Reply to Anonymous Referee #2

First, we would like to thank Anonymous Referee #2 for his/her helpful comments.

In the manuscript "Four years of ground-based MAX-DOAS observations of HONO and NO_2 in the Beijing area" written by Hendrick et al. long term observations of HONO and NO_2 in Beijing and Xianghe are presented. Some interesting details about the MAX-DOAS retrieval are given and generally the paper is written in a clear and well structured manner. The variations of HONO, NO_2 VCDs and ground concentrations, and their influence on OH production rates are analyzed on different time scales. Since the topic is relevant for ACP I recommend publications after some revisions.

General comments:
There are almost no validation studies presented in this paper. Especially for the novel HONO retrieval the results should be compared to other data sets. The only comparison mentioned is the comparison with the diurnal cycles of NO_2 with similar measurements by Ma et al (2013) and they don't seem to agree well. A comparison of the NO_2 column densities with satellite derived VCDs might be very useful. Regarding the HONO results a comparison with chemical transport model output could also share some light about the uncertainties. The authors should try to find some independent data set to validate their results in order to show that the observed patterns and estimated uncertainties are reasonable.

We agree on this general comment about the lack of validation in the ACPD version of the manuscript. In the revised version, we have added comparisons between our retrieved AODs and sunphotometer data and between our retrieved NO_2 VCDs and Ma et al. (2013) data. In both cases, the agreement with correlative data is very good, even with Ma et al. (see our reply to the specific comment **'**P10635, Sec.3.2' here below). In the case of HONO, the problem is that, to our knowledge, co-located data at Beijing and Xianghe stations are not available. However, we have added in the revised manuscript a comparison with other daytime HONO measurements performed in or in the vicinity of megacities in East Asia. A similar comparison has been already done by Li et al. (2012) but in that case, all published HONO measurements in Asia and in the rest of the world were compared. As can be seen in Table 4 of the revised manuscript, the agreement with our HONO data is found to be relatively good, suggesting that our HONO retrieval gives reasonable results. Comparing our HONO and NO_2 observations with 3D-CTM simulations could be surely valuable, but it would be beyond the scope of this first study. The main purpose here is to describe and characterize the first measurements of HONO by the MAX-DOAS technique in or in the vicinity of a megacity and to briefly discuss the observed features. The

data set presented here can certainly be used to constrain further HONO modeling studies, as suggested in the conclusions of the paper.

Specific comments:

P10625, L15f: The sentence reads like you are saying that MAX-DOAS has a higher sensitivity to absorbers close to the ground than LP-DOAS, is that what you meant? Maybe you mean compared to other scattered sunlight DOAS techniques (e.g. ZSLDOAS)? Please clarify

We meant that MAX-DOAS has a higher sensitivity to absorbers close to the ground compared to other scattered sunlight DOAS technique. We have clarified this point in the revised manuscript.

P10626, L21f and P10628, L15f: First you mention that the MAX-DOAS instrument is described in Clemer et al (2010), but in Clemer et al (2010) four wavelength regions (360, 477, 577, and 630 nm) are used for the O4 retrieval, whereas you are using only two wavelength regions (around 360 and 477). The uncertainties of the DOAS retrievals at larger wavelengths might increase but the results still contain useful information. Please explain why you are using a different retrieval algorithm (if you do). If your algorithm differs from the one described by Clemer et al., you should provide more details about the aerosol profile retrieval and show some examples.

We used the Clemer et al. (2010) algorithm but instead of retrieving aerosol extinction profiles at four wavelengths (360, 477, 577, and 630 nm), we performed two independent aerosol retrievals: one at 360 nm and one at 477 nm (wavelengths used for the retrieval of HONO and NO_2, respectively). We mention this point in the revised manuscript and, as also requested by Referee #1, we have added aerosol retrieval results (retrieval example as well as AOD diurnal and seasonal cycles).

P10627, L2f and L11: You write "Scattered light is collected at various elevation and azimuth angles. . .", but the azimuth direction is fixed (L11). Maybe you mean something else, so please clarify.

Our instrument is able to perform azimuthal scans but at Beijing and Xianghe, this option was not activated and measurements have been performed at a fixed azimuth angle corresponding to the north direction. We have clarified this point in the revised manuscript.

P10629, Eq.1: Did you check whether this profile shape is appropriate for an urban environment? Are the deviations shown in Fig. 2 and 3 typical or do they average out over time? If they do not average out, you could use an average profile shape scaled by the VCD of the 30 deg scan as the a priori. Did you try alternatives? Another option would be to use the last accepted profile shape of the time series as the a priori for the next inversion. Please justify why you chose this profile shape.

To our knowledge, there were only a few attempts in the past to measure the HONO vertical distribution in big cities (e.g. Villena et al., 2011; Wong et al., 2012). Moreover, these measurements were limited to two altitude levels (ground + top of a building, so roughly the 0-200m range), which is certainly a great achievement but not sufficient to provide us with vertical profiles which can be used as a priori in our retrievals. Models could also provide information on the vertical distribution of HONO. However, as mentioned in Elshorbany et al. (2012), realistic simulations of HONO levels are difficult due to the fact that the HONO photochemistry is still not well known and that each region/city requires a specific parameterization. Because of the lack of relevant existing data sets, we decided to follow the pragmatic approach of using an exponentially decreasing profile shape with a fixed scaling height (SH=0.5 km) for both HONO and NO_2 retrievals. Such an approach has been already used in the past in the absence of correlative data for the a priori profiles (see e.g. Friess et al., 2011). It should also be noted that the impact of the a priori on the retrieved HONO and NO_2 surface concentrations and VCDs has been evaluated by using a SH of 1 km instead of 0.5 km. We obtained the following impacts on the retrieved quantities: ~20% on HONO VCD, ~10% on HONO surface concentration, ~10% on NO_2 VCD, and ~12% on NO_2 surface concentration. These uncertainties related to the choice of the a priori profiles were included in the error budget.

P10630, L3: The geometric approximation is based solely on Rayleigh scattering, so if there are large amounts of aerosol present and hence a high degree of Mie scattering, this approximation becomes inaccurate. However, you have this information from the first part of the MAX-DOAS inversion, so why not use it here?

We agree that using the geometric approximation in presence of large aerosol loads can lead to inaccuracies. Ma et al. (2013) compared tropospheric NO_2 VCDs estimated using the simple geometric approximation to those derived in a more robust way by calculating air mass factors for different aerosol conditions. They found that the biases between both methods are rather small, i.e. within 10% in most cases and within 20% in all cases. So, the simple geometric approximation applied on MAX-DOAS measurements at 30° elevation in Beijing and Xianghe provides tropospheric NO_2 VCDs with a sufficiently high level of accuracy, especially given the fact that they are not exploited geophysically but only used as a priori information in the OEM-based retrieval.

P10632, L3f: The rejection of contaminated retrievals is a very crucial issue, so some more information might be useful here. For example, did you check what is causing the rejection using independent data like cloud observations? How many scans are rejected based on each of the three criteria and overall? Why don't you use the color index as additional information as presented by one of your colleagues at the EGU meeting this year ("Development of a cloud-screening method for MAX-DOAS measurements", C. Gielen et al., 2013)?

In the revised manuscript, we have added a fourth criterion: scans with very large AOD values (>6) are also rejected given the significantly larger uncertainties on the trace gas retrievals in such conditions. This new criterion has only a limited impact at Beijing and no impact in Xianghe. Using our four criteria, we reject about 35% of the scans at both stations.

We have used the cloud screening approach developed by C. Gielen from BIRA-IASB to make some statistics on the sky conditions corresponding to the selected and rejected scans as well as to the entire sets of scans at both stations. To our opinion, in the case of trace gas retrievals, such a cloud screening tool must be used more as ancillary information on the selected scans rather than as a scan selection tool given the fact that a cloudy situation does not necessarily result in a bad trace gas retrieval. Gielen's method is based on the color index, i.e. the ratio of the intensities at 405 and 670 nm in zenith geometry. Normalized color index for Beijing and Xianghe are presented in Fig. 1.

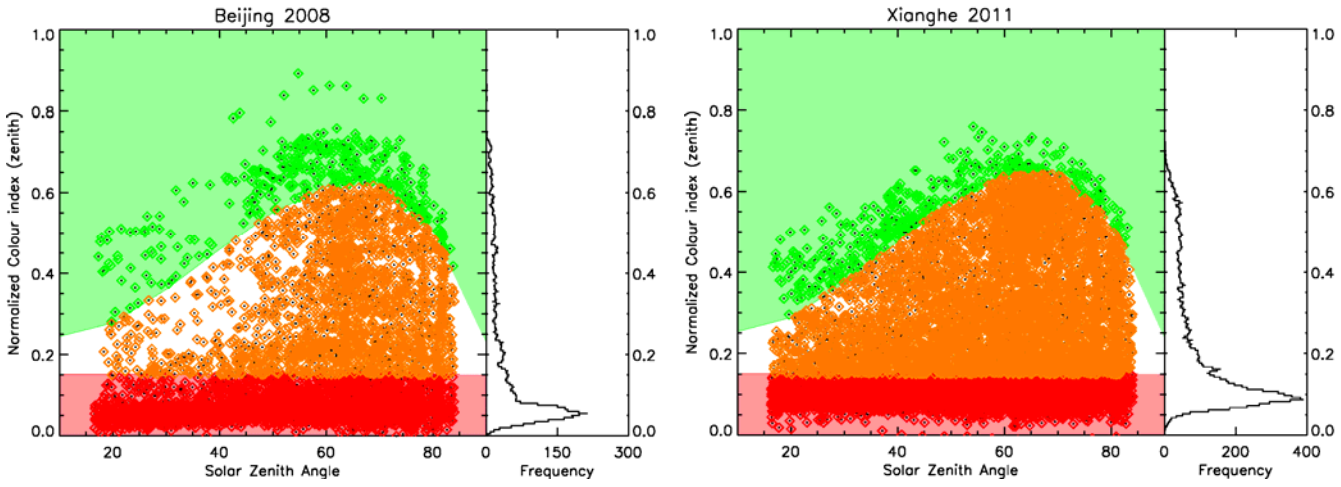

Figure 1: Normalized color index for MAX-DOAS scans at Beijing in 2008 (left plot) and at Xianghe in 2011 (right plot). Three regimes are defined based on color index simulations for different sky conditions: clear-sky (CS; green area), fully cloudy/extreme aerosol load (FC; red area), and intermediate conditions (orange area). The attached right plot indicates the frequency of occurrence of color index values.

Given the definition of our color index, values close to 1 means clear-sky conditions and when it is close to zero, it indicates fully cloudy conditions. Three

regimes have been defined based on color index simulations for different sky conditions: clear-sky (CS; green area), fully cloudy including extreme aerosol loads (FC; red area), and intermediate conditions (orange area). We see from Fig. 1 that in the Beijing area, the number of clear-sky scans is very small compared to the fully cloudy and intermediate cases. Fig. 2 shows the statistics of the sky conditions corresponding to the selected and rejected scans using our four criteria and to the entire sets of scans at both stations.

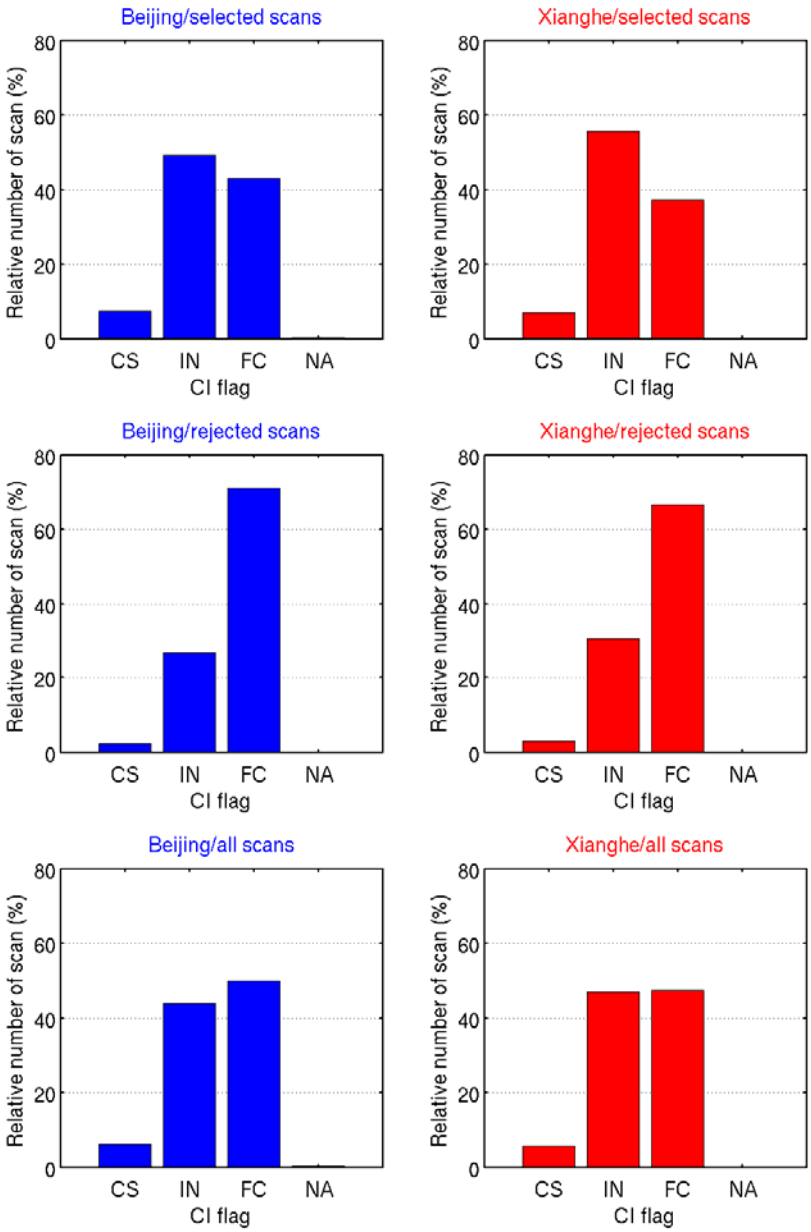

Figure 2: Statistics of the sky-conditions derived using the color index approach and corresponding to the selected (upper plots) and rejected (middle plots) scans using our four criteria and to the entire sets of scans (lower plots) at both stations (blue for Beijing, red for Xianghe). CS is for clear-sky, IN is for intermediate, FC is for fully cloudy, and NA is for unsuccessful cloud retrieval.

As can be seen, our four selection criteria lead at both stations to a larger relative contribution of scans corresponding to clear-sky and intermediate conditions compared to the entire sets of scans. Regarding the rejected scans, the contribution of fully cloudy scans significantly dominates the two others.

So, the cloud screening approach gives useful indications on sky conditions corresponding to the selected and rejected scans. However, since this cloud screening method is not used for the selection of useful scans in the present study and has not yet been published so far, we decided to not include the above discussion in the revised manuscript.

P10635, Sec.3.2: Why are there no rush hour peaks in your NO2 signal? LP-DOAS measurements of NO2 ground concentrations e.g. in Hong Kong can see a significant increase during rush hour time (see Chan et al., NO2 measurements in Hong Kong using LED based long path differential optical absorption spectroscopy, AMT, 2012), so it should look similar for Beijing, at least to some extend in your surface concentration results. Could you elaborate on what you think is causing the difference in diurnal variations compared to the results by Ma et al.(2013)? The spring and summer cycles don't agree at all and I wouldn't call a correlation coefficient of 0.6 for fall to "agree well". Did you compare only the average for the entire period or could you compare the seasonal averages on an annual basis? It would be interesting to see if the discrepancies change over time or not.

In the ACPD paper, we extracted the NO_2 VCD diurnal variations published in the Ma et al. (2013) paper and those were directly compared to our retrieved values. The problem is that the periods corresponding to both data sets were not the same: July 2008-April 2009 in our case and August 2008-September 2011 in Ma et al. (2013). The revised manuscript now includes a comparison of NO_2 VCDs for the same period, thanks to data provided by Jianzhong Ma from the Chinese Academy of Meteorological Sciences in Beijing (now co-author of the article). As can be seen in the revised paper, the agreement is very good, lending confidence to our NO_2 retrieval.

The rush hour peaks are indeed not well marked in our observations, as is also the case in the Ma et al. (2013) data set. So far, we have no clear explanation for this feature. Maybe it could be simply related to the fact that given the horizontal representativeness of MAX-DOAS observations (generally around 5-10 km or less in a polluted region as Beijing), the effective location probed by the measurements is less sensitive to the traffic.

References:

Clémer, K., Van Roozendael, M., Fayt, C., Hendrick, F., Hermans, C., Pinardi, G., Spurr, R., Wang, P., and De Mazière, M.: Multiple wavelength retrieval of

tropospheric aerosol optical properties from MAXDOAS measurements in Beijing, Atmos. Meas. Tech., 3, 863-878, 2010.

Elshorbany, Y. F., Steil, B., Brühl, C., and Lelieveld, J.: Impact of HONO on global atmospheric chemistry calculated with an empirical parameterization in the EMAC model, Atmos. Chem. Phys., 12, 9977-10000, 2012.

Li, X., Brauers, T., Häseler, R., Bohn, B., Hofzumahaus, A., Holland, F., Lu, K. D., Rohrer, F., Hu, M., Zeng, L. M., Zhang, Y. H., Garland, R., Su, H., Nowak, A., Takegawa, N., Shao, M., and Wahner, A.: Exploring the atmospheric chemistry of nitrous acid (HONO) at a rural site in Southern China, Atmos. Chem. Phys., 12, 1497-1513, 2012.

Ma, J. Z., Beirle, S., Jin, J. L., Shaiganfar, R., Yan, P., and Wagner, T.; Tropospheric NO_2 vertical column densities over Beijing: results of the first three-years of ground-based MAX-DOAS measurements (2008–2011) and satellite validation, Atmos. Chem. Phys., 13, 1547-1567, 2013.

Villena, G., Kleffmann, J., Kurtenbach, R., Wiesen, P., Lissi, E., Rubio, M., Croxatto, G., and Rappenglück, B.: Vertical gradients of HONO, NO_x and O_3 in Santiago de Chile, Atmos. Environ., 45, 3867–3873, doi:10.1016/j.atmosenv.2011.01.073, 2011.

Wong, K. W., Tsai, C., Lefer, B., Haman, C., Grossberg, N., Brune, W. H., Ren, X., Luke, W., and Stutz, J.: Daytime HONO vertical gradients during SHARP 2009 in Houston, TX, Atmos. Chem. Phys., 12, 635-652, 2012.