Atmos. Chem. Phys. Discuss., 11, C2608–C2618, 2011
www.atmos-chem-phys-discuss.net/11/C2608/2011/
© Author(s) 2011. This work is distributed under
the Creative Commons Attribute 3.0 License.

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Evaluation of satellite-derived HCHO using statistical methods" *by* J. H. Kim et al.

J. H. Kim et al.

lihuawang@nsstc.uah.edu

Received and published: 3 May 2011

We believe that the reviewer provided valuable comments that are enable to improve this paper. We answered all reviewers' questions and comments.

Reviewer's question-1-a: Barkley et al. (GRL, vol. 36, L04803, 2009) presented a very similar analysis based on GOME and SCIAMACHY HCHO data sets. The added-value of this paper compared to the Barkley's paper is not really clear. In any case, the authors should compare and discuss their results to those from Barkley. In particular, Barkley et al. attribute part of the HCHO signal in Amazonia to biogenic sources on contrary to the present study.

Answer: The Barkley et al (2009) paper is about analyses of GOME and SCIAMACHY

HCHO, not evaluation of the data sets. In that paper, they concluded biogenic activity is the main source of HCHO over the globe. Especially, they focused on finding how the HCHO seasonality is influenced only by biogenic activity over Amazon rainforest. To do this, they removed biomass burning influence by using fire counts and satellite NO2 column measurements. The reviewer indicated that this result is different from what we found that biomass burning is the main driving mechanism of HCHO.

The EOF analysis is to show the variability of a species after removing a mean value. Over the globe, we know that the main source of HCHO as well as its variability is biogenic activity. Because of this, EOF mode 1 shows the dominant signal coming from biogenic activity. However, over the South America, there are two main HCHO sources; biogenic and biomass burning activity. No matter which is the largest source for HCHO, EOF mode-1 will pick up the signal from the source with the strongest variability. Our analysis shows the strongest variability of HCHO over South America coming from biomass burning activity. This is not different from Barkley et al (2009). As the reviewer indicated, this part was not clearly explained. We modified this part and discussed our results to those from Barkley (2009) as well as we referenced Barkley et al. (2009).

Reviewer's question-1-b: Also, they use the three first modes to explain most of the HCHO variability. In the present manuscript, the authors only use the first mode which represents a small fraction of the GOME and SCIA HCHO variability. Can the following modes provide useful information?

The following figures (Figure 1, 2 and 3) are of the EOF mode-2 of GOME, SCIAMACHY, and OMI. OMI only explains negligible amounts of variance, 10%, relative to mode-1, 56%. There is a pattern along the east-west direction, but we don't find any source that matches this pattern. Because mode-2 is an orthogonal Eigenvector to mode-1, the pattern in mode-2 could be just the mathematical derivative. The second figure is the mode-2 of GOME, with variance of 7%. The north-south stripes are identified. This is in accordance with satellite track most likely due to occur from low spatial and temporal coverage of GOME. The third is the mode-2 of SCIAMACHY. No meaningful signal is observed.

When we added the analyses of all modes, the paper became too long. These are the reasons why we only discuss mode-1.

Reviewer's question-2-a: The title of the manuscript and the abstract are misleading. They suggest that a validation of various HCHO data sets based on a novel method is realized. To me, the statistical tools presented in this work are used to interpret seasonalities and spatial patterns of HCHO and to establish links with possible sources. Very little is done in terms of intercomparison of the different data sets. The fact that the various data sets show roughly the same structures in the EOF mode 1 is too qualitative to be considered as validation. In addition, interpretation of the features of the EOF mode 1 is missing for the different data sets. For example, why does this mode represent 50% of the OMI variability and only 20% of that of GOME and SCIA?

Answer: We would like to point out that this paper is not about the analyses, but the evaluation of satellite data analyses. Most papers analyze satellite data and draw conclusions under the assumption that the data are flawless. However, satellite data have many errors. This paper uses a new statistical approach not to analyze, but evaluate the satellite data by examination of consistency between our understanding of atmospheric chemistry and physics, and various satellite data. According to this kind of evaluation study, we believe there have been the improvements in satellite trace gases retrievals. This paper is a further study about satellite data evaluation that was originally started by Kim et al. (2008, JGR): Singular value decomposition analyses of tropical tropospheric ozone determined from TOMS.

The results from EOF analysis can be different depending on the grid-size of data sampling, the spatial and temporal resolution of the satellite data, and the study area, e.g., global or local. For example, OMI has much higher spatial and temporal resolution than GOME and SCIAMACHY. This is the reason why OMI shows stronger variability than the others.

Reviewer's question-2-b: Why is the GOME EOF mode 1 signal much weaker compared to the other instruments? Why is the amplitude of the corresponding expansion coefficient much weaker? In the Barkley's paper, the time series of the principal component of mode 1 shows only a small discontinuity when SCIAMACHY replaces GOME.

Answer: Barkley et al.(2009) used the merged two satellite data sets between GOME and SCIAMACHY covering from 1996-2008, while we used an individual data set. If you used longer period of data set, you can get better and clearer signal in variance and corresponding expansion coefficient than the data set with shorter period. For our case, we included GOME data up to the year of 2003 when GOME data suffered from instrument degradation. Therefore, the discontinuity between our work and Barkley's is due to data in the year 2003. As we said before, because the purpose of this paper is to evaluate satellite data, we included available data. For our case, we normalized data sets to identify the variability. This does not affect the results.

Reviewer's question-3: The authors compare the spatial and temporal variations of CO and HCHO observations to those of fire counts from ATSR. There is no reference for this product. Since the spatial and temporal patterns of the fire count product is largely used and discussed by the authors, a figure illustrating these features should be added.

Answer: We referenced ATSR data and its location on line 15 of page 8010, "Along-Track Scanning Radiometer (ATSR) (http://www.atsr.rl.ac.uk/)". Because it is widely used, we did not include it. However, because of reviewer's comment, we added the EOF mode-1 in the revised version that shows fire location and seasonality in Figure 2 and Figure 4.

Reviewer's question-4: It is not clear how the longer CO lifetime can explain the lag between the peak in the ATSR fire count product and the maximum in the CO observations. Since CO is directly emitted by fires, the two maxima should be in phase whatever the CO lifetime. This lag could be caused by fires emitting CO which are undetected by ATSR. The seasonality in the fire count product could be compared to the GFED inventory. Another explanation for this lag could be that the maximum in the CO observations is partly due to alternative sources. A large part (_50% at the global scale) of CO originates from oxidation of CH4 and NMVOCs (e.g. Hooghiemstra et al., ACPD, 11, 341–386, 2011). The authors do not discuss this secondary source and establish a direct link between fires and CO maximum without any convincing argument for the lag.

Answer: The sources of CO could be something other than biomass burning as the reviewer indicated. The reviewer suggested that the other possible sources are undetected by ATSR: Oxidation of CH4 and NMVOCs could be the sources of CO. But, as Barkley et al. (2009) showed, the background HCHO produced by the oxidation of non-isoprene does not have seasonality. Because CO can be the product from oxidation of hydrocarbon, we expect that CO seasonality entirely due to hydrocarbon oxidation must be marginal. As we have clearly stated, EOF and SVD shows the variability of sources, not amounts of sources. Even though ATSR might not observe some of the fires, the outcome of the EOF will not be changed assuming the fire variability is stronger than biomass burning variability. Therefore, mode-1 of CO variability must come from biomass burning activity. When we analyzed correlation among NO2, ozone, and CO over Africa, we found a similar lag among NO2, ozone, and CO (Kim et al., JGR, 2008). In the paper, we found the maximum of NO2 and ozone always occurred one or two months earlier than CO. The lifetime of these species was the most likely explanation for the lag as we discussed in this paper.

Reviewer's question-5: The same comment is true for HCHO. In principle, a maximum in the HCHO observations should be detected when the biomass burning activity is the largest. For example, in the auxiliary material of the Barkley's paper, a figure shows the clear temporal correlation between fire counts and high HCHO columns. The lag between ATSR fire count and HCHO maxima can't be explained by the HCHO lifetime as claimed by the authors.

Answer: Over Africa, we expect to have maximum HCHO in January when biomass

burning activity is the strongest over north central equatorial Africa, where more intensive burning takes place, because of a shorter lifetime of HCHO relative to CO. Conversely, the observations revealed that the maximum HCHO was found further downwind from where maximum CO was observed. In addition, the fire counts show that burning is the highest in January during the northern burning season (http://www.atsr.rl.ac.uk/); whereas, HCHO shows a temporary minimum in January between two peaks in December and March. These spatial and temporal patterns are different from those seen over southern tropical Africa during the southern burning season.

As we indicated in the text, a significant difference between SCIAMACHY, GOME and OMI HCHO comes from the seasonality observed in the northern tropical region during the northern biomass burning season. The double-peak feature seen in OMI HCHO seasonality is marginally observed in SCIAMACHY and GOME HCHO seasonality. The difference in HCHO seasonality between datasets could be caused by the difference in spatial and temporal resolution between the instruments. However, this explanation does not seem to be the likely cause because the difference in seasonality is observed over a large area and for many years. Analyzing climatological wind and precipitation over this region reveals that the wind is northeasterly and the ITCZ is located in south of the equator during the boreal winter. No special weather pattern was observed over northern tropical Africa in January (http://iridl.ldeo.columbia.edu/maproom/.Regional/.Africa). Another possibility for the cause could be the difference in equator crossing time between OMI (01:45 p.m.), and SCIAMACHY (10:00 a.m.). (http://www.knmi.nl; http://envisat.esa.int). The diurnal cycle of biomass-burning activity and HCHO-related chemistry may be the cause of the difference (Benning and Wahner, 1998; Palmer et al., 2007; Stavrakou et al., 2009a).

We found some discrepancies among satellite data as well as between satellite data and atmospheric chemistry. Again, this is not analysis of satellite data, but evaluation of the data. The detailed chemical analysis is beyond the scope of this study.

Reviewer's question-6: In the "introduction" and "data" sections, the authors suggest that the same algorithm is applied to OMI, GOME and SCIAMACHY data. To my knowledge, the data provided on the "temis" and "mirador.gsfc.nasa" websites are not retrieved in the same way. A description of the GOME and SCIA products is probably missing. On the TEMIS website is provided a GOME-2 data set. This instrument has a better spatial coverage than GOME and SCIAMACHY and could have been considered in this study. Also, the MOPITT CO product is not described and there is no reference for it.

Answer: The HCHO from satellite is retrieved by the concept described between line 24 on page 8005 and line 7 on page 8006, as well as between line 6 and 13 on page 8009. Dr. Kurosu and Dr. I. De Smedt, co-authors on this paper, implement HCHO retrieval from OMI, and GOME and SCIAMACHY, respectively. Thus, we think the description of satellite HCHO retrievals is good enough.

Reviewer's question-7: The discussion about the retrieval error sources in the introduction should be clarified and moved to the next section. The description of the statistical methods should be extended. In particular, the differences between the EOF and SVD analyses should be further explained, especially because the contribution of the results from the SVD analysis is not clear in the discussion, these results appearing very similar to those.

Answer: This appears to be relocation of some of contents. Because this paper is about the evaluation of satellite data, we prefer to put the retrieval error in the INTRODUCTION to emphasize the purpose as well as goal of this paper. However, because the reviewer suggested to describe the EOF and SVD, we explain those in detail and add more references . Reviewer's question-8: The EOF analysis allows detecting area with high variance compared to a mean value. Possible spatial and temporal correlations of the EOF signal for HCHO or CO with biomass burning activity does not mean that biomass burning is the strongest source of the corresponding specie but that biomass burning is the main cause for its variability. The authors should be more

cautious in the formulation they use for the conclusions they draw from their analyses. In particular, their conclusions are in opposition to several studies showing that biogenic emissions are an important source for HCHO production in tropical regions (e.g. Stavrakou et al., ACP, 9, 1037–1060, 2009; Guenther et al., Atmos. Chem. Phys., 6, 3181–3210, 2006).

Answer: As is explained, the EOF analysis allows detecting areas with high variance after removing a mean value. The SVD examines the coupled variability of two fields. Each pair of singular vectors describes a fraction of the square covariance (SCF) between the two variables.

As we pointed out in the second answer above, many scientists use the satellite data under the assumption that the data are flawless. However, satellite data have many errors. This paper is not about the analyses of satellite data, but evaluation of satellite data. Our results are not quite different from the works that that were listed by the reviewer.

We agree that there are some of explanations were not clear and somewhat confused. We corrected these in the revised version.

Interactive comment on Atmos. Chem. Phys. Discuss., 11, 8003, 2011.

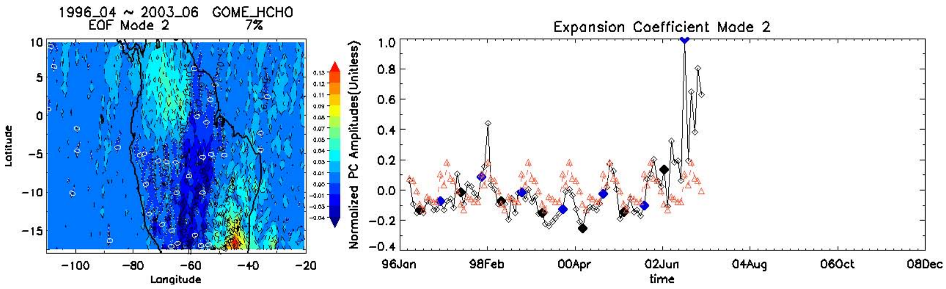

Fig. 1. EOF mode-2 of GOME

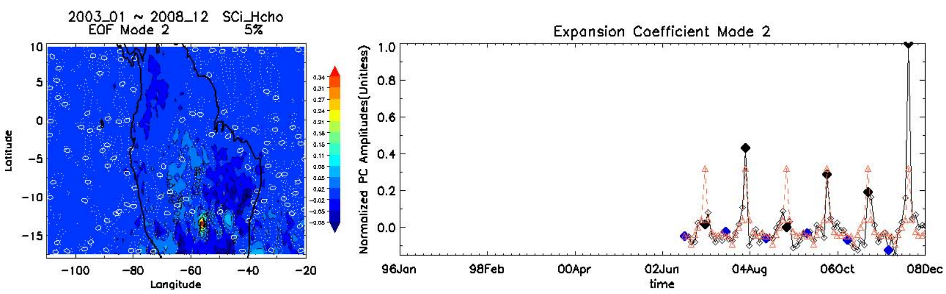

Fig. 2. EOF mode-2 of SCIAMACHY

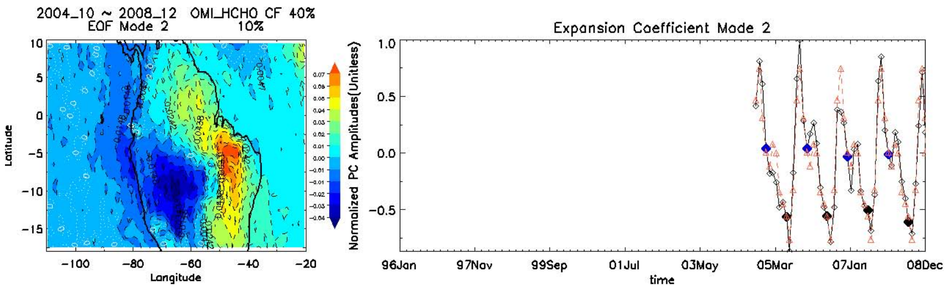

Fig. 3. EOF mode-2 of OMI