AGU PUBLICATIONS

Geophysical Research Letters

RESEARCH LETTER
10.1002/2015GL063204

Key Points:
- A new method for HCHO retrievals with small noise and artifacts
- OMPS suitable for HCHO retrievals despite relatively low resolution
- Biomass burning as a main factor driving seasonal and regional patterns in HCHO

Supporting Information:
- Figures S1–S15 and Table S1

Correspondence to:
C. Li,
can.li@nasa.gov

Citation:
Li, C., J. Joiner, N. A. Krotkov, and L. Dunlap (2015), A new method for global retrievals of HCHO total columns from the Suomi National Polar-orbiting Partnership Ozone Mapping and Profiler Suite, *Geophys. Res. Lett.*, *42*, 2515–2522, doi:10.1002/2015GL063204.

Received 21 JAN 2015
Accepted 17 MAR 2015
Accepted article online 21 MAR 2015
Published online 8 APR 2015

A new method for global retrievals of HCHO total columns from the Suomi National Polar-orbiting Partnership Ozone Mapping and Profiler Suite

Can Li[1,2], Joanna Joiner[2], Nickolay A. Krotkov[2], and Laura Dunlap[3]

[1]Earth System Science Interdisciplinary Center, University of Maryland, College Park, Maryland, USA, [2]NASA Goddard Space Flight Center, Greenbelt, Maryland, USA, [3]Department of Atmospheric and Oceanic Science, University of Maryland, College Park, Maryland, USA

Abstract We introduce a new method for retrieving formaldehyde (HCHO) based on principal component analysis (PCA) of satellite-measured radiances. Applying the technique to the Suomi National Polar-orbiting Partnership/Ozone Mapping and Profiler Suite (S-NPP/OMPS) radiances between 328.5 and 356.5 nm, we extract principal components (PCs) associated with various physical processes (e.g., ozone absorption and rotational Raman scattering) and measurement details (e.g., wavelength shift). These PCs, along with precomputed HCHO Jacobians, are utilized in spectral fitting to estimate HCHO loading and reduce interferences. Comparisons with model simulations and independent Ozone Monitoring Instrument (OMI) retrievals indicate that our algorithm can detect enhanced HCHO signals over source regions such as the southeast U.S., producing HCHO total columns with similar spatial distributions and seasonal patterns. While our OMPS retrievals are ~15–20% lower than OMI retrievals from a different algorithm, the differences may be attributed to several instrumental and algorithmic factors. This study demonstrates the potential of PCA algorithms and of OMPS for continuing the long-term satellite HCHO data record.

1. Introduction

Formaldehyde (HCHO) is an intermediate oxidation product of volatile organic compounds. It has a short lifetime of just hours [e.g., *Shim et al.*, 2005] and consequently has a highly inhomogeneous spatio-temporal distribution. In remote oceanic regions, HCHO is predominantly produced from methane oxidation with concentrations typically in the range of ~0.2–1 ppb [e.g., *Singh et al.*, 2001]. For regions having significant biogenic, anthropogenic, or biomass burning sources of nonmethane volatile organic compounds (NMVOCs), HCHO concentrations of ~10 ppb or greater have been reported [e.g., *Zhu et al.*, 2014, and references therein]. The spatial distributions and seasonal patterns of HCHO provide valuable insights into the sources of NMVOCs (e.g., isoprene) that are important precursors of tropospheric ozone and organic aerosols.

Since the 1990s, the launch of several spaceborne UV-VIS spectrometers, including GOME (Global Ozone Monitoring Experiment), Scanning Imaging Absorption Spectrometer for Atmospheric Chartography, OMI (Ozone Monitoring Instrument), and GOME-2, has made it possible to estimate HCHO total column globally [e.g., *Chance et al.*, 2000; *De Smedt et al.,* 2008, 2012; *González Abad et al.*, 2015]. HCHO data from these sensors have helped to infer the strength of biogenic isoprene emissions from various regions [e.g., *Barkley et al.*, 2008; *Fu et al.*, 2007; *Marais et al.*, 2012; *Millet et al.*, 2008; *Palmer et al.*, 2003] and to estimate anthropogenic sources of NMVOCs [e.g., *Fu et al.*, 2007; *Zhu et al.*, 2014]. The ratio between satellite-observed HCHO and NO_2 also offers hints about the chemical regimes of ozone production [*Duncan et al.,* 2010; *Martin et al.*, 2004]. However, satellite retrievals of HCHO still suffer from relatively large noise and uncertainty owing to relatively weak signals and strong interferences. Instrument degradation over time has prompted the inclusion of instrument-related parameters in the spectral fitting in the latest algorithms in order to stabilize long-term data records [e.g., *De Smedt et al.*, 2012; *González Abad et al.*, 2015]. As effective as these approaches are, sizable differences remain between different instruments and algorithms [e.g., *De Smedt et al.*, 2012].

In this study, we introduce a new principal component analysis (PCA) spectral fitting algorithm for HCHO retrievals. It is based on an algorithm that has already been successfully implemented in the operational production of the new generation OMI SO_2 product [*Li et al.*, 2013]. Compared with the previous product, the OMI SO_2 product from the PCA algorithm improves the detection limit of anthropogenic SO_2 point sources

©2015. American Geophysical Union. All Rights Reserved.

by a factor of 2 [*Fioletov et al.*, 2015]. Here we demonstrate that the PCA algorithm can well capture the HCHO signals over source regions. We also demonstrate that the recently launched Ozone Mapping and Profiler Suite (OMPS) aboard the NASA/NOAA Suomi National Polar-orbiting Partnership (S-NPP) spacecraft is well suited for global observations of HCHO.

2. Data and Methodology

2.1. Method Overview

The PCA retrieval approach has been discussed in detail elsewhere [*Joiner et al.*, 2013; *Li et al.*, 2013] and we only provide a brief introduction here. Assuming a species of interest, g (e.g., SO_2 or HCHO), that has no or very small signals outside of certain source regions, a PCA analysis of satellite-measured radiance spectra over large background regions (e.g., the equatorial Pacific) can yield a series of spectral features (principal components, PCs). These PCs capture much of the variance in the measured radiance data, and the leading ones (which explain a very high fraction of the spectral variance) are generally associated with physical processes (e.g., ozone absorption, rotational Raman scattering, etc.) or measurement artifacts (e.g., wavelength jitter and undersampling) that often interfere with the signals of interest. For a given field of view (FOV), we can simultaneously estimate the vertical column density (VCD) of g (Ω_g) and the coefficients (ω_i) of the PCs ($v_{i,k}$), by fitting the first n_v PCs along with the Jacobian of g

$$N_k(\omega, \Omega_g) = \sum_{i=1}^{n_v} \omega_i v_{i,k} + \Omega_g \frac{\partial N_k}{\partial \Omega_g}, \tag{1}$$

where $k = 1 \ldots j$, represents the j wavelengths in the spectral range. N is the measured N value spectra for this particular FOV, defined as $N = -100 \times \log_{10}(I/I_0)$, with I and I_0 being the earthshine radiances and solar irradiances at the top of the atmosphere, respectively. The Jacobian ($\partial N_k/\partial \Omega_g$) is defined as the partial derivative of N value at the kth wavelength with respect to Ω_g. In some cases, signals from some interfering species are too weak to be readily resolved by the PCA technique and are not distinctly represented by the leading PCs. To account for these species, their cross sections can be included as additional reference spectra on the right-hand side of equation (1).

2.2. HCHO Retrievals With S-NPP/OMPS

Carrying five scientific instruments including OMPS, the S-NPP spacecraft has been orbiting the Earth in a Sun-synchronous polar orbit since October 2011. The spacecraft has an ascending node equator crossing time of ~13:30, close to the Aura spacecraft that hosts OMI. Similar to OMI, the nadir mapper (NM) of the OMPS instrument contains a 2-D CCD detector and provides near daily global coverage, although at coarser spatial ($50 \times 50\ km^2$ versus $13 \times 24\ km^2$ at nadir) and spectral (~1 nm versus ~0.5 nm) resolution in the nominal mode that is operating most of the time. A recent analysis indicates that the NM sensor has a signal-to-noise ratio of ~2000:1 or better for the wavelength range of 320–370 nm [*Seftor et al.*, 2014]. One day per week, the OMPS NM operates in a high spatial resolution mode, providing measurements at $10 \times 10\ km^2$ resolution for fewer wavelengths. In this study, we use data from the nominal mode.

For HCHO retrievals, we use the OMPS-measured radiances between 328.5 and 356.5 nm, similar to what has been used in OMI retrievals [*González Abad et al.*, 2015; *De Smedt et al.*, 2012; *Kurosu et al.*, 2004]. This spectral window was selected to minimize the effects of SO_2 absorption at shorter wavelengths and the strong absorption peak of O_2-O_2 collision complex near 360 nm. We found that extending the spectral window to include the HCHO absorption peak at ~326 nm led to comparable HCHO retrievals under most conditions, but artificial HCHO signals were produced in strong volcanic plumes due to interference by SO_2.

Following our approach developed for OMI SO_2 retrievals [*Li et al.*, 2013], we process OMPS data from its 36 rows (cross-track positions) separately to better account for the unique instrumental characteristics in each row of the CCD detector. In order to minimize the impacts of orbit-to-orbit changes in dark current and other effects, the PCA is also conducted on an orbital basis. For a given row of an orbit, we start by excluding FOVs in the descending node and those having high slant column O_3 (>2500 Dobson Unit or DU, where 1 DU = 2.69×10^{16} molecules/cm^2). The slant column O_3 (S_{O3}) is calculated from the solar zenith angle (θ_0), viewing zenith angle (θ) and O_3 VCD (Ω_{O3}) from the level 2 OMPS total O_3 product [e.g., *Kramarova et al.*, 2014] as $S_{O3} = \Omega_{O3} \times [\sec(\theta_0) + \sec(\theta)]$. This effectively removes FOVs in high latitudes during winter that are

©2015. American Geophysical Union. All Rights Reserved.

expected to have very weak HCHO signals and greater retrieval uncertainty due to reduced sensitivity. The remaining ~300 FOVs out of the ~400 FOVs in the entire row are subject to a first step PCA, and the resulting PCs are used along with a fixed HCHO Jacobian (calculated assuming midlatitude O_3 and temperature profiles, HCHO mostly in the boundary layer, $\Omega_{O3} = 325$ DU, $\theta_0 = 30°$, $\theta = 0°$, surface reflectivity = 0.05, and cloud-free conditions) to obtain an initial estimate of the HCHO VCD for each FOV (Ω_{HCHO_ini}). A second step PCA is then conducted for FOVs having relatively small Ω_{HCHO_ini} (within $\pm 1.5\sigma$ of the row/orbit) to derive updated PCs from regions with no discernable HCHO signals. We then update our estimates of Ω_{HCHO} by using the new PCs and the same Jacobian to fit the measured *N* value spectra again. In the final and third step of the process, we repeat the data filtering and PCA procedures in the second step but use HCHO Jacobians from a lookup table (see below) in the spectral fitting for the final estimates of Ω_{HCHO}. It should also be mentioned that in the final fitting we take a two-step approach first proposed by *Richter et al.* [2011]: at first we fit the radiances without including the HCHO Jacobians. We then examine the residuals to remove wavelengths with large spikes (over ±0.06 N value) before executing a second fitting that uses the HCHO Jacobians. This approach reduces noise over the South Atlantic Anomaly region but has negligible effects elsewhere.

For HCHO Jacobians, we built a lookup table from radiative transfer calculations using the Vector LInearized Discrete Ordinate Radiative Transfer (VLIDORT) model [*Spurr*, 2008]. The lookup table allows us to account for the effects of several factors on HCHO Jacobians, including the viewing geometry (θ_0, θ, and relative azimuth angle, φ), Ω_{O3}, surface reflectivity (*R*), terrain height, the effective cloud fraction (f_c), and the cloud optical centroid pressure (OCP). During data processing, we obtain Ω_{O3} from the level 2 OMPS total O_3 product, and f_c and cloud OCP from the level 2 OMPS cloud product [*Vasilkov et al.*, 2014]. For surface reflectivity, we use the same climatology derived from Total Ozone Mapping Spectrometer measurements [*Herman and Celarier*, 1997] as used in cloud retrievals. A multidimensional interpolation scheme is then used to determine the HCHO Jacobian for each FOV, based on the level 2 input data as well as monthly climatological HCHO profiles from the Global Modeling Initiative (GMI), a global chemical transport model using assimilated meteorological fields [*Rodriguez*, 1996]. For this study, GMI was run for 3 years (2005–2007) at a horizontal resolution of 2° × 2.5° for 72 vertical layers and included NMVOCs emitted from anthropogenic, biogenic, and biomass burning sources. The monthly mean HCHO profiles were generated by averaging model output sampled at the OMI overpass time (cf. Figures S1 in the supporting information for example model profiles). As discussed in section 3, GMI also provides background HCHO values that are used in our final product.

As with the PCA SO_2 retrieval algorithm, an important consideration here is the number of PCs (n_v) that should be included in the fitting. In the SO_2 algorithm, we checked the PCs for SO_2 signatures to avoid collinearity in equation (1), as even PCs obtained from remote oceanic regions may sometimes be contaminated with SO_2 signatures from strong transient volcanic plumes. For HCHO retrievals, this is less of a concern as HCHO signals are generally very weak outside of several major source regions. We conducted tests with different n_v and concluded that $n_v \geq 6$ was necessary to obtain realistic results (see Figure 1 for example PCs). We also found that using $n_v = 6$ sometimes produced negative retrievals over cloudy scenes or in high latitudes (e.g., around the Hudson Bay). Increasing n_v to 8 results in relatively small changes in the retrieved Ω_{HCHO} over known source regions but reduces negative biases over background regions. The difference in the retrieved Ω_{HCHO} between $n_v = 8$ and $n_v = 10$ is also relatively small. Next, we focus on results obtained using $n_v = 8$. As we find no distinct BrO-associated features in the first 10 PCs, we also include the BrO cross sections [*Fleischmann et al.*, 2004] as a reference spectrum in the spectral fitting to minimize its potential interference. Changes in the retrieved Ω_{HCHO} due to the inclusion of BrO in the fitting are minor.

3. Results

Figures 1a–1e show typical leading PCs extracted from OMPS radiance data in the spectral range of 328.5–356.5 nm. As with our previous analysis on OMI radiances at shorter wavelengths (310.5–340 nm) [*Li et al.*, 2013], the first PC (not shown) from OMPS data is the mean radiance spectrum and explains more than 99% of the overall variance. The second (Figure 1a) and third (Figure 1b) PCs are significantly correlated with O_3 cross sections and appear to be associated primarily with O_3 absorption and surface reflectance, respectively. A notable difference between the two analyses is that from OMPS data we were able to derive two distinct spectral features correlated with rotational Raman scattering (RRS, or the Ring effect, Figures 1c and 1d), whereas the RRS signature in the OMI data within 310.5–340 nm was embedded primarily in the third PC

©2015. American Geophysical Union. All Rights Reserved.

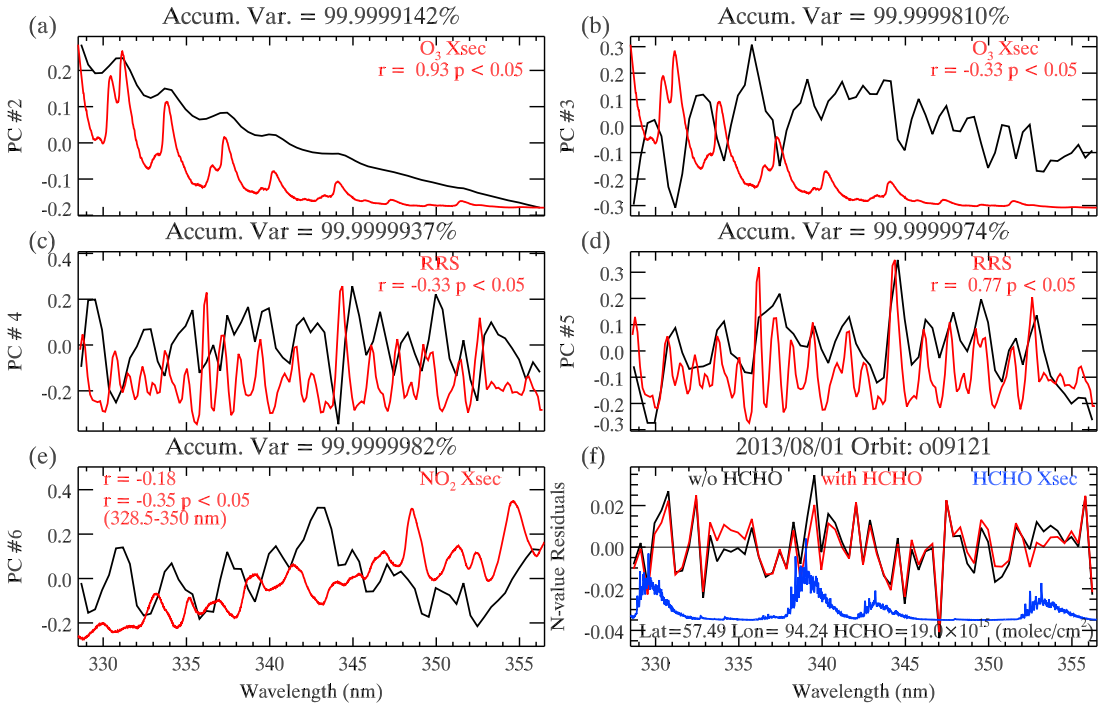

Figure 1. (a–e) Black lines: the second to sixth PCs extracted from the S-NPP OMPS radiance data from row 20 of orbit 9121 that passed over Siberia on 1 August 2013. The first PC (not shown) is essentially the mean radiance spectrum. Red lines represent O_3 cross sections at 243 K in Figures 1a and 1b, representative spectrum of rotational Raman scattering in Figures 1c and 1d, and the NO_2 cross sections at 294 K in Figure 1e. Cumulative percentage of the variance explained by the PCs and the results of the correlation analysis between PCs and reference spectra (convolved with the OMPS slit function) are given for each panel. The correlation analysis in Figure 1e was conducted after the removal of the fourth and first degree fitted polynomials from the sixth PC and the NO_2 cross sections, respectively, for both the entire spectral window ($r = -0.18$) and wavelengths < 350 nm ($r = -0.35$). (f) Fitting residuals with (red) and without (black) HCHO Jacobians included for an FOV centered at 57.49°N, 94.24°E, a Siberian location influenced by biomass burning plumes on that day. Blue line: HCHO cross sections at 298 K. All reference spectra have been scaled and shifted for more convenient comparison with the PCs and residuals.

[cf. *Li et al.*, 2013, Figure 1c]. The sixth PC from OMPS (Figure 1e) may be associated with additional trace gas absorbers such as NO_2, given the significant correlation between the two spectra at wavelengths < 350 nm. All together, just these six leading PCs account for over 99.99999% of the spectral variance contained in the OMPS-measured radiances for the row. We further examined the trailing PCs (not shown) and found no clear evidence of any of them being linked to known geophysical processes. These PCs account for only a very small fraction of the variance and may represent contributions from both measurement artifacts (e.g., wavelength shift, undersampling, etc.) and physical processes that are not resolved by PCA.

As discussed in section 2.2, for this study we use eight PCs and the BrO cross sections in the spectral fitting. As an example, Figure 1f gives the fitting residuals for a radiance spectrum acquired over Siberia on 1 August 2013. Comparing the residuals between spectral fits with and without the HCHO Jacobians included, one may notice that the inclusion of HCHO Jacobians reduces the fitting residuals at wavelengths with relatively large HCHO cross sections or sensitivity. The estimated Ω_{HCHO} from this fitting is 19×10^{15} molecules/cm^2, suggesting a substantial amount of HCHO produced from the fires that were visible in satellite images from the Aqua/Moderate Resolution Imaging Spectroradiometer instrument.

For further evidence that the OMPS PCA algorithm is capable of detecting HCHO signals, consider the monthly Ω_{HCHO} maps in Figure 2. The GMI simulation (Figure 2a) reveals several regions with active biogenic or biomass burning sources in August, including the southeast U.S., the Amazon basin, tropical Africa, Siberia, and eastern China. OMPS retrievals in Figure 2b use no a priori information from GMI and assume a fixed HCHO profile with HCHO mostly in the planetary boundary layer (PBL). The resulting retrieved Ω_{HCHO}, while smaller than model results (Figure 2b), also shows enhanced signals over the same regions. The GMI simulation used fire emissions for 2005–2007, and this explains the different locations of the Siberian HCHO hot spots between our retrievals and the model. The generally consistent spatial distributions between Figures 2a and 2b suggest that

©2015. American Geophysical Union. All Rights Reserved.

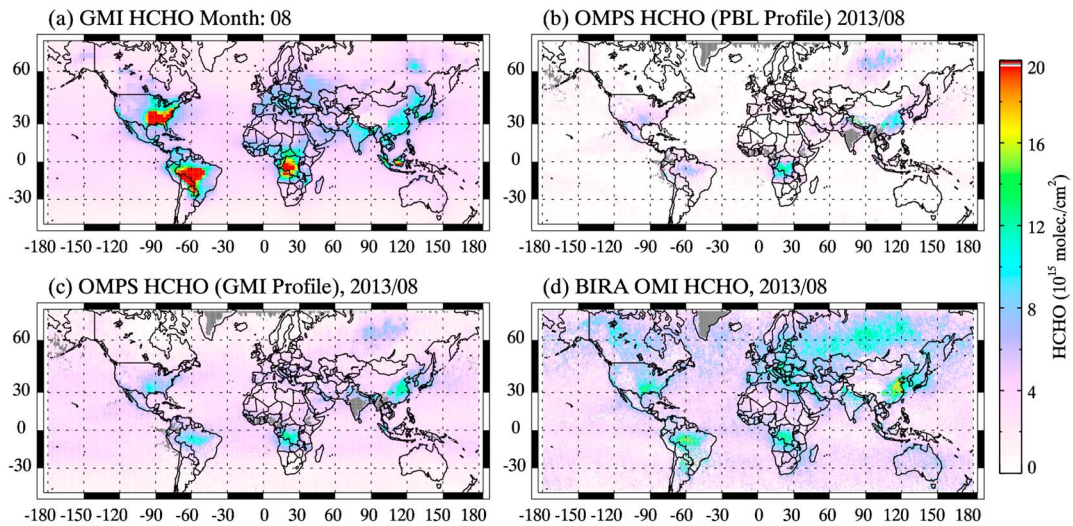

Figure 2. (a) GMI-simulated monthly mean Ω_{HCHO} for August. (b) Monthly mean Ω_{HCHO} for August 2013 retrieved from OMPS using our PCA algorithm and assuming the same a priori HCHO profile everywhere, with HCHO mostly in the planetary boundary layer. (c) The same as Figure 2b but using a priori profiles from GMI and adding the GMI-simulated Ω_{HCHO} over the background Pacific region (150–180°W) to all longitudes. (d) Monthly mean Ω_{HCHO} retrieved from OMI by the BIRA group [*De Smedt et al.*, 2012] for August 2013. OMI data are gridded at 0.25° × 0.25° resolution after excluding FOVs with cloud fraction > 0.4. OMPS data are gridded at 0.5° × 0.5° resolution after excluding FOVs in the four outermost rows or having cloud radiance fraction > 0.3. Gray-shaded grid cells in OMPS maps have less than five data points for the month.

the PCA algorithm indeed can independently capture enhanced HCHO signals. On the other hand, the PCA-retrieved Ω_{HCHO} over the remote Pacific (150–180°W) is practically zero on average. For the five latitude bands (30–60°S, 10–30°S, 10°S–10°N, 10–30°N, and 30–60°S) within the region, the ranges of daily mean Ω_{HCHO} are −0.5 to 0.6, 0.2 to 1.3, −0.4 to 1.2, −0.7 to 0.8, and 0.1 to 1.3×10^{15} molecules/cm^2, respectively, for the month. The corresponding ranges of daily standard deviation are 2.9 to 4.3, 2.6 to 3.6, 2.5 to 3.5, 2.2 to 3.0, and 2.3 to 3.3×10^{15} molecules/cm^2, respectively. This indicates that signals of background HCHO from methane oxidation may be too weak to be reliably observed. Another possibility is that, by fitting the mean spectrum (the first PC), the PCA approach effectively removes the background signals and only detects enhancements over the global background.

To account for profile effects of HCHO on the Jacobians, we use GMI HCHO profiles as a priori data in Figure 2c. To correct for the low bias over background regions, we also add monthly mean GMI Ω_{HCHO} over the remote Pacific Ocean (150–180°W, cf. Figure S2 for latitudinal dependence of this correction) to all longitudes. This effectively applies the reference sector correction technique [*Khokhar et al.*, 2005] to the PCA retrievals, making them more comparable with other algorithms that apply similar adjustments tied to chemical transport models [e.g., *De Smedt et al.*, 2012]. As can be seen from Figure 2c, the spatial pattern of OMPS HCHO with GMI model input remains similar to those in Figures 2a and 2b. It is also fairly consistent with the OMI HCHO retrievals (Figure 2d) using the Belgian Institute for Space Aeronomy (BIRA) differential optical absorption spectroscopy (DOAS) algorithm [*De Smedt et al.*, 2012]. Quantitatively, the OMPS PCA Ω_{HCHO} retrievals are about 15–20% smaller than those from BIRA OMI in several source regions during the peak season. For example, the regional mean Ω_{HCHO} for the southeast U.S. for August 2013 is $(7.4 \pm 2.3) \times 10^{15}$ molecules/cm^2 (mean ± 1σ) or ~21% smaller than BIRA OMI HCHO of $(9.3 \pm 3.1) \times 10^{15}$ molecules/cm^2. The largest relative difference is found for Siberia, where the mean OMPS PCA Ω_{HCHO} is $(5.2 \pm 1.7) \times 10^{15}$ molecules/cm^2 for the month and 56% lower than the BIRA OMI Ω_{HCHO} of $(9.3 \pm 2.1) \times 10^{15}$ molecules/cm^2. For months with lower HCHO loading, the two retrievals generally agree better in the tropics than in the middle latitudes (cf. Table S1 in the supporting information for more comparisons). This may reflect greater uncertainty in both retrievals under more challenging conditions during winter such as larger solar zenith angle and weaker signals. Despite these differences, the overall agreement between the OMPS and OMI retrievals is encouraging and should lend confidence to both products, given that they are independently produced from two different instruments using fundamentally different approaches (PCA versus DOAS). For example, OMPS has lower spatial and spectral resolution but higher signal-to-noise ratio

©2015. American Geophysical Union. All Rights Reserved.

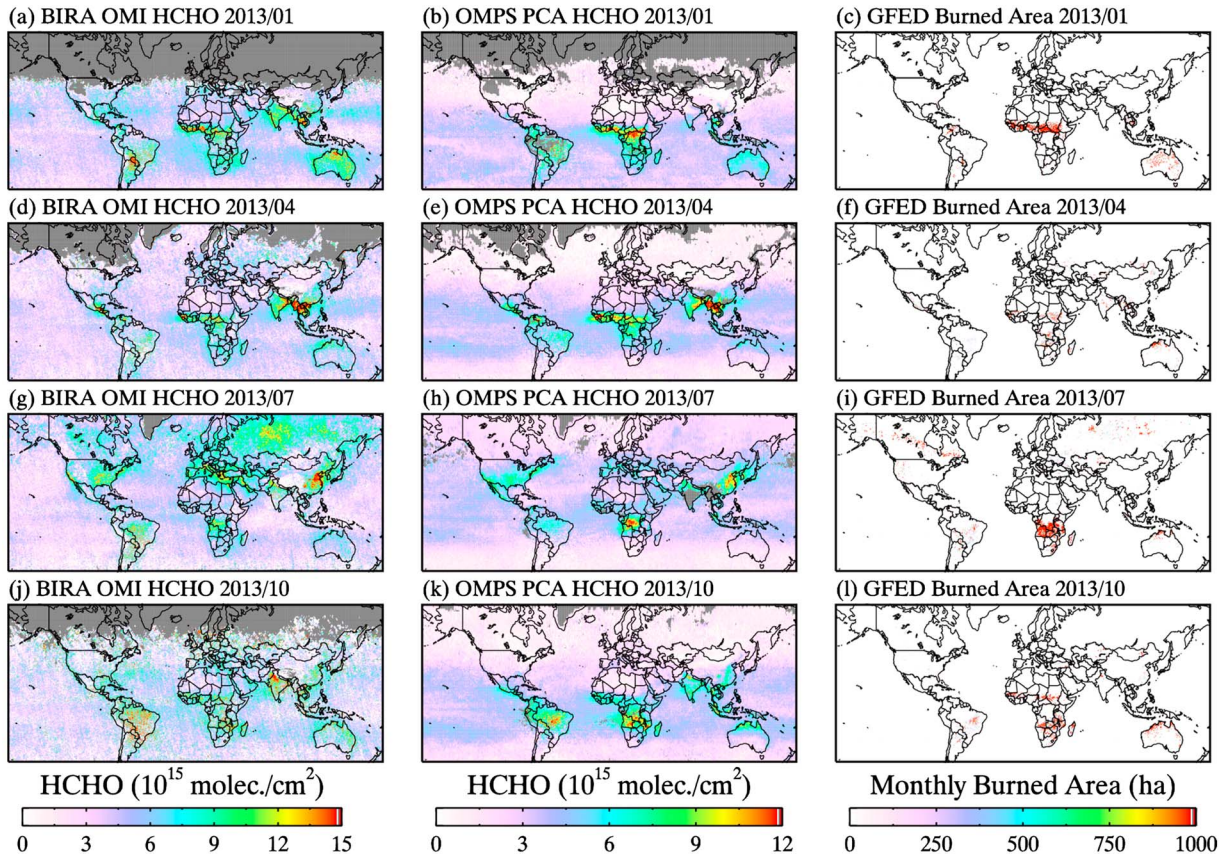

Figure 3. (a) Monthly mean OMI Ω_{HCHO} for January 2013 retrieved with the BIRA DOAS algorithm. (b) Monthly mean OMPS Ω_{HCHO} for January 2013 retrieved with our PCA algorithm. (c) GFED burnt area for January 2013. (d–l) The same as Figures 3a–3c but for April (Figures 3d–3f), July (Figures 3g–3i), and October 2013 (Figures 3j–3l). Note the different color scales for BIRA OMI HCHO (Figures 3a, 3d, 3g, and 3j) and OMPS PCA HCHO (Figures 3b, 3e, 3h, and 3k). The data filtering criteria and gridding resolution are the same as in Figure 2.

than OMI. The input data are also different between the two algorithms, including the a priori HCHO profiles, surface reflectivity, cloud data, and also background HCHO corrections. More detailed analyses will be necessary to reconcile the quantitative differences seen between the two data sets. We also compared our HCHO retrievals with those from the new Smithsonian Astrophysical Observatory (SAO) OMI product [*González Abad et al.*, 2015]. The HCHO burden in SAO retrievals is greater than both OMPS PCA and BIRA OMI retrievals, while the spatial patterns are mostly consistent between the three (Figure S4). Our preliminary OMI PCA HCHO retrievals also show consistent spatial distribution with OMPS (cf. section 4).

Figure 3 compares the seasonal patterns of BIRA OMI and OMPS PCA HCHO. Also presented in the figure is the estimated burnt area in the Global Fire Emission Database (GFED) as an indicator of biomass burning activities [*Giglio et al.*, 2013]. In general, the two HCHO retrievals demonstrate similar seasonal changes, and fire emissions appear to be an important process driving these changes. Examples include seasonal fires in Australia in January and October (Figures 3c and 3l), in Sahel in January (Figure 3c), in central and southern Africa in July and October (Figures 3i and 3l), in Southeast Asia in April (Figure 3f), and in Siberia in July (Figure 3i). Stronger biogenic emissions and photochemical processes likely contribute to the greater HCHO loading over the southeast U.S. in July, whereas anthropogenic sources likely also play a significant role in eastern China. More detailed regional maps of OMPS HCHO and GFED burnt area are available in the supporting information (Figures S5–S11).

4. Discussion

From the noise in retrievals over the remote Pacific using a fixed PBL HCHO profile (~3×10^{15} molecules/cm^2, cf. section 3), we estimate that the detection limit of the PCA retrieval approach is ~12×10^{15} molecules/cm^2

©2015. American Geophysical Union. All Rights Reserved.

for an instantaneous FOV (signal-to-noise ratio 2:1 for $\pm 1\sigma$ noise). This is comparable to the detection limit of 10×10^{15} molecules/cm^2 estimated by *González Abad et al.* [2015] for their OMI retrievals based on fitting residuals. It is also much greater than the expected signal level over clean background regions, indicating that the PCA approach limits the possibility of detecting background HCHO. For typical Ω_{HCHO} values of 6 to 15×10^{15} molecules/cm^2 over polluted regions, the retrieval noise translates into relative uncertainty of 40–100% on a single FOV basis. Additionally, there is about 20–50% relative uncertainty in HCHO Jacobians, mainly attributed to errors in HCHO profiles (up to 40%) and cloud pressure (20–30%, see Figures S12–S14 for examples), as well as interpolation error (up to ~10% for large θ_0 and θ). The overall typical retrieval uncertainty is estimated at between 45% and 110%, similar to the 45–105% uncertainty range given by *González Abad et al.* [2015] based on estimated uncertainty in fitting and AMF (air mass factor). A more comprehensive, rigorous error analysis of the OMPS PCA retrievals requires thorough understanding of the instrument performance and will be the subject of future studies.

The comparable detection limit and level of uncertainty between the two approaches should not come as a surprise, given that both methods attempt to exploit the same spectral content in order to minimize the effects of the same interfering physical processes (e.g., absorption by O_3 and BrO and rotational Raman scattering). One unique feature of the PCA approach, however, is that it does not require explicit corrections to the measured radiances. Unlike the DOAS approach [e.g., *De Smedt et al.,* 2012; *González Abad et al.*, 2015], no instrument-related parameters are explicitly included in our spectral fitting. This facilitates the production of consistent, long-term data records spanning different satellite missions. Indeed, preliminary test demonstrates that the PCA algorithm can produce consistent HCHO retrievals between OMI and OMPS (cf. Figure S15 for preliminary OMI results).

In terms of processing time, it takes about 10 min to process an entire OMPS orbit using a single state-of-the-art CPU, with about 95% of the time dedicated to the interpolation of HCHO Jacobians at the many wavelengths used in the retrieval. The processing time can be reduced by several times by using a DOAS-like approach of first fitting the HCHO cross sections to derive the slant column amount and then calculating the AMF for just a few wavelengths, given the relatively small spectral dependence of AMF in the fitting window.

5. Conclusions

In this study, we apply a PCA spectral fitting algorithm to HCHO retrievals with the S-NPP/OMPS nadir mapper. We demonstrate that the PCA technique is capable of extracting various spectral features (i.e., PCs) that are associated with both physical processes (e.g., O_3 absorption and RRS) and measurement details (e.g., wavelength shift) from OMPS radiance data. Use of these PCs derived directly from the measurement data in the spectral fitting allows us to accurately account for these potential interferences in HCHO retrievals, leading to retrievals with relatively small noise and artifacts. While our PCA algorithm finds no distinct HCHO signals over background oceanic areas, it is able to detect enhanced signals over source regions such as the southeast U.S., the Amazon basin, and tropical Africa. This is evidenced by the generally consistent spatial patterns between modeled and OMPS-retrieved HCHO. In comparison with DOAS HCHO retrievals applied to the Aura/OMI sensor [*De Smedt et al.*, 2012], the OMPS PCA retrievals are on average about 15–20% smaller in most source regions (e.g., eastern China and the southeast U.S.). Despite this difference, probably attributable to instrumental and algorithmic factors, the two retrievals show similar spatial distributions and consistent seasonal changes. There is also good agreement in the spatial patterns between the OMPS HCHO and GFED fire products particularly for Australia, South Asia, Southeast Asia, and tropical Africa, all areas with active seasonal fires. Overall, the PCA algorithm shows great potential as a new technique for global observations of HCHO. The OMPS nadir mapper, despite having lower spectral and spatial resolution as compared with OMI, appears to be suitable for HCHO retrievals and for continuing the long-term global HCHO data record.

Acknowledgments
We thank the NASA GSFC GMI group for GMI HCHO model results and two anonymous reviewers for comments that helped improve the manuscript. BIRA OMI HCHO data were downloaded from http://h2co.aeronomy.be. SAO OMI HCHO data were obtained from http://disc.sci.gsfc.nasa.gov. GFED data were downloaded from http://www.globalfiredata.org. OMPS L1B and L2 O_3 and cloud research products are available at https://ozoneaq.gsfc.nasa.gov. L.D. acknowledges graduate research assistantship partially supported by the NASA Earth Science New Investigator Program (grant NNX14AI02G). OMPS HCHO retrieval results are available upon request from the corresponding author.

The Editor thanks two anonymous reviewers for their assistance in evaluating this paper.

References

Barkley, M. P., P. I. Palmer, U. Kuhn, J. Kesselmeier, K. Chance, T. P. Kurosu, R. V. Martin, D. Helmig, and A. Guenther (2008), Net ecosystem fluxes of isoprene over tropical South America inferred from Global Ozone Monitoring Experiment (GOME) observations of HCHO columns, *J. Geophys. Res.*, *113*, D20304, doi:10.1029/2008JD009863.

Chance, K., P. I. Palmer, R. J. D. Spurr, R. V. Martin, T. P. Kurosu, and D. J. Jacob (2000), Satellite observations of formaldehyde over North America from GOME, *Geophys. Res. Lett.*, *27*, 3461–3464, doi:10.1029/2000GL011857.

©2015. American Geophysical Union. All Rights Reserved.

De Smedt, I., J. F. Műller, T. Stavrakou, R. van der A, H. Eskes, and M. Van Roozendael (2008), Twelve years of global observations of formaldehyde in the troposphere using GOME and SCIAMACHY sensors, *Atmos. Chem. Phys.*, *8*, 4947–4963, doi:10.5194/acp-8-4947-2008.

De Smedt, I., M. Van Roozendael, T. Stavrakou, J.-F. Müller, C. Lerot, N. Theys, P. Valks, N. Hao, and R. J. van der A (2012), Improved retrieval of global tropospheric formaldehyde columns from GOME-2/MetOp-A addressing noise reduction and instrumental degradation issues, *Atmos. Meas. Tech.*, *5*, 2933–2949.

Duncan, B. N., et al. (2010), Application of OMI observations to a space-based indicator of NO_x and VOC controls on surface ozone formation, *Atmos. Environ.*, *44*, 2213–2223.

Fioletov, V. E., C. A. McLinden, N. Krotkov, and C. Li (2015), Lifetimes and emissions of SO_2 from point sources estimated from OMI, *Geophys. Res. Lett.*, *42*, doi:10.1002/2015GL063148, in press.

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. Photochem. Photobiol., A*, *168*, 117–132.

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, doi:10.1029/2006JD007853.

Giglio, L., J. T. Randerson, and G. R. van der Werf (2013), Analysis of daily, monthly, and annual burned area using the fourth-generation Global Fire Emissions Database (GFED4), *J. Geophys. Res. Biogeosci.*, *118*, 317–328, doi:10.1002/jgrg.20042.

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*, 19–32.

Herman, J. R., and E. A. Celarier (1997), Earth surface reflectivity climatology at 340–380 nm from TOMS data, *J. Geophys. Res.*, *102*(D23), 28,003–28,011, doi:10.1029/97JD02074.

Joiner, J., L. Guanter, R. Lindstrot, M. Voigt, A. P. Vasilkov, E. M. Middleton, K. F. Huemmrich, Y. Yoshida, and C. Frankenberg (2013), Global monitoring of terrestrial chlorophyll fluorescence from moderate-spectral-resolution near-infrared satellite measurements: Methodology, simulations, and application to GOME-2, *Atmos. Meas. Tech.*, *6*, 2803–2823, doi:10.5194/amt-6-2803-2013.

Khokhar, M. F., C. Frankenberg, M. Van Roozendael, S. Beirle, S. Kühl, A. Richter, U. Platt, and T. Wagner (2005), Satellite observations of atmospheric SO_2 from volcanic eruptions during the time-period of 1996–2002, *Adv. Space Res.*, *36*, 879–887.

Kramarova, N. A., E. R. Nash, P. A. Newman, P. K. Bhartia, R. D. McPeters, D. F. Rault, C. J. Seftor, P. Q. Xu, and G. J. Labow (2014), Measuring the Antarctic ozone hole with the new Ozone Mapping and Profiler Suite (OMPS), *Atmos. Chem. Phys.*, *14*, 2353–2361.

Kurosu, T. P., K. Chance, and C. E. Sioris (2004), Preliminary results for HCHO and BrO from the EOS-Aura Ozone Monitoring Instrument, *Proc. Soc. Photo Opt. Instrum. Eng.*, *5652*, 116–123, doi:10.1117/12.578606.

Li, C., J. Joiner, N. A. Krotkov, and P. K. Bhartia (2013), A fast and sensitive new satellite SO_2 retrieval algorithm based on principal component analysis: Application to the ozone monitoring instrument, *Geophys. Res. Lett.*, *40*, 6314–6318, doi:10.1002/2013GL058134.

Marais, E. A., et al. (2012), Isoprene emissions in Africa inferred from OMI observations of formaldehyde columns, *Atmos. Chem. Phys.*, *12*, 6219–35.

Martin, R. V., A. M. Fiore, and A. Van Donkelaar (2004), Space-based diagnosis of surface ozone sensitivity to anthropogenic emissions, *Geophys. Res. Lett.*, *31*, L06120, doi:10.1029/2004GL019416.

Millet, D. B., D. J. Jacob, K. F. Boersma, T. M. Fu, T. P. Kurosu, K. Chance, C. L. Heald, and A. Guenther (2008), Spatial distribution of isoprene emissions from North America derived from formaldehyde column measurements by the OMI satellite sensor, *J. Geophys. Res.*, *113*, D02307, doi:10.1029/2007JD008950.

Palmer, P. I., D. J. Jacob, A. M. Fiore, R. V. Martin, K. Chance, and T. P. Kurosu (2003), Mapping isoprene emissions over North America using formaldehyde column observations from space, *J. Geophys. Res.*, *108*, 4180, doi:10.1029/2002JD002153.

Richter, A., M. Begoin, A. Hilboll, and J. P. Burrows (2011), An improved NO_2 retrieval for the GOME-2 satellite instrument, *Atmos. Meas. Tech.*, *4*, 1147–1159, doi:10.5194/amt-4-1147-2011.

Rodriguez, J. M. (1996), Chapter 6 in atmospheric effects of aviation: First report of the subsonic assessment project, in *Global Modeling Initiative*, NASA Reference Publication, vol. 1385, edited by A. M. Thompson, R. R. Friedl, and H. L. Wesoky, National Aeronautics and Space Administration, Washington, D. C.

Seftor, C. J., G. Jaross, M. Kowitt, M. Haken, J. Li, and L. E. Flynn (2014), Postlaunch performance of the Suomi National Polar-orbiting Partnership Ozone Mapping and Profiler Suite (OMPS) nadir sensors, *J. Geophys. Res. Atmos.*, *119*, 4413–4428, doi:10.1002/2013JD020472.

Shim, C., Y. Wang, Y. Choi, P. I. Palmer, D. S. Abbot, and K. Chance (2005), Constraining global isoprene emissions with Global Ozone Monitoring Experiment (GOME) formaldehyde column measurements, *J. Geophys. Res.*, *110*, D24301, doi:10.1029/2004JD005629.

Singh, H., 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*, 1078–1081.

Spurr, R. (2008), *LIDORT and VLIDORT: Linearized Pseudo-Spherical Scalar and Vector Discrete Ordinate Radiative Transfer Models for use in Remote Sensing Retrieval Problems*, Light Scattering Reviews, vol. 3, edited by A. Kokhanovsky, Springer, Berlin, doi:10.1007/978-3-540-48546-9.

Vasilkov, A., J. Joiner, and C. Seftor (2014), First results from a rotational Raman scattering cloud algorithm applied to the Suomi National Polar-orbiting Partnership (NPP) Ozone Mapping and Profiler Suite (OMPS) nadir mapper, *Atmos. Meas. Tech.*, *7*, 2897–2906.

Zhu, L., D. J. Jacob, L. J. Mickley, E. A. Marais, D. S. Cohan, Y. Yoshida, B. N. Duncan, G. González Abad, and K. V. Chance (2014), Anthropogenic emissions of highly reactive volatile organic compounds in eastern Texas inferred from oversampling of satellite (OMI) measurements of HCHO columns, *Environ. Res. Lett.*, *9*, 114004, doi:10.1088/1748-9326/9/11/114004.

©2015. American Geophysical Union. All Rights Reserved.