fall meeting, New Orleans, 2017.

(5) First MAX-DOAS observations of formaldehyde and glyoxal in Phimai, Thailand, Poster Presentation, ACAM meeting,China, 2017.

(6) First MAX-DOAS observations of formaldehyde and glyoxal in Southeast and South Asia, Oral Presentation, Japan Meteorological Society meeting, Hokkaido, 2017.

(7) First MAX-DOAS observations of formaldehyde and glyoxal in Southeast and South Asia, Oral Presentation, JpGU meeting, Chiba, 2018.

Acknowledgements

I am very glad and grateful for being able to conduct my doctoral studies at the Center for Environmental Remote Sensing (CEReS) in Chiba University. I would like to express my benevolent gratitude to Dr. Hitoshi Irie for giving me the chance to be a part of this wonderful research institute. I am also grateful to him for being my thesis supervisor and a wonderful mentor. I am also thankful to him for giving me a chance to connect with the international DOAS community, which has helped me to acquire a lot of knowledge regarding the scientific application of DOAS. I would like to thank Dr. Alessandro Damiani for his magnificent support especially during writing scientific articles and answering the critical reviewer comments. I would also like to appreciate the technical assistance of Okamoto San, who has also taught me a lot of aspects of UNIX systems. Special thanks to Kumakawa San for helping me in all aspect of my life in Japan.

I am grateful to Dr. Isabelle DeSmedt and Christopher Lerot of the BIRA group for providing us the satellite HCHO and CHOCHO data. I would also like to acknowledge the contribution of Dr. Kengo Sudo for his assistance with the modelling studies.

Lastly, I would like to thank all the members of the Irie Laboratory (past and present members) for the technical and non-technical discussions, parties, fun, and overall for all the sweet memories which I will reminisce all throughout my life.