The First Evaluation of Formaldehyde Column Observations by Pandora Spectrometers during the KORUS-AQ Field Study

Elena Spinei[1,2], Andrew Whitehill[3], Alan Fried[4], Martin Tiefengraber[5,6], Travis N.
5 Knepp[7,8], Scott Herndon[9], Jay R. Herman[1,10], Moritz Müller[5,6], Nader Abuhassan[1,10],
Alexander Cede[1,5], Petter Weibring[4], Dirk Richter[4], James Walega[4], James Crawford[7],
James Szykman[3,7], Lukas Valin[3], David J. Williams[3], Russell Long[3], Robert J. Swap[1],
Youngjae-Lee[11], Nabil Nowak[2], Brett Poche[2].

Correspondence author Elena Spinei Lind: eslind@vt.edu

10 [1] NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA
[2] Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA
[3] US EPA, Research Triangle Park, Durham, North Carolina 27709, USA
[4] Institute of Arctic and Alpine Research (INSTAAR) at the University of Colorado,
Boulder, Colorado 80303, USA
15 [5] LuftBlick, Kreith 39A, 6162 Mutters, Austria
[6] Institue of Atmospheric and Cryospheric Sciences, University of Innsbruck, Innsbruck,
Austria
[7] NASA Langley Research Center, Hampton, Virginia 23681, USA
[8] Science Systems and Applications, Inc., Hampton, VA 23681, USA
20 [9] Aerodyne Research, Inc., Billerica, MA 01821, USA
[10] University of Maryland Baltimore County, Baltimore, MD, USA
[11] Korean National Institute of Environmental Research (NIER), South Korea

40

Abstract

The KORUS-AQ field study conducted during May-June 2016 offered the first opportunity to evaluate direct-sun observations of formaldehyde (HCHO) total column densities with improved Pandora spectrometer instruments. The measurements

5   highlighted in this work were conducted both in the Seoul megacity area at the Olympic Park site (latitude: 37.5232° N, longitude: 127.1260° E, 26 m a.s.l.) and at a nearby rural site downwind of the city at the Mount Taehwa Research Forest site (latitude: 37.3123° N, longitude: 127.3106° E, 160 m a.s.l.). Evaluation of these measurements was made possible by concurrent ground-based in situ observations of HCHO at both sites as well

10   as overflight by the NASA DC-8 research aircraft. The flights provided in situ measurements of HCHO to characterize its vertical distribution in the lower troposphere (0 - 5 km).

Diurnal variation in HCHO total column densities followed the same pattern at both sites, with the minimum daily values typically observed between 6:00-7:00 local

15   time, gradually increasing to a maximum between 13:00 and 17:00 before decreasing into the evening. Pandora vertical column densities were compared with those derived from the DC-8 HCHO in-situ measured profiles augmented with in-situ surface concentrations below the lowest altitude of the DC-8 in proximity to the ground sites. A comparison between 49 column densities measured by Pandora versus aircraft integrated in-situ data

20   showed that Pandora values were larger by 16% (intercept = 0.22 DU, R^2 = 0.68). Pandora HCHO columns were also compared with columns calculated from the surface in-situ measurements over Olympic Park by assuming a well-mixed lower atmosphere up to a ceilometer measured mixed-layer height (MLH) and various assumptions about the small residual HCHO amounts in the free troposphere up to the tropopause. The best

25   comparison (slope = 1.03 ± 0.03, intercept = 0.29 ± 0.02 DU and R^2 of 0.78±0.02) was achieved assuming equal mixing within ceilometer measured MLH combined with an exponential profile shape. These results suggest that diurnal changes in HCHO surface concentrations can be reasonably estimated from the Pandora total column and information on the mixed-layer height. More work is needed to understand the bias in the

30   intercept and the slope relative to columns derived from the in-situ aircraft and surface measurements.

1. Introduction

Formaldehyde (HCHO) is a key constituent in tropospheric chemical cycling. Its
abundance is dominated by secondary formation through the oxidation of methane and
non-methane hydrocarbons. It is also short lived, undergoing photolysis or oxidation by
OH within a few hours under typical daytime conditions. As such, HCHO provides an
important indicator of the integrated oxidation of hydrocarbons that contributes to
tropospheric ozone production in the presence of nitrogen oxides. The degradation of
HCHO can also constitute an important secondary source of HO_x ($HO + HO_2$), serving to
amplify oxidation rates in polluted atmospheres. A more detailed discussion of HCHO
chemistry can be found in Fried et al., (2011) and references therein.

The attributes described above make HCHO an important test species in
evaluating our mechanistic understanding of tropospheric oxidation reactions as well as a
valuable proxy for hydrocarbon emissions. Remote sensing of HCHO promises valuable
insight into the emissions and processes driving tropospheric chemistry. For instance,
satellite measurements of HCHO by the Global Ozone Monitoring Experiment (GOME)
(Fu et al., 2007; Palmer, 2003; Palmer et al., 2006; Shim et al., 2005), SCanning Imaging
Absorption SpectroMeter for Atmospheric CHartographY (SCIAMACHY) (Wittrock et
al., 2006) , and Ozone Monitoring Instrument (OMI) (Marais et al., 2012) have been
used to map the isoprene emissions on a global scale. In combination with remote
sensing of NO_2, satellite observations of HCHO have been explored for their utility in
assessing the factors contributing to ozone pollution by mapping areas of NOx-controlled
versus VOC-controlled ozone formation (Jin et al., 2017; Jin and Holloway, 2015;
Schroeder et al., 2016)

With the promise of both temporal and spatial information for HCHO on the
horizon from a constellation of geostationary satellites (Zoogman et al., 2017), other
possible uses for satellite observations of HCHO are emerging. Recent work by
Schroeder et al., (2016) suggests that column HCHO shows promise as a proxy for
surface ozone. Valin et al., (2016) examine the relationship between column HCHO and
its dependence on OH production and VOC reactivity, demonstrating the importance of
this information to improving satellite-derived emissions estimates for isoprene and other

hydrocarbons. These efforts to further develop and improve the use of future satellite observations elevates the need for ground-based remote sensing to validate satellite measured HCHO columns. Remote sensing Differential Optical Absorption Spectroscopy (DOAS) has been widely used to measure HCHO from ground (Lee et al., 2005; MacDonald et al., 2012; Vlemmix et al., 2015), aircraft (Baidar et al., 2013) and satellite (Bauwens et al., 2016; De Smedt et al., 2015) platforms. In this paper we present HCHO total columns from DOAS measurements of unscattered, direct-sun photons using NASA/GSFC (Goddard Space Flight Center) Pandora instruments and in-situ measurements over two sites during the Korea-United States Air Quality Study (KORUS-AQ) conducted in May-June 2016 in South Korea. These Pandora measurements are then evaluated extensively with both ground-based and airborne in situ observations of HCHO available during the study.

2. KORUS-AQ Measurements

KORUS-AQ fielded a multi-perspective suite of observations including both remote sensing and *in situ* observations of air quality at ground sites across the peninsula and on research aircraft collecting valuable data on conditions aloft. Pandora spectrometers were used to observe total columns of HCHO at five locations, but two sites in particular also included ground-based *in situ* measurements of HCHO and frequent atmospheric profiling overflights by the NASA DC-8 aircraft with an *in situ* measurement of HCHO onboard.

The first site was located in the Seoul megacity at Olympic Park (37.5232° N, 127.1260° E, 26 m a.s.l.), which the DC-8 overflew routinely during the study, visiting the site three times per day at the beginning, middle, and end of many research flights. These overflights were typically at 300 m during a descent over the city that ended below 30 m during a "missed approach" over the runway at Seoul Air Base approximately 8 km to the south.

The second site was at Mount Taehwa (37.3123° N, 127.3106° E, 160 m a.s.l.), a research forest site located approximately 29 km southeast of the Olympic Park site and in the predominant downwind sector of transport for the Seoul megacity plume. Overflight of Mt. Taehwa was routinely performed following a missed approach at Seoul

Air Base and was followed by a spiral ascent to 7.6 km altitude to provide a complete profile of in situ HCHO in the lower atmosphere over Korea.

2.1. Pandora Measurements

5   Pandora instruments are field grade spectroscopic UV-VIS systems (Herman et al. 2009). They are part of the joint National Aeronautics and Space Administration (USA) and European Space Agency sponsored Pandonia Global Network (PGN). The main goal of PGN is to provide consistent ground-based total NO_2, HCHO and O_3 columns for satellite validation.

10   The instrument consists of a small Avantes low stray light spectrometer (280 – 525 nm with 0.6 nm spectral resolution with 5 times oversampling) connected to an optical head by a 400-micron core diameter single strand fiber optic cable. The optical head is attached to a small 2-axis positioner, capable of accurate pointing for tracking the sun's center (±0.2°). A diffuser is included in the optical path to minimize the effect

15   small pointing errors.

      Pandora spectra are automatically collected and submitted to NASA/Luftblick servers for centralized uniform processing by the Blick Software Suite (Cede, 2017). All standard operational Pandora data products are available at http://pandonia.net/data (note, HCHO is not a standard product at the time of this publication).

20   Pandoras deployed during KORUS-AQ were retrofitted with new UV grade fused silica windows with broadband antireflection coating (ARC, 250-700 nm). This modification from the earlier versions of Pandora was necessary to decrease spurious spectral structure in direct sun (DS) spectra. This new ARC window improved NO_2 and O_3 measurements and made HCHO retrieval from Pandora DS measurements possible.

25   Pandoras measure unscattered solar photons in a narrow cone (2.1° field of view (FOV) full width half maximum with a diffuser and 1.6° FOV without the diffuser) at a specific solar azimuth and zenith direction that changes from East in the morning to West in the evening. Figure 1 shows the schematics of DS observation geometry where detected photons travel through the atmosphere in a slant path.
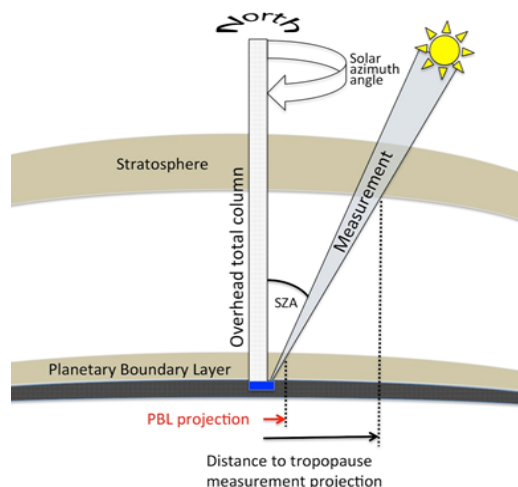
Figure 1. Schematic of direct sun observation geometry.

2.1.1. Pandora HCHO vertical column retrieval

HCHO total vertical column densities are calculated from Pandora measurements of non-scattered sun photons (with visible light blocked by a U340 filter) using the DOAS technique. The analysis consists of the following steps described in detail by Cede, (2017):

A) Correction of the DS collected spectra (L0 to L1 level) for dark current, CCD non-linearity, latency effect, pixel response non uniformity, filter transmission, instrument temperature sensitivity, stray light, wavelength shift, etc.

B) Selection of the reference spectrum, ideally, a Pandora measured reference spectrum with the smallest possible HCHO absorption and highest signal-to-noise ratio. In this study all spectra with low measurement noise collected around local noon (±30 min) during the entire campaign were averaged to create a single reference spectrum.

C) Calculation of HCHO differential slant column densities (ΔSCD) relative to the reference spectrum using DOAS equation (Cede, 2017):

$$\ln F_{0i} - \ln(F_i - P_{OFFSi}) - \tau s_{FIXi} = \sum_{j=1}^{n_{GAS}} \tau s_{ji} (q s_j, T_j) + P_{SMOi} + P_{RESCi} \qquad (1)$$

where i is the index for pixel inside the limits of the fitting window, i=1 to n, the center-wavelength of pixel i is λ_i; F_{0i} is the reference spectrum at pixel i used in the fitting; F_i is the L1 data for pixel i; P_{OFFSi} is the offset polynomial evaluated at pixel i; τs_{FIXi} is the

"known" slant optical depth at pixel i; j is the index for atmospheric absorber, j=1 to n_{GAS}; τs_{ji} is the slant optical depth of absorber $_j$ for slant column qs_j and effective temperature T_j at pixel i; P_{SMOi} is the smooth-part polynomial evaluated at pixel i; P_{RESCi} is the resolution change polynomial evaluated at pixel i. Since the reference spectrum contains an

5   unknown amount of HCHO, retrieved slant columns are the difference between true HCHO slant columns (qs) and slant columns in the reference spectrum (qs_{REF}). We will use SCD for qs_{HCHO} notation in the rest of this paper.

The fitting window used in this study to calculate HCHO columns is 332 -359 nm (P_{SMO} = 4). In addition to HCHO, other gases present in the atmosphere absorb in the

10   selected fitting window: ozone (O_3), nitrogen dioxide (NO_2), oxygen collision complex (O_2O_2), and bromine monoxide (BrO). Their high-resolution molecular absorption cross sections were convolved with the Pandora instrument slit function prior to DOAS fitting and listed in Table 1 (for convolution details see Cede, 2017).

15   Table 1. DOAS fitting parameters used to calculate HCHO ΔSCD from Pandora direct sun measurements

Fitting wavelength window: 332 – 359 nm		
Polynomial order: 4		
Offset and wavelength shift polynomial order: 1		
Species	Temperature [K]	Citation of high resolution cross section
O_3	225	Malicet et al., 1995
NO_2	263	Vandaele et al., 1998
O_2-O_2	262	Hermans et al., 2003
BrO	223	Fleischmann et al., 2004
HCHO	298	Meller and Moortgat, 2000

D) Calculation of the air mass factor for DS observation geometry (AMF$_{DS}$) according to Eq. (2):

$$AMF_{DS} = sec\left\{arcsin\left[\left(\frac{R_{Earth}}{R_{Earth}+h_{eff}}\right) \cdot sin\ (SZA^*)\right]\right\} \qquad (2)$$

Where R_{Earth} is the distance from the center of the Earth to the measurement location, SZA^* is the geometrical solar zenith angle corrected for refraction, and h_{eff} is the effective, profile concentration weighted height of a "typical" background HCHO

25   distribution (4.3 km). Dependence of AMF$_{DS}$ on h_{eff} is very small at SZA < 80°.

E) Estimation of HCHO slant column density in the reference spectrum (SCD$_{REF}$)
using Modified Langley Extrapolation method (MLE). Herman et al., (2009) reported its
application to NO$_2$ measurements. MLE is a statistical method based on the assumption
that during a sufficiently long measurement period, the trace gas of interest (here HCHO)

5   VCDs will reach a certain minimum level at different times of the day (AMF). This
assumption might not hold, especially if the species in question has a systematic diurnal
pattern. Practical implementation of MLE consists in subsetting all the measured ΔSCD
data in AMF bins and performing a linear regression on the lowest 2-percentile of low
noise L1 data. The Intercept from the linear regression represents SCD$_{REF}$ at AMF = 0.

10   Depending on the data filtering for instrumental/atmospheric noise and upper and
lower AMF limits, SCD$_{REF}$ over Olympic Park ranges from 1.035±0.06 DU to 1.22±0.04
DU (Dobson Unit = 2.69 x10^{16} molecules/cm^2); and over Mt. Taehwa from 0.74±0.03
DU to 0.82±0.04 DU. In all cases the quality of the linear fit is very high (R^2 > 0.97).
Higher SCD$_{REF}$ were the result of data limitation to AMF between 1.2 and 4. In this study

15   we estimated SCD$_{REF}$ = 1.035±0.18 DU for Olympic Park and 0.74±0.08 DU at Mt.
Taehwa. We selected a smaller SCD$_{REF}$ from the range since it is more representative of
the "standard" implementation of MLE by the PGN personnel, mainly based on NO$_2$ data
processing experience.

F) Calculation of the HCHO vertical column density (VCD, L2 level Pandora data)

20   VCD = (ΔSCD + SCD$_{REF}$) / AMF   (3)

2.1.2. Pandora HCHO VCD uncertainty budget

The total error in the Pandora direct-sun HCHO VCD (ε_{VCD}) combines errors in
ΔSCD and SCD$_{REF}$ calculation and errors in AMF determination. ε_{VCD} can be estimated by
summing the corresponding errors in quadrature according to Eq. (4) (assuming all the

25   components are independent of each other).

$$\varepsilon_{VCD} = \sqrt{\left(\frac{\varepsilon_{\Delta SCD}}{AMF}\right)^2 + \left(\frac{\varepsilon_{SCD_{REF}}}{AMF}\right)^2 + \left(\frac{\varepsilon_{AMF} \cdot SCD_{REF}}{AMF^2}\right)^2 + \left(\frac{\varepsilon_{AMF} \cdot \Delta SCD}{AMF^2}\right)^2}$$ (4)

A. Uncertainties in ΔSCD are due to (1) statistical errors of the DOAS fitting and (2)
systematic errors in: the laboratory measured molecular absorption cross-sections and

their temperature dependence, wavelength calibration, and the cross correlation between absorption cross-sections of different molecules (choice of fitting wavelength window) *(Platt and Stutz, 2008; Stutz and Platt, 1996)*. In this study statistical errors of the DOAS fitting are calculated by BlickSFA algorithm (Cede, 2017) that accounts for atmospheric

5 and instrumental noise.

Selection of a fitting scenario (332-359 nm, see Table 1) can result in an error of ±10%. This error was determined by performing DOAS fittings using different scenarios (324-359, 336-359, variation of P_{SMO} in Eq. (1), inclusion and exclusion of BrO absorption, and fitting O_3 temperature dependent cross section). All of the scenarios

10 resulted in comparable DOAS fitting residual optical depth rms, but varied by 20% in HCHO ΔSCD. The selected fitting scenario was in the middle of the spread.

Uncertainties due to the laboratory measured high-resolution molecular cross-sections (σ) used in DOAS fitting (after convolution) propagate into the retrieved HCHO SCD. The effect of other gas σ errors depends on cross correlation between different

15 cross sections and instrumental noise in a specific fitting window. Pinardi et al. (2013) reported that for multi-axis DOAS geometry (336.5-359 nm), error due to σ(O_3) selection (Bogumil et al., 2003 vs. Malicet et al., 1995) can result in HCHO ΔSCD error of 13% (ΔSCD); σ(NO_2) selection (Vandaele et al., 1998 vs. Burrows et al., 1998)) - up to 5%, and σ(BrO) selection (Fleischmann et al., 2004 vs. Wilmouth et al., 1999) - about 2%.

20 We adopt Pinardi et al. (2013) estimates for Pandora HCHO total error calculation in this study and will perform more sensitivity studies with Pandora data in future.

Uncertainty due to extraneous spectral structure in DS spectra (even with the new ARC window) is harder to evaluate and will be subject of future studies.

B. Uncertainties in the SCD_{REF} ($\varepsilon_{SCD_{REF}}$) calculated from MLE are driven by the data

25 availability and diurnal changes in HCHO optical depth relative to the reference time optical depth. MLE requires at least 2 weeks of measurements during relatively cloud free conditions that are collected uniformly at all SZA from minimum up to at least 75° (AMF = 3.5). Based on different data filtering for MLE we assume that the error is about 17 ± 4 % over Olympic Park and about 14 ± 5 % over Mt. Taehwa. SCD_{REF} is not truly

30 independent of the uncertainty in ΔSCD, especially the systematic component (e.g. selection of the fitting scenario).

C. Uncertainty in the DS AMF is less than 1% at SZA smaller than 80°. At large SZAs (> 80°) the error arises from: 1) uncertainty in the assumed species profile; 2) uncertainty in the appropriate measurement time and, therefore, the SZA itself due to longer integration times; 3) AMF wavelength dependence due to atmospheric refraction; 4) uncertainty in effective SZA calculation due to refraction; and 5) larger contribution of the scattered photons at longer integration times. In this study only DS measurements taken at SZA < 80° were considered to reduce AMF error to < 0.5%.

We estimate that the total error in DS Pandora HCHO total column measurements during KORUS-AQ is [± 6 (statistical) ± 25 (systematic)] % at SZA < 80°. Table 2 summarizes all the error sources.

Table 2. Pandora HCHO total column error budget from direct sun measurements.

Uncertainty type and source	Olympic Park	Mt. Taehwa
Statistical noise [%]:		
Instrumental and atmospheric	6 ± 3	6 ± 4
Systematic [%]:		
Choice of fitting DOAS scenario	± 10	± 10
Laboratory measured cross-sections:		
HCHO	± 9	± 9
Other (O_3, NO_2, BrO)	± 14	± 14
SCD_{REF} [%]:	17 ± 4	14 ± 5
AMF_{DS} at SZA < 80° [%]:	< 0.5	< 0.5
Overall uncertainty [%]	**±6 (stat)±26(sys)**	**±6 (stat)±24(sys)**

Figure 2 demonstrates dependence of the total HCHO error on the measurement time (AMF) according to Eq. (4). The V-shape is mostly due to the error in SCD_{REF}. The "direction" of V-shape depends on whether SCD_{REF} is overestimated (Λ) or underestimated (V). Since the errors were added in quadrature, Fig. 2 shows an overestimation effect.
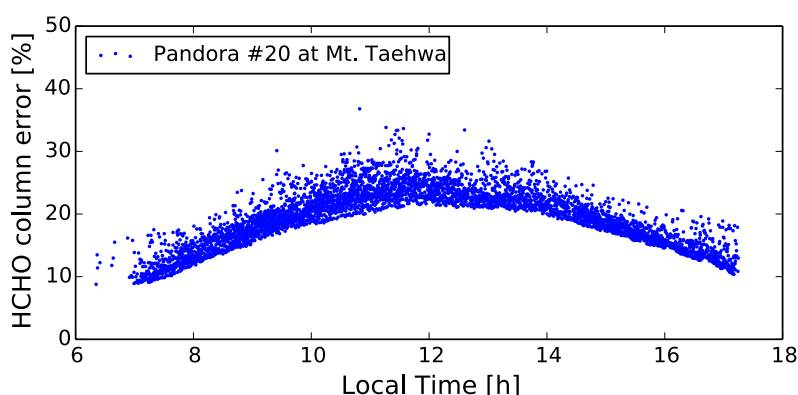
Figure 2. Estimation of total HCHO column errors from Pandora direct sun
measurements during KORUS-AQ (May-June 2016) at Mt. Taehwa.

2.2. Ground-based in situ measurements at Olympic Park and Mt. Taehwa

Surface HCHO concentrations were measured at Mt. Taehwa and Olympic Park
by Quantum Cascade Laser (QCL, from Aerodyne Research, Inc) in-situ instruments
(Hottle et al., 2009) resulting in measured HCHO VMR every 60 seconds with estimated
precision of 0.06 ppb and accuracy of approximately 10%. In-situ HCHO measurements
were conducted by the U.S. Environmental Protection Agency (EPA) at the Olympic
Park Research Site, and by the Aerodyne Research, Inc. at the Mt. Taehwa Site.

Light from 1765 cm^{-1} (Olympic Park) and 2831.6 cm^{-1} (Mt. Taehwa) quantum
cascade lasers were passed through a 0.5 liter absorption cell with an effective path
length of 76 m. Air was sampled at 12 liters per minute from a heated glass inertial inlet
system located at a height of around 10 and 15 meters above ground level for Olympic
Park and Mt. Taehwa sites, respectively.

Figure 3 shows time coincident in situ surface HCHO volume mixing ratios (vmr)
at Olympic Park and Mt. Taehwa. The average vmr during the campaign at Mt Taehwa
was 2.68±1.45 ppb (min=-0.74 ppb, max=9.22 ppb, median=2.39 ppb, Q25=1.59 ppb,
Q75=3.51 ppb). Somewhat higher vmr were observed at Olympic Park: 3.46±1.59 ppb
(min=0.07ppb, max=12.73ppb, median=3.35ppb, Q25=2.38ppb, Q75=4.40ppb). In
general, HCHO surface diurnal variation followed the same pattern at both sites with the
minimum daily HCHO concentrations typically observed between 6:00-7:00 local time
and gradual increase to the maximum between 13:00 and 17:00. The largest differences

between the sites were detected during night and morning hours (from about 21:00 to 11:00 local time). While these statistics provide a valuable overview of surface HCHO at these two sites, a deeper exploration of this behavior is beyond the scope of this paper and will be provided in other KORUS-AQ manuscripts exploring the details of air quality chemistry during the study.
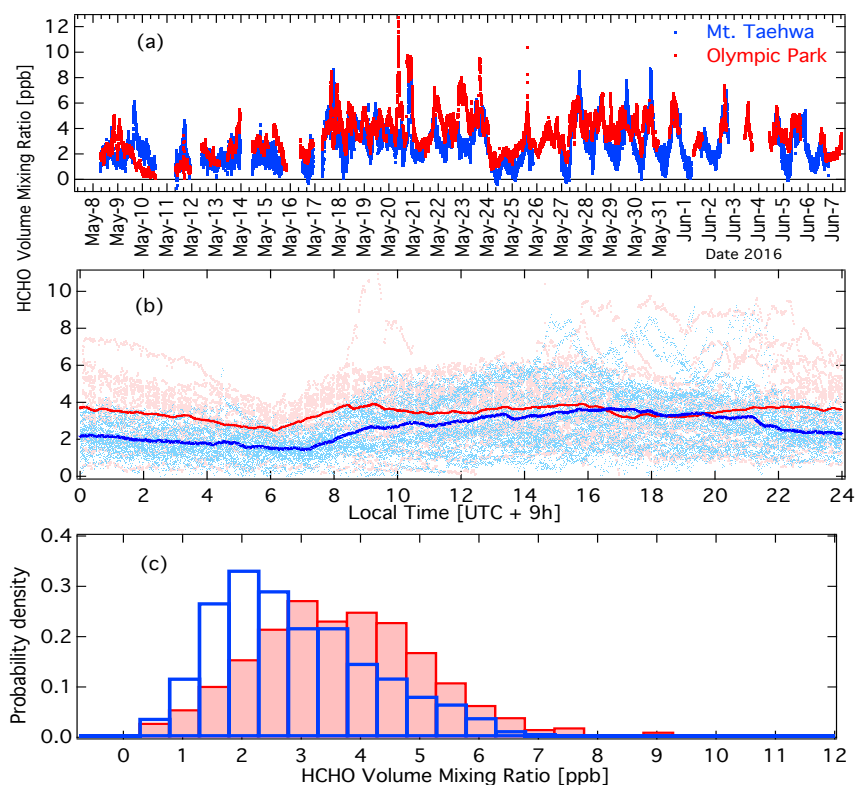
Figure 3. (a) In situ HCHO volume mixing ratios measured at Olympic Park and Mt. Taehwa; (b) Diurnal variation of HCHO volume mixing ratios at Olympic Park and Mt. Taehwa (solid lines represent running averaged data); (c) Overlapping histograms of the HCHO distribution at Olympic Park and Mt. Taehwa during KORUS-AQ.

2.3. Airborne in-situ measurements onboard the NASA DC-8

The Compact Atmospheric Multispecies Spectrometer (CAMS) is a dual channel infrared laser absorption spectrometer that provided measurements of HCHO with 1-second time resolution on the NASA DC-8. A comprehensive description of CAMS can be found in Richter et al., (2015). Briefly, mid-IR laser light at 3.53 μm (2831.6 cm⁻¹) is

generated by non-linear mixing of near-IR lasers in a periodically poled lithium niobate (PPLN) crystal. The combined beams are directed through a multipass Herriott absorption cell (pathlength of 89.6 m) through which ambient air is continuously sampled at a pressure of 50 torr. The lasers are modulated and scanned through an isolated

5   vibrational-rotational HCHO absorption line (2831.6 cm^{-1}). A one standard deviation limit of detection in ambient air is 30 to 50 pptv in 1-second. Based upon the accuracy of our standards along with other factors, we estimate an overall accuracy of 4 to 6% in determining the ambient mixing ratio.

There were a total of 20 local flights of the DC-8 over Korea from 2 May to 10

10   June 2016. As described earlier, flights included routine overflight of the two sites as well as vertical profiling in their vicinity multiple times per day. Figure 4 shows a summary of all flight trajectories and measured HCHO over and near the two sites. We "assigned" data collected below 3 km to a respective site if the ground distance from the site to the aircraft was less then 15 km (Fig 4(a) and 4(c)). This resulted in a total of 38 DC-8

15   measured profiles over Mt. Taehwa and 43 over Olympic Park.

Most DC-8 measurements directly above the Mt. Taehwa site were done at altitude 0.6 - 1 km a.s.l., reaching the minimum altitudes in a narrow path when approaching the site from the west and descending to the east before conducting the spiral ascent (Fig. 4(a)). In-line overpasses over Olympic Park extended to a maximum height

20   of 2 - 3 km north of the site with a variable minimum altitude (0.1±0.17 km) south of the site (Fig. 4(c)). Measurements above 3.5 km have little impact on HCHO vertical column variability (mean = 0.23 ± 0.14 ppb) during the entire study (see Fig. 4b). Considering the short distance between the Olympic Park and Mt. Taehwa sites and the minimal variability of free tropospheric HCHO compared to boundary layer variability, we

25   complement the in-line overpasses over Olympic Park (up to 2-3 km) with the higher altitude portion from profiles over Mt. Taehwa (from 2-3 km to 6-8 km) that are measured within 30 min from the end of the in-line Olympic Park overpass. Figure 4 (b and d) shows diurnal changes in the HCHO vertical distribution. The largest variability in HCHO was observed in the lowest 1 km at both sites as a function of time of day. HCHO

30   was confined to shallow layers in the morning (0.5-0.8 km) and then expanded to up to 2-3 km around 15-16 h local time due to enhanced production and vertical mixing. Morning

HCHO profiles and profiles with low mixing ratios tend to have an exponential function shape. Conversely, profiles during mid-afternoon can be described by a uniform value in the mixed layer with exponential decay to a minimum free tropospheric concentration around 4 km.
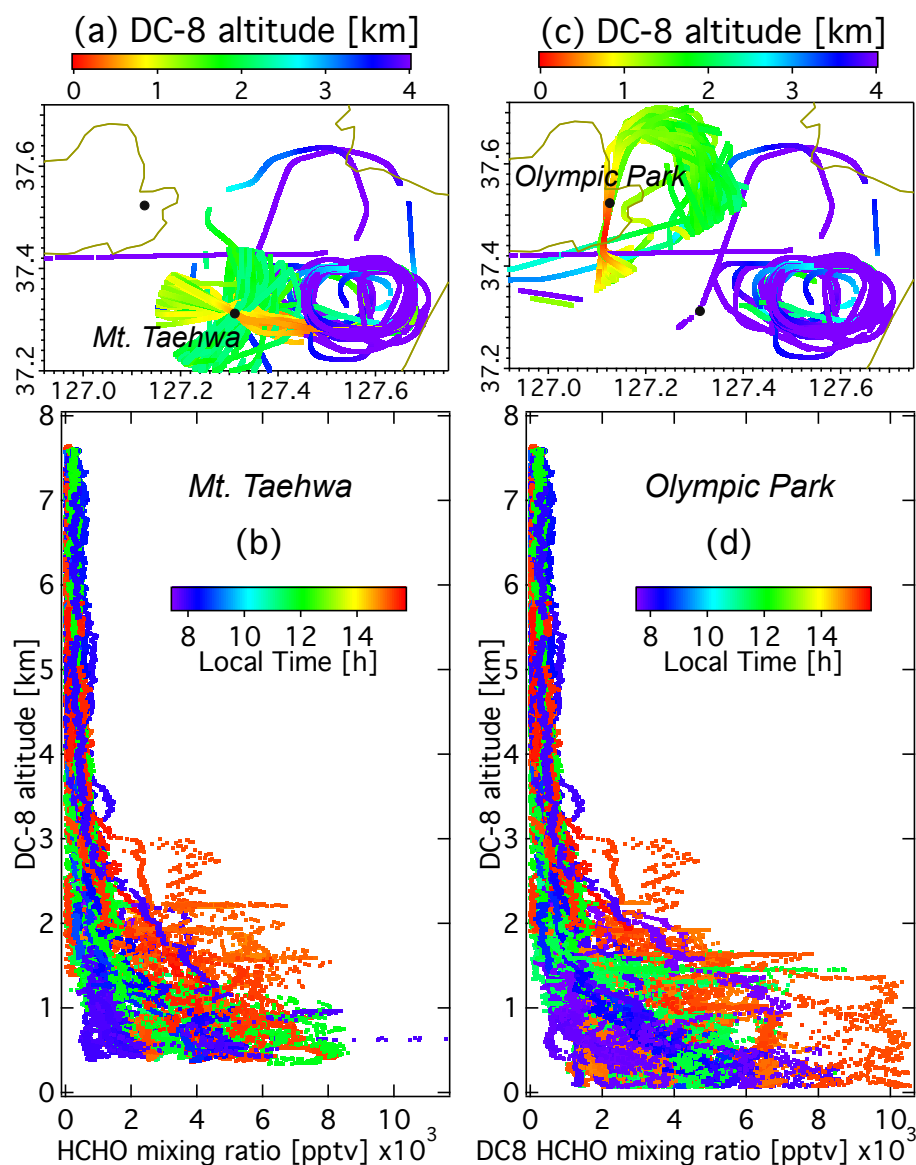
Figure 4. Summary of all DC-8 flights over the Mt. Taehwa and Olympic Park sites: (a) and (c) DC-8 GPS altitude; (b) and (d) HCHO mixing ratios measured onboard the DC-8 as a function of as a function of altitude.

3. Estimation of HCHO vertical column densities from in situ measurements

3.1. Integration of airborne measurements to determine column densities

Figure 5 shows the linear correlation between the in situ ground-based measurements at each site and the aircraft measurements averaged over the lowest 200 m in proximity to Olympic Park and the lowest 600 m a.s.l. near Mt. Taehwa sites. The total duration of flight time needed to sample the corresponding vertical distances was between 0.5 to 3 min. This resulted in ground distance coverage of 17.4 ± 5.9 km around Mt. Taehwa and 9.1 ± 1.5 km around Olympic Park. Generally, near-surface averaged aircraft observations were lower than the in-situ measured surface concentrations (slope of 0.92 at Olympic Park and 0.81 at Mt. Taehwa). DC-8 altitude at the closest site distance (< 0.2 km) was 0.44±0.02 km over Olympic Park and 0.65±0.10 km over Mt. Taehwa.

The absolute difference between the averaged near-surface DC-8 and in situ measurements were 0.74±0.65 ppb for Olympic Park and 0.62±0.40 ppb for Mt. Taehwa. The correlation (R^2) between the in situ ground-based and near surface DC-8 measured HCHO concentrations is 0.69 for Olympic Park and 0.80 for Mt. Taehwa. This suggests some spatial HCHO heterogeneity in the vertical (surface to 200/470 m) and horizontal (up to 23 km) directions.

To account for the partial column between the surface and the lowest aircraft altitude, we complement DC-8 profiles with the in-situ surface measurements. Air density at the surface was calculated from the WS501 Lufft measurements of temperature and pressure at Mt. Taehwa. There were no pressure measurements at Olympic Park so we scaled pressure from Osan Air Base to the Mt. Taehwa altitude. Temperature measurements at Olympic Park did not cover the entire campaign so we used temperature from Mt. Taehwa (+ 2°K) on a few missing data days. We also exclude in-situ HCHO measurements at Mt. Taehwa on 2 June 2016 due to the unreasonably low measurements during the whole day. The partial column above the aircraft was calculated using the mixing ratio measured at the highest DC-8 altitude up to the tropopause height which varied around 12.77 ± 1.63 km. The tropopause height was calculated from the radiosonde temperature profiles launched from Osan Air Base during the campaign. We

estimate that the partial column above the aircraft altitude is about 0.05 to 0.07 DU. These added partial columns above DC-8 maximum altitudes up to the tropopause accounted for about 5±4 % of the total columns. The added column below the lowest DC-8 altitudes down to the site surface accounted for 16±7% over Mt. Taehwa and 3±1% over Olympic Park.

Total columns from DC-8 HCHO profiles were determined by numerical integration of the volume number density from the lowest to the highest altitudes. Errors in derived DC-8 HCHO total columns are comprised of the instrumental uncertainties of the measurements, errors in temperature and pressure profiles, errors due to spatial and

10 temporal heterogeneity of the HCHO distribution in the sampled air relative to the specific volume over the site, and errors due to extrapolation to the parts of the atmosphere not sampled by the aircraft. In this study we approximate errors due to spatial and temporal heterogeneity of the HCHO distribution by comparing DC-8 measurements within the lowest 200 m for Olympic Park and 470 m for Mt. Taehwa to the in-situ

15 surface mixing ratios (see Fig. 5). This uncertainty source leads to a potential underestimation of 8% for Olympic Park and 19% for Mt. Taehwa. Instrumental errors are random and are on the order of 4-6%. We assume that the uncertainty in the partial column above the DC-8 is 50%, which translates to about 2.5% of the total column. We assume that the uncertainty in the partial column below the DC-8 minimum altitude is

20 dominated by the uncertainty due to heterogeneity. Another source of error in the calculated columns over Olympic Park is the potential heterogeneity above the highest DC-8 altitude above Olympic Park (2-3km) and Mt. Taehwa. Considering all these sources the total error in derived VCD from the aircraft measurements is about (–11 ± 6)% for Olympic Park and (–19 ± 6)% over Mt. Taehwa. Negative errors indicate

25 underestimation of the total column.
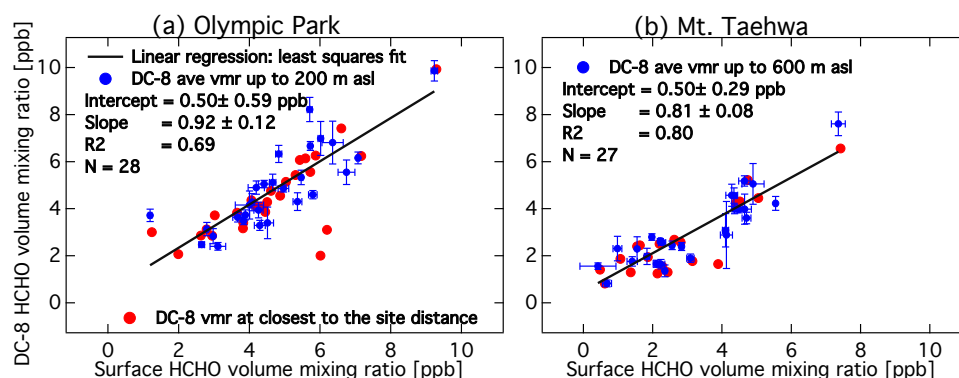
Figure 5. Correlation between surface in situ and near-surface DC-8 measurements at Olympic Park (a) and Mt. Taehwa (b) during KORUS-AQ. All available aircraft measurements were averaged from the lowest DC-8 altitude up to 200 m above sea level (asl) at Olympic Park and up to 600 m asl at Mt. Taehwa (blue circles). Measurements at the closest DC-8 location to the sites are also shown (red circles).

3.2 Deriving column densities from surface measurements and mixing layer height

Given the broader availability of Pandora observations and surface HCHO measurements without the benefit of complementary airborne sampling, we also developed estimates for column densities depending only on in situ surface measurements and information on mixing layer height (MLH) derived from Vaisala Ceilometer CL51 backscatter profiles at 910 nm (Knepp et al., 2017). The main assumption is that most of the HCHO column is located in the well-mixed layer. Figure 6 shows MLH derived from the backscatter profiles at Mt. Taehwa and Olympic Park. The estimated MLH diurnal changes are very similar at both sites. The minimum MLH (300 – 500 m) is during the night and early morning hours (22 – 8 h). PBL growth typically starts around 7-8 in the morning and reaches its maximum (1.5 – 2 km) around 15-16 hours local time. On some days however the estimated MLH peaks later (around 18 h) and is significantly higher (around 3 km). Measured MLH, however, are somewhat lower at Mt. Taehwa compared to Olympic Park in the morning and late afternoon.

To estimate the total column from in-situ surface concentrations we (1) filtered and averaged the MLH data for both sites to generate "measured" MLH and (2) created a median MLH as a function of local time of day from all measurements. A median MLH was used to test the hypothesis of whether a "generic seasonal" estimation of MLH can be applied to relate in-situ surface and column HCHO measurements.

Ceilometer measured MLH can exhibit sporadic variations that are not related to the true changes of MLH. We have examined effect of several filtering schemes on the total columns: (1) no filtering with 5 min averaging of raw MLH; (2) running median (±300 points); (3) running median (±150 points); (4) averaging raw MLH over 5 min

5   after removing data that have large difference with the running median (300 points) > 300 m; (5) averaging raw MLH on 5 min after removing data that have difference with the running median (150 points) > 300 m. In general, filtering of MLH has a small effect on the agreement between the remote sensing columns and columns derived from the in situ measurements (R^2 standard deviation of 0.02). Running median (±300 points) produced

10   the best agreement and is used in the calculation of ground-up columns.
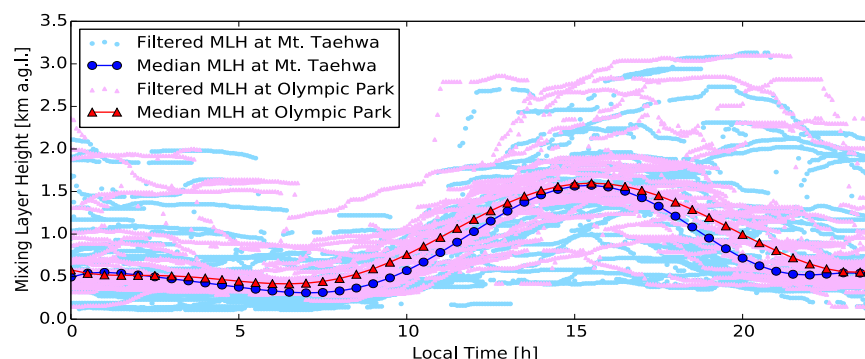
Figure 6. Mixing Layer Height (MLH) above ground level derived from Vaisala Ceilometer CL51 backscatter profiles (910 nm) over Olympic Park and Mt. Taehwa during KORUS-AQ.

We calculated total columns from in-situ measurements (ground-up VCD) using 4 different profile shapes: (1) a uniform HCHO mixing ratio up to the median MLH with a free tropospheric mixing ratio of 0.23 ppb from the MLH to the average tropopause height of 12.77 km; (2) same as (1) but using the "measured" MLH; (3) a uniform HCHO

20   mixing ratio up to the median MLH with a free tropospheric mixing ratio that exponentially decreases above the MLH to 0.23 ppb within 3xMLH or 4 km (whichever is smaller) and remains constant up to the average tropopause height of 12.77 km; (4) same as (3) but using the "measured" MLH.

Free troposphere vmr of 0.23 ± 0.14 ppb is derived from DC-8 in situ

25   measurements during KORUS-AQ. For a location with no aircraft measurements free troposphere vmr can be estimated from chemical transport models (see Fried et al., 2011)

A single temperature and pressure profile for the whole campaign was generated from all available radiosonde measurements. This profile was scaled to account for surface temperature and pressure changes during the campaign within the MLH.

4. Diurnal variability of HCHO columns

5   Total columns from direct-sun Pandora, DC-8 aircraft profiles, and surface measurements (using 4 profile shape assumptions) are shown in Fig. 7 for Olympic Park and Fig. 8 for Mt. Taehwa. Days with no or limited data (e.g. cloud screened Pandora data) were excluded from these figures. Diurnal changes for all of the column estimations show similar trends with minimum VCD typically in early morning and maximum VCD

10   around 14-16 h local time. Figures 7 and 8 show effect of the assumed profile shapes on the derived "ground-up" columns. As expected the profile shapes (2-grey) and (4-light blue) that use measured MLH result in the largest VCD when MLH is larger than the median values.
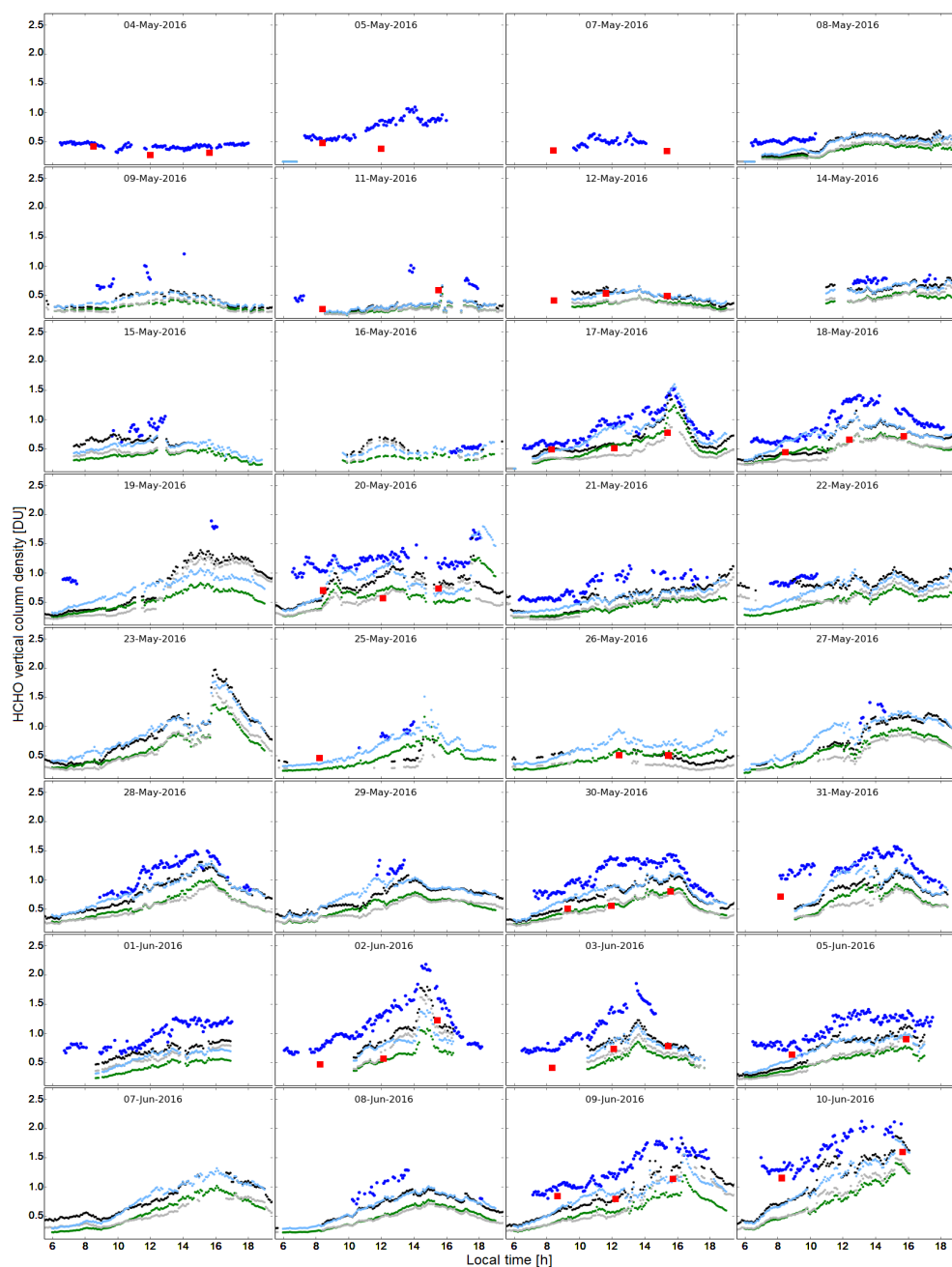
Figure 7. Vertical column densities at Olympic Park during KORUS-AQ derived from: Pandora direct-sun measurements (•); DC-8 (from surface to 12.77 km (■); surface concentrations and profile shapes: (1) box with a median MLH (green); (2) box with a measured MLH (grey); (3) box/exp with a median MLH (light blue); (4) box/exp with a measured MLH (black).
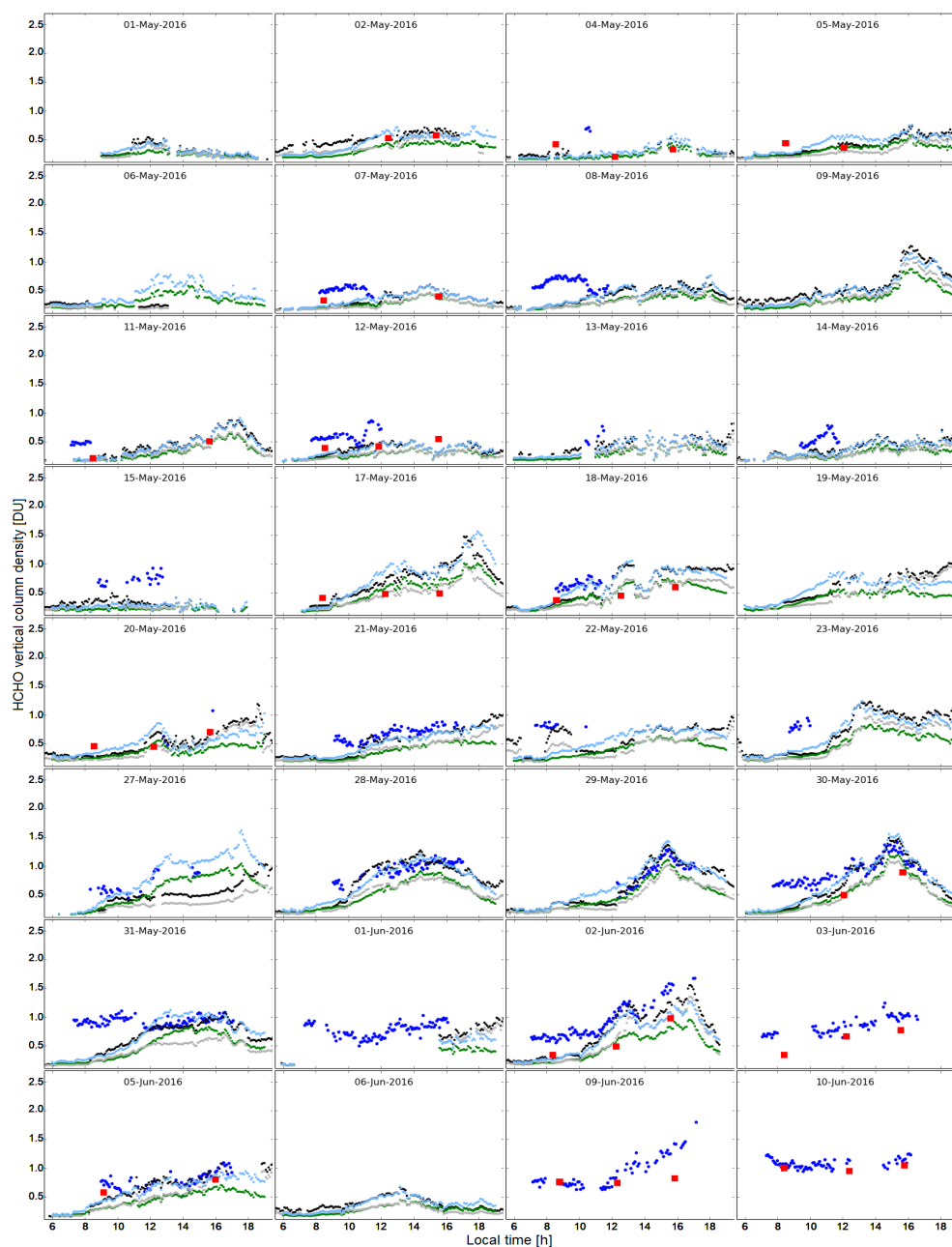
Figure 8. Vertical column densities at Mt. Taehwa during KORUS-AQ derived from: Pandora direct-sun measurements (•); DC-8 (from surface to 12.77 km ■); surface concentrations and profile shapes: (1) box with a median MLH (green); (2) box with a measured MLH (grey); (3) box/exp with a median MLH (light blue); (4) box/exp with a measured MLH (black).

This is obvious on 19 May 2016 (Olympic Park) when measured MLH in the afternoon was 3 km compared to a median MLH of 1.5 km. Considering that exponential function addition to the box shape is limited to 4 km (or 3 x MLH, whichever is smaller) the larger the MLH the smaller the difference between the derived VCD from the

5 corresponding box and box/exponential profile shapes. This is also demonstrated by the afternoon data on 19 May 2016 when "ground-up" VCD from box profile shape (2) and box/exponential profile shape (4) have very small offset between them. However, when MLH is significantly lower than 4 km, exponential decay from the surface measured concentration to 0.23 ppb can add a substantial amount to the total column.

10 "Ground-up" and Pandora columns both exhibit similar HCHO changes on a smaller scale (e.g. 20 May 2016 around 18:00 at Olympic Park). The absolute values however are different. In addition, Pandora total columns tend to have a smaller rate of change between 6 and 10 in the morning compared to the "ground-up" columns at both sites. This could be an indication of underestimation of Pandora SCD_{REF} or inability of

15 the selected profile shapes to capture true HCHO vertical distribution. Support for the later reason was seen in the ceilometers measurements (see Fig. 9), where on the majority of days the ceilometer captured residual layers above the morning mixing layer (ML). The residual layers persisted until late morning when growth of the ML reached the top of the residual layers. The rapid growth of the ML, which is typical in early morning

20 hours, would also explain the larger rate of change in the morning hours in the "ground up" columns.
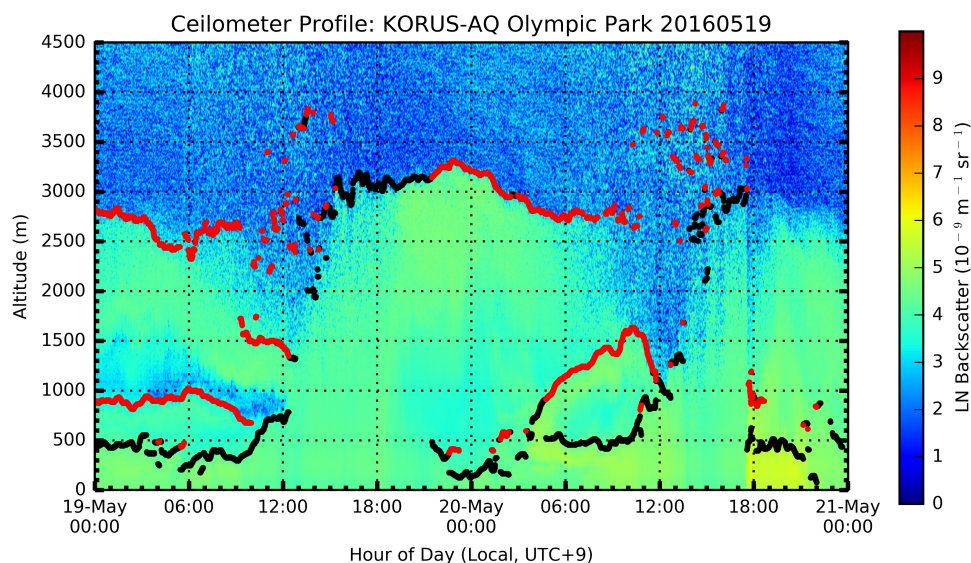
Figure 9. Vaisala Ceilometer CL51 backscatter measurements (910 nm) on 19 and 20 May 2016 at Olympic Park, during KORUS-AQ. Black dots represent MLH used in this study.

5. Comparison of DC-8 HCHO columns with Pandora and "ground-up" columns.

DC-8 integrated columns tend to be within the variability of the "ground-up" columns from the four profile shapes and are typically smaller than the Pandora measurements. Figure 10(a) shows linear regression between Pandora and DC-8 HCHO columns at the two sites with the slope equal to 1.16 ± 0.23, intercept $= 0.22 \pm 0.16$ DU and $R^2 = 0.68$ (49 measurements). Interpretation of the differences, however, is not straightforward since there are multiple occasions when the agreement is very good (e.g. 5, 10 June 2016 at Mt. Taehwa and 4 May 2016 at Olympic Park). On the other hand, there are days (e.g. 20 and 30 May 2016 at Olympic Park) when the differences between the DC-8 and Pandora VCD are 0.5-1 DU. Such large differences for some days and small differences for other days cannot be explained by the errors in SCD_{REF} or SCD measurements alone, and most likely are the result of spatial and temporal heterogeneity of HCHO distribution and differences in volume sampling by DC-8 and Pandora. Pandora column overestimation of 16% relative to DC-8 HCHO columns is within the potential underestimation errors in DC-8 columns of $\sim 11\%$ for Olympic Park and 19%

for Mt. Taehwa. It also can indicate a potential error due to DOAS fitting scenario selection.

Figure 10(b) shows linear regression analysis results for "ground-up" columns best agreeing with the DC-8 columns (box profile shape with measured MLH, (2)). This

5 profile shape has linear regression correlation $R^2 = 0.69$, slope 0.99 ± 0.18 and intercept - 0.17 ± 0.12 DU. The error in surface derived columns represents a standard deviation between the 4 different profile shapes used to create the columns.
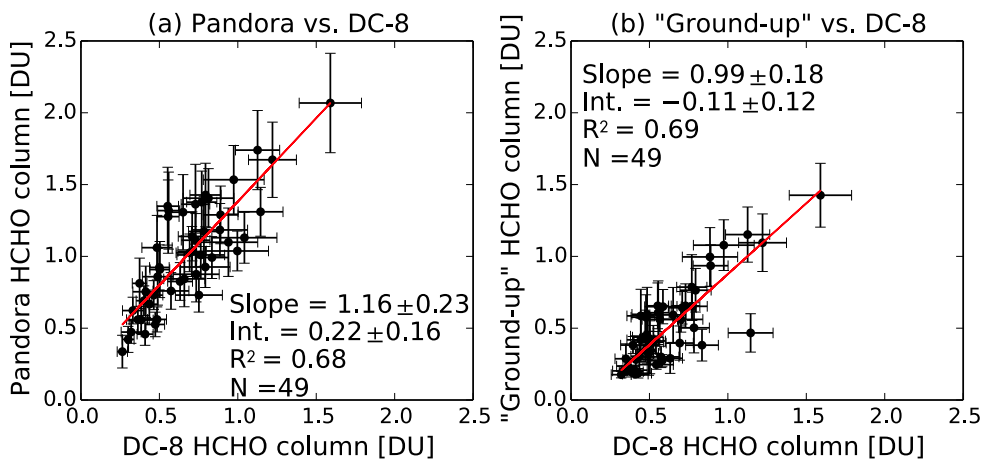
Figure 10. Correlation between HCHO columns for (a) Pandora and DC-8 integrated
10 vertical columns and (b) "ground-up" (surface vmr within measured MLH, box profile shape) and DC-8 integrated columns at Olympic Park and Mt. Taehwa during KORUS-AQ.

Table 3 shows that the agreement is much worse between DC-8 HCHO VCD and
15 "ground-up" VCD for other profile shapes. This discrepancy can be an indication that the chosen profile shapes are not representative of the actual HCHO distribution especially for very shallow MLH.

Table 3. Linear correlation between HCHO total columns from Pandora direct sun measurements and columns calculated from in situ surface concentration measurements based on different profile shape assumptions (Y) vs. HCHO integrated columns from DC-8 (X).

VCD source	Profile shape	Slope	Intercept	R^2	N
In-situ[1] (ground-up)	Box: measured MLH	0.99 ± 0.18	-0.11 ± 0.12	0.69	49
	Box + exponential: measured MLH	1.13 ± 0.25	0.00 ± 0.17	0.62	49
	Box: median MLH	0.71 ± 0.17	0.05 ± 0.11	0.55	57
	Box + exponential: median MLH	1.07 ± 0.25	0.02 ± 0.16	0.48	58
Pandora[2]	NA	1.16 ± 0.23	0.22 ± 0.16	0.68	49

5 [1]All measurements are averaged over 5 min intervals.
[2] Pandora measurements are filtered based on the maximum relative instrumental/atmospheric HCHO column error of 15% and maximum normalized residual rms of 0.002.

10 **6. Comparison of Pandora HCHO columns with "ground-up" columns.**

Based on the DC-8 measured HCHO profile discussion and diurnal changes in the ceilometer-determined MLH, we do not expect any meaningful correlation between the Pandora total columns and in-situ surface concentrations. Indeed, Figure 11 (a) and (c) show a general correspondence between surface HCHO measurements at Olympic Park

15 and Mt. Taehwa and Pandora column measurements, but the relationship is too diffuse to allow surface values to be derived from column measurements or vice versa.
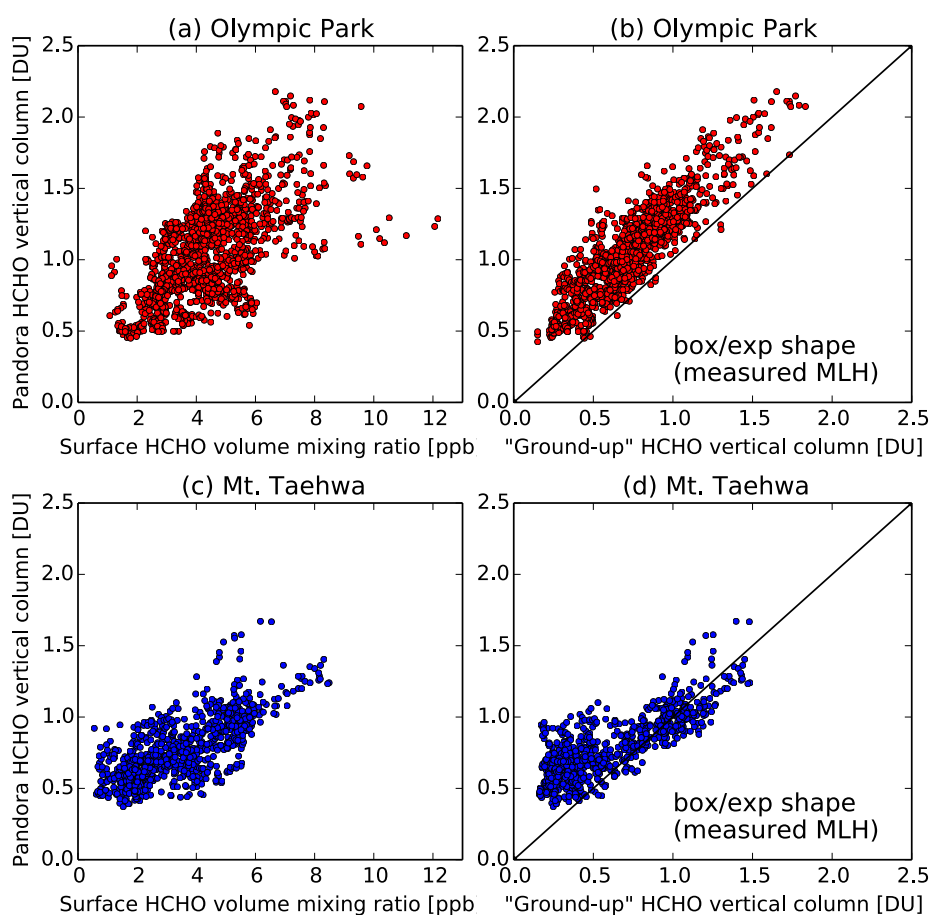
Figure 11. Pandora measured HCHO vertical column densities vs. surface in situ mixing ratios and columns calculated from the surface vmr at Olympic Park and Mt. Taehwa during KORUS-AQ (May-June 2016).

Linear regression analysis in Fig. 11 was done between Pandora (Y) and "Ground-up" (X) HCHO columns to identify which profile shape is more representative of Pandora column measurements. The best correlation ($R^2 = 0.78\pm0.02$) and slope (1.03 ± 0.03) were determined for profile shape 4 (box and exponential profile with measured MLH) at Olympic Park (see Fig. 11(b)). The intercept of 0.29 ± 0.02 DU could be the result of incorrect selection of DOAS fitting window, which can cause a constant offset. To improve the agreement between Pandora and "ground-up" column diurnal patterns a more in depth analysis is required to determine if a larger SCDREF is needed, causing a

larger offset, or the "ground up" columns systematically underestimated due to elevated layers not captured in the "group up" model.

Table 4 summarizes linear regression results for all profile shapes. Standard deviations in Table 4 for the profile shapes with measured MLH represent the effect of

5   different filtering of MLH data. In general, the effect of MLH filtering is very small. For columns derived from the box and exponential profile shape with the measured MLH and only 5 min MLH averaging the correlation with Pandora columns (R^2) is 0.76. The same profile shape but using a 300-point running median resulted in correlation (R^2) of 0.80. In the absence of measured MLH a median MLH combined with an exponential function

10 still can be relatively accurately used to estimate a near surface concentration from Pandora HCHO columns ($R^2 = 0.68$ and slope = 1.06).

The correlation between Pandora and "ground-up" columns at Mt. Taehwa is worse than at Olympic Park since there were fewer Pandora full day measurements at Mt. Taehwa because of instrumental issues early in the campaign. There were several days

15 that had only morning Pandora measurements. During morning hours measured MLH was relatively shallow (~ 300 m) at Mt. Taehwa and the agreement between very small "ground-up" columns and Pandora columns was poor. Figure 11(d) shows that for "ground-up" columns greater than 0.7 DU the agreement is significantly improved (negligible offset, and slope close to 1). More investigation is needed to understand the

20 differences between the two sites.

Table 4. Linear correlation between HCHO total columns from Pandora direct sun measurements (Y) and columns calculated from in situ surface concentration measurements based on different profile shape assumptions (X). All measurements are averaged over 5 min intervals. Least absolute residual fitting method is used.

Site	Profile shape	Slope	Intercept	R^2	N
Olympic Park	Box: measured MLH	1.33 ± 0.05	0.34 ± 0.03	0.73 ± 0.03	1280 - 1469
	Box + exponential: measured MLH	1.03 ± 0.03	0.29 ± 0.02	0.78 ± 0.02	1280 - 1469
	Box: median MLH	1.34 ± 0.01	0.33 ± 0.01	0.67	1547
	Box + exponential: median MLH	1.06 ± 0.01	0.25 ± 0.01	0.68	1547
Mt. Taehwa	Box: measured MLH	0.73 ± 0.02	0.47 ± 0.01	0.64 ± 0.02	737 - 826
	Box + exponential: measured MLH	0.51 ± 0.01	0.46 ± 0.01	0.67 ± 0.01	737 - 826
	Box: median MLH	0.67 ± 0.01	0.47 ± 0.00	0.59	879
	Box + exponential: median MLH	0.50 ± 0.00	0.47 ± 0.00	0.57	879

6. Conclusions and Discussion

We have presented a first evaluation of Pandora total column HCHO measurements collected in continuous direct-sun observation mode during the KORUS-AQ 2016 field study. The total column measurements were compared to the integrated DC-8 in-situ profile measurements and in-situ scaled columns assuming different profile shapes.

The following observations were made:

1. The largest sources of uncertainty in Pandora HCHO DS column measurements are from:

(A) Systematic errors due to selection of the fitting window and choice of the cross sections. The combined error is on the order of ± 25% and is responsible for an offset in vertical column that is mostly independent of AMF.

(B) Estimation of SCD_{REF} in the reference spectrum using MLE and is on the order of 14-17%. This error depends on the diurnal variation of HCHO optical depth for DS AMF. Further studies are needed to understand the effect of MLE on derived SCD_{REF} for conditions with very small HCHO production rates and/or very systematic diurnal

patterns. Error in SCD_{REF} introduces AMF dependent error in HCHO total columns resulting in up or down bowing of the diurnal changes.

2. The statistical HCHO total column errors were $\pm (6 \pm 4)\%$. This indicates that using new head sensor ARC window significantly reduced the spurious spectral structure present in the previous versions of Pandora DS measurements.

3. DC-8 in-situ profile measurements were done over limited altitude ranges. On average the DC-8 integrated columns were complemented with (a) about 5 ± 4 % of the total columns from the maximum DC-8 altitude to tropopause over Mt. Taehwa and (b) column below the lowest DC-8 altitudes down to the site surface of $16\pm7\%$ over Mt. Taehwa and $3\pm1\%$ over Olympic Park. No profile measurements were conducted above 2-3 km over Olympic Park. An assumption was made that HCHO vertical distributions above DC-8 maximum altitude over Olympic Park are the same as those 25-40 km southeast over Mt. Taehwa.

4. DC-8 in-situ profile measurements (< 3 km) used in this evaluation were within 15 km radius of each site. The DC-8 flight trajectories did not coincide with Pandora east-south-west direct sun line of sight.

5. DC-8 measurements in the lowest 200 m and 470 m above Olympic Park and Mt. Taehwa were on average 8% and 19% lower than the time coincident surface concentrations at the corresponding sites indicating spatial (vertical and horizontal) heterogeneity of HCHO distribution within 15-20 km.

6. Pandora HCHO total columns were on average 16% larger than DC-8 integrated profiles with an offset bias of 0.22 DU and correlation coefficient (R^2) of 0.68. The source of this difference will require further evaluation since underestimation of DC-8 integrated final total columns and overestimation of Pandora total columns are possible. This issue can be potentially resolved in future by placing Pandora at the "touch down" location and sky scanning in addition to DS measurements in the aircraft direction.

7. DC-8 measured morning HCHO profiles and profiles with low mixing ratios had an exponential function shape. Profiles during mid-afternoon can be described by a uniform value in the mixed layer with exponential decay to a minimum free tropospheric concentration around 4 km (0.23 ppb).

8. Based on DC-8 profile shape and Ceilometer backscatter estimation of MLH we calculated total columns from in-situ measurements ("ground-up" VCD) using 4 different profile shapes: (1) a uniform HCHO vmr up to the median MLH with a free tropospheric mixing ratio of 0.23 ppb from the MLH to the average tropopause height of 12.77 km; (2) same as (1) but using the measured MLH; (3) a uniform HCHO mixing ratio up to the median MLH with a free tropospheric mixing ratio that exponentially decreases above the MLH to 0.23 ppb within 3xMLH or 4 km (whichever is smaller) and remains constant up to the average tropopause height of 12.77 km; (4) same as (3) but using the measured MLH. The main goal was to determine whether any of these profile shapes can be used to convert column measurement into surface concentrations

9. Comparison between Pandora and "ground-up" columns over Olympic Park suggested that profile shape (4) with measured MLH and exponential decay produced the best agreement (slope = 1.03 ± 0.03, intercept = 0.29 ± 0.02 DU and R2 = 0.78±0.02) The source of the offset bias is not clear at this point.

10. Pandora HCHO columns and "ground-up" columns disagree the most early in the morning when MLH are very shallow, and the ceilometers detect elevated residual layers. This disagreement is likely due the the tested shapes not adequately capturing the elevated layers during these conditions (aerosol driven MLH is not representative of HCHO distribution when elevated layers are present)

11. Based on DC-8 and "ground-up" comparison Pandoras were able to capture diurnal variation of HCHO column with some positive bias. This makes Pandora an excellent validation instrument for TEMPO.

Acknowledgments:

All data used in this work can be downloaded from: https://www-air.larc.nasa.gov/missions/korus-aq/. The authors would like to acknowledge the National Institute of Environmental Research (NIER) for their tremendous effort during KOUR-AQ. In particular we would like to thank NIER for their effort to establish the Olympic Park research site, and their assistance with logistics for measurements at both Olympic Park and Mt. Taewha Forest. This work was supported under the NASA Tropospheric Chemistry Program and the EPA Air, Climate, and Energy Research

Program. Although this paper has been reviewed by the EPA and approved for publication, it does not necessarily reflect EPA policies or views. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

References

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

10   Bauwens, M., Stavrakou, T., Müller, J.-F., De Smedt, I., Van Roozendael, M., van der Werf, G. R., Wiedinmyer, C., Kaiser, J. W., Sindelarova, K. and Guenther, A.: Nine years of global hydrocarbon emissions based on source inversion of OMI formaldehyde observations, Atmospheric Chemistry and Physics, 16(15), 10133–10158, doi:10.5194/acp-16-10133-2016, 2016.

15   Bogumil, K., Orphal, J., Homann, T., Voigt, S., Spietz, P., Fleischmann, O. ., Vogel, A., Hartmann, M., Kromminga, H., Bovensmann, H., Frerick, J. and Burrows, J. .: 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, Journal of Photochemistry and Photobiology A: Chemistry, 157(2–
20   3), 167–184, doi:10.1016/S1010-6030(03)00062-5, 2003.

Burrows, J. P., Dehn, A., Deters, B., Himmelmann, S., Richter, A., Voigt, S. and Orphal, J.: ATMOSPHERIC REMOTE-SENSING REFERENCE DATA FROM GOME: PART 1. TEMPERATURE-DEPENDENT ABSORPTION CROSS-SECTIONS OF NO2 IN THE 231–794 nm RANGE, Journal of Quantitative Spectroscopy and Radiative Transfer,
25   60(6), 1025–1031, doi:10.1016/S0022-4073(97)00197-0, 1998.

Cede, A.: Manual for Blick Software Suite 1.3, [online] Available from: http://pandonia.net/media/documents/BlickSoftwareSuite_Manual_v7.pdf, 2017.

De Smedt, I., Stavrakou, T., Hendrick, F., Danckaert, T., Vlemmix, T., Pinardi, G., Theys, N., Lerot, C., Gielen, C., Vigouroux, C., Hermans, C., Fayt, C., Veefkind, P.,
30   Müller, J.-F. and Van Roozendael, M.: Diurnal, seasonal and long-term variations of global formaldehyde columns inferred from combined OMI and GOME-2 observations, Atmospheric Chemistry and Physics, 15(21), 12519–12545, doi:10.5194/acp-15-12519-2015, 2015.

Fleischmann, O. C., Hartmann, M., Burrows, J. P. and Orphal, J.: New ultraviolet
35   absorption cross-sections of BrO at atmospheric temperatures measured by time-windowing Fourier transform spectroscopy, Journal of Photochemistry and Photobiology A: Chemistry, 168(1–2), 117–132, doi:10.1016/j.jphotochem.2004.03.026, 2004.

Fried, A., Cantrell, C., Olson, J., Crawford, J. H., Weibring, P., Walega, J., Richter, D., Junkermann, W., Volkamer, R., Sinreich, R., Heikes, B. G., O'Sullivan, D., Blake, D. R., Blake, N., Meinardi, S., Apel, E., Weinheimer, A., Knapp, D., Perring, A., Cohen, R. C., Fuelberg, H., Shetter, R. E., Hall, S. R., Ullmann, K., Brune, W. H., Mao, J., Ren, X.,

5   Huey, L. G., Singh, H. B., Hair, J. W., Riemer, D., Diskin, G. and Sachse, G.: Detailed comparisons of airborne formaldehyde measurements with box models during the 2006 INTEX-B and MILAGRO campaigns: potential evidence for significant impacts of unmeasured and multi-generation volatile organic carbon compounds, Atmospheric Chemistry and Physics, 11(22), 11867–11894, doi:10.5194/acp-11-11867-2011, 2011.

10   Fu, T.-M., Jacob, D. J., Palmer, P. I., Chance, K., Wang, Y. X., Barletta, B., Blake, D. R., Stanton, J. C. and Pilling, M. J.: Space-based formaldehyde measurements as constraints on volatile organic compound emissions in east and south Asia and implications for ozone, Journal of Geophysical Research, 112(D6), doi:10.1029/2006JD007853, 2007.

Herman, J., Cede, A., Spinei, E., Mount, G., Tzortziou, M. and Abuhassan, N.: NO_2

15   column amounts from ground-based Pandora and MFDOAS spectrometers using the direct-sun DOAS technique: Intercomparisons and application to OMI validation, Journal of Geophysical Research, 114(D13), doi:10.1029/2009JD011848, 2009.

Hermans, C., Vandaele, A. C., Fally, S., Carleer, M., Colin, R., Coquart, B., Jenouvrier, A. and Merienne, M.-F.: Absorption Cross-section of the Collision-Induced Bands of

20   Oxygen from the UV to the NIR, in Weakly Interacting Molecular Pairs: Unconventional Absorbers of Radiation in the Atmosphere, edited by C. Camy-Peyret and A. A. Vigasin, pp. 193–202, Springer Netherlands, Dordrecht., 2003.

Hottle, J. R., Huisman, A. J., DiGangi, J. P., Kammrath, A., Galloway, M. M., Coens, K. L. and Keutsch, F. N.: A Laser Induced Fluorescence-Based Instrument for In-Situ

25   Measurements of Atmospheric Formaldehyde, Environmental Science & Technology, 43(3), 790–795, doi:10.1021/es801621f, 2009.

Jin, X. and Holloway, T.: Spatial and temporal variability of ozone sensitivity over China observed from the Ozone Monitoring Instrument: OZONE SENSITIVITY OVER CHINA, Journal of Geophysical Research: Atmospheres, 120(14), 7229–7246,

30   doi:10.1002/2015JD023250, 2015.

Jin, X., Fiore, A. M., Murray, L. T., Valin, L. C., Lamsal, L. N., Duncan, B., Folkert Boersma, K., De Smedt, I., Abad, G. G., Chance, K. and Tonnesen, G. S.: Evaluating a Space-Based Indicator of Surface Ozone-NO_x-VOC Sensitivity Over Midlatitude Source Regions and Application to Decadal Trends: Space-Based Indicator of O_3

35   Sensitivity, Journal of Geophysical Research: Atmospheres, 122(19), 10,439-10,461, doi:10.1002/2017JD026720, 2017.

Knepp, T. N., Szykman, J. J., Long, R., Duvall, R. M., Krug, J., Beaver, M., Cavender, K., Kronmiller, K., Wheeler, M., Delgado, R., Hoff, R., Berkoff, T., Olson, E., Clark, R., Wolfe, D., Van Gilst, D. and Neil, D.: Assessment of mixed-layer height estimation from

single-wavelength ceilometer profiles, Atmospheric Measurement Techniques, 10(10), 3963–3983, doi:10.5194/amt-10-3963-2017, 2017.

Lee, C., Kim, Y. J., Hong, S.-B., Lee, H., Jung, J., Choi, Y.-J., Park, J., Kim, K.-H., Lee, J.-H., Chun, K.-J. and Kim, H.-H.: Measurement of Atmospheric Formaldehyde and Monoaromatic Hydrocarbons using Differential Optical Absorption Spectroscopy during Winter and Summer Intensive Periods in Seoul, Korea, Water, Air, and Soil Pollution, 166(1–4), 181–195, doi:10.1007/s11270-005-7308-6, 2005.

MacDonald, S. M., Oetjen, H., Mahajan, A. S., Whalley, L. K., Edwards, P. M., Heard, D. E., Jones, C. E. and Plane, J. M. C.: DOAS measurements of formaldehyde and glyoxal above a south-east Asian tropical rainforest, Atmospheric Chemistry and Physics, 12(13), 5949–5962, doi:10.5194/acp-12-5949-2012, 2012.

Malicet, J., Daumont, D., Charbonnier, J., Parisse, C., Chakir, A. and Brion, J.: Ozone UV spectroscopy. II. Absorption cross-sections and temperature dependence, Journal of Atmospheric Chemistry, 21(3), 263–273, doi:10.1007/BF00696758, 1995.

Marais, E. A., Jacob, D. J., Kurosu, T. P., Chance, K., Murphy, J. G., Reeves, C., Mills, G., Casadio, S., Millet, D. B., Barkley, M. P., Paulot, F. and Mao, J.: Isoprene emissions in Africa inferred from OMI observations of formaldehyde columns, Atmospheric Chemistry and Physics, 12(14), 6219–6235, doi:10.5194/acp-12-6219-2012, 2012.

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

Palmer, P. I.: Mapping isoprene emissions over North America using formaldehyde column observations from space, Journal of Geophysical Research, 108(D6), doi:10.1029/2002JD002153, 2003.

Palmer, P. I., Abbot, D. S., Fu, T.-M., Jacob, D. J., Chance, K., Kurosu, T. P., Guenther, A., Wiedinmyer, C., Stanton, J. C., Pilling, M. J., Pressley, S. N., Lamb, B. and Sumner, A. L.: Quantifying the seasonal and interannual variability of North American isoprene emissions using satellite observations of the formaldehyde column, Journal of Geophysical Research, 111(D12), doi:10.1029/2005JD006689, 2006.

Pinardi, G., Van Roozendael, M., Abuhassan, N., Adams, C., Cede, A., Clémer, K., Fayt, C., Frieß, U., Gil, M., Herman, J., Hermans, C., Hendrick, F., Irie, H., Merlaud, A., Navarro Comas, M., Peters, E., Piters, A. J. M., Puentedura, O., Richter, A., Schönhardt, A., Shaiganfar, R., Spinei, E., Strong, K., Takashima, H., Vrekoussis, M., Wagner, T., Wittrock, F. and Yilmaz, S.: MAX-DOAS formaldehyde slant column measurements during CINDI: intercomparison and analysis improvement, Atmospheric Measurement Techniques, 6(1), 167–185, doi:10.5194/amt-6-167-2013, 2013.

Platt, U. and Stutz, J.: Differential optical absorption spectroscopy: principles and applications, Springer, Berlin., 2008.

Richter, D., Weibring, P., Walega, J. G., Fried, A., Spuler, S. M. and Taubman, M. S.: Compact highly sensitive multi-species airborne mid-IR spectrometer, Applied Physics
5   B, 119(1), 119–131, doi:10.1007/s00340-015-6038-8, 2015.

Schroeder, J. R., Crawford, J. H., Fried, A., Walega, J., Weinheimer, A., Wisthaler, A., Müller, M., Mikoviny, T., Chen, G., Shook, M., Blake, D. R., Diskin, G., Estes, M., Thompson, A. M., Lefer, B. L., Long, R. and Mattson, E.: Formaldehyde column density measurements as a suitable pathway to estimate near-surface ozone tendencies from
10 space: O3-CH2O RELATIONSHIP, Journal of Geophysical Research: Atmospheres, 121(21), 13,088-13,112, doi:10.1002/2016JD025419, 2016.

Shim, C., Wang, Y., Choi, Y., Palmer, P. I., Abbot, D. S. and Chance, K.: Constraining global isoprene emissions with Global Ozone Monitoring Experiment (GOME) formaldehyde column measurements, Journal of Geophysical Research, 110(D24),
15 doi:10.1029/2004JD005629, 2005.

Stutz, J. and Platt, U.: Numerical analysis and estimation of the statistical error of differential optical absorption spectroscopy measurements with least-squares methods, Applied Optics, 35(30), 6041, doi:10.1364/AO.35.006041, 1996.

Valin, L. C., Fiore, A. M., Chance, K. and González Abad, G.: The role of OH
20 production in interpreting the variability of CH_2O columns in the southeast U.S.: OH, VOC and CH_2O Columns in the SOUTHEAST U.S.A., Journal of Geophysical Research: Atmospheres, 121(1), 478–493, doi:10.1002/2015JD024012, 2016.

Vandaele, A. C., Hermans, C., Simon, P. C., Carleer, M., Colin, R., Fally, S., Mérienne, M. F., Jenouvrier, A. and Coquart, B.: Measurements of the NO2 absorption cross-
25 section from 42 000 cm−1 to 10 000 cm−1 (238–1000 nm) at 220 K and 294 K, Journal of Quantitative Spectroscopy and Radiative Transfer, 59(3–5), 171–184, doi:10.1016/S0022-4073(97)00168-4, 1998.

Vlemmix, T., Hendrick, F., Pinardi, G., De Smedt, I., Fayt, C., Hermans, C., Piters, A., Wang, P., Levelt, P. and Van Roozendael, M.: MAX-DOAS observations of aerosols,
30 formaldehyde and nitrogen dioxide in the Beijing area: comparison of two profile retrieval approaches, Atmospheric Measurement Techniques, 8(2), 941–963, doi:10.5194/amt-8-941-2015, 2015.

Wilmouth, D. M., Hanisco, T. F., Donahue, N. M. and Anderson, J. G.: Fourier Transform Ultraviolet Spectroscopy of the $A\,{}^2\Pi_{3/2} \leftarrow X\,{}^2\Pi_{3/2}$ Transition of BrO[†], The
35 Journal of Physical Chemistry A, 103(45), 8935–8945, doi:10.1021/jp991651o, 1999.

Wittrock, F., Richter, A., Oetjen, H., Burrows, J. P., Kanakidou, M., Myriokefalitakis, S., Volkamer, R., Beirle, S., Platt, U. and Wagner, T.: Simultaneous global observations of

glyoxal and formaldehyde from space, Geophysical Research Letters, 33(16), doi:10.1029/2006GL026310, 2006.

Zoogman, P., Liu, X., Suleiman, R. M., Pennington, W. F., Flittner, D. E., Al-Saadi, J. A., Hilton, B. B., Nicks, D. K., Newchurch, M. J., Carr, J. L., Janz, S. J., Andraschko, M.

5 R., Arola, A., Baker, B. D., Canova, B. P., Chan Miller, C., Cohen, R. C., Davis, J. E., Dussault, M. E., Edwards, D. P., Fishman, J., Ghulam, A., González Abad, G., Grutter, M., Herman, J. R., Houck, J., Jacob, D. J., Joiner, J., Kerridge, B. J., Kim, J., Krotkov, N. A., Lamsal, L., Li, C., Lindfors, A., Martin, R. V., McElroy, C. T., McLinden, C., Natraj, V., Neil, D. O., Nowlan, C. R., O′Sullivan, E. J., Palmer, P. I., Pierce, R. B., Pippin, M.

10 R., Saiz-Lopez, A., Spurr, R. J. D., Szykman, J. J., Torres, O., Veefkind, J. P., Veihelmann, B., Wang, H., Wang, J. and Chance, K.: Tropospheric emissions: Monitoring of pollution (TEMPO), Journal of Quantitative Spectroscopy and Radiative Transfer, 186, 17–39, doi:10.1016/j.jqsrt.2016.05.008, 2017.