We would like to thank both referees for their positive review of our paper and the constructive comments which we address below, one by one. For the sake of clarity, the original text from referees is first returned in bold italic blue font.

Answers to Anonymous Referee 1:

The paper by Pinardi et al. is a thorough assessment of state-of-the-art remote sensing of atmospheric HCHO abundances by ground-based multi-axis differential optical absorption spectroscopy. The study evaluates the performance of remote sensing instrumentation and data reduction techniques among various research groups who made a joint effort of a dedicated comparison campaign. Further, the paper provides an indepth analysis of potential error sources including correlations between the retrieval parameters, systematic and random errors. A recipe how to improve the retrieval setup is proposed.

The paper is very well written, the employed methods are well thought and well documented.
The topic will be of interest to AMT readers, in particular to the DOAS community.
Therefore, I recommend publication with only minor modifications according to some of my comments listed below.

Comments:
p.6681,l.9: molec cm2 -> molec cm-2. This error occurs several times in the manuscript, in particular in section 5. Please check the whole manuscript.

Reply: This error was not noticed during the proof-reading. It has now been corrected in the whole manuscript.

p.6681,l.21: integration time -?-> exposure time. To my knowledge the correct term would be "exposure" not "integration". If so, please check the whole manuscript.

Reply: Although we agree that "exposure time" is a perfectly valid term, we would prefer to keep "integration time" as used in the initial manuscript. The main reason is to maintain the coherency with the terminology used in the CINDI overview paper by Piters et al 2012 (table A1) as well as in Roscoe et al. (2010).

p.6684,l.15: evaluations -> retrievals.

Reply: Done.

Paragraph 3: Later in the manuscript, the authors mention that others (Roscoe et al.) found that the accuracy of the pointing information plays a crucial role for the NO2 and O4 comparison study and when converting DSCDs into VCDs. Pointing errors would also affect the comparison between the different instruments. Please shortly comment on this issue.

Reply: Pointing errors are a potentially important source of uncertainty for MAXDOAS measurements as discussed in Roscoe et al. (2012). This error is particularly important for trace gases having sharp concentration gradients in the first layers above the surface, such as NO_2. In the case of HCHO, which is typically more smoothly distributed in the boundary layer, the sensitivity to pointing errors is slightly less critical. Nevertheless this is an important issue which has been addressed early in the CINDI campaign, actually before the HCHO comparison was started. Owing to specific field measurements, the alignment was checked for each instrument and pointing issues were minimized. A sentence has been added at the end of Sect. 5.2. to make this clear.

p.6690,l.20: as it's the information used -> since this the information used

Reply: Done.

p.6691,l.12: The general reader might not be aware of the AMF concept. Explain shortly.

Reply: The following sentence (bold) has been added in the text:

… and second, from direct conversion of the zenith-sky observations using appropriate air mass factors (AMFs). **For one given observation, the AMF represents the geometrical enhancement factor that results from the geometry of the (MAX)DOAS observation and from the scattering properties of the atmosphere. AMFs are calculated using radiative transfer models accounting for multiple scattering and earth sphericity. They allow for conversion of the measured slant columns into equivalent vertical columns.** For the present analysis, and considering that our test data were recorded under clear-sky aerosol-free conditions, zenith-sky HCHO AMFs were calculated using the UVspec/DISORT model

p.6692: What is the physical difference between the methods for calculating the Ring effect, why is there such large differences in the Ring baseline?

Reply: The Ring cross-sections used in this study have been generated using slightly different (although related) methods. In all cases, one assumes that rotational Raman scattering (RRS) by molecular nitrogen and oxygen is the main cause of the Ring effect. The baseline case follows the method described in Chance and Spurr (1997). It consists in a simple convolution of a reference solar spectrum using RRS cross-sections. For the second case (case A), the approach is the one proposed by Wagner et al. (2009). It is very similar to Chance and Spurr (1997) except that the convolution is performed using normalised RRS cross-sections instead of absolute ones, which leads to slightly different spectral structures in the resulting Ring cross-section. For cases B and C, the Ring cross-sections were obtained using full radiative transport calculations including the effect of RRS and its temperature dependence. In case B, a simple Rayleigh atmosphere was assumed, while for case C, a principal component analysis approach following Vountas et al.(1998) was applied. These different approaches of calculating a Ring effect cross-section all approximate the actual physics of the problem, but in slightly different ways. We note that in each case slightly different misfit effects are obtained resulting in different biases on small absorbers such as HCHO and BrO.

p.6692,l.22: polynomials -> polynomial

Reply: Done.

p.6695,l.24: details -> detail

Reply: Done.

p.6697, l.16: as it's the main parameter -> as these are the main input parameters for retrieving profile information

Reply: Done.

Fig.2: Explain the panels and the two colors in the upper panel.

Reply: A legend has been added in the first panel and the caption has been improved.

Fig.5: The black circles are too big. In my printout, there is only a big black blob visible. The units of the regression parameters might be obvious, but for completeness one could consider mentioning (likewise Fig.6, Fig.7).

Reply: The units have been added and the figure has been plotted with smaller dots. However, due to the position of the points, this is not entirely solving the "blob" impression.

Fig.8, caption: "after straight line fits of each instrument's data" I do not understand this comment. Isn't it just subtracting the respective data from the reference data?

Reply: The caption has been changed to:
"Histograms of the HCHO DSCD absolute deviations (10^{15} molec/cm²) of each instrument's data compared to the reference set, for the case of measurements at 4° elevation angle, and for the whole campaign."

Fig. 9, 10, 11, 16: Units not explicitly mentioned.

Reply: Units have been added in the figure captions.

Answers to Anonymous <u>Referee 2:</u>

This paper presents intercomparison results for HCHO slant columns measurements obtained during a dedicated campaign. The paper is well written and the methodology is very good. All the sources of uncertainties are well evaluated and discussed. This is a nice presentation of how intercomparison analysis must be conducted. I strongly support the publication of the paper, if the minor comments and improvements listed below are taken into account.

Page 5: Can you explain what is a "semi-blind intercomparison"?

Reply: A semi-blind intercomparison is an intercomparison campaign monitored by a referee that follows a pre-defined protocol for the data comparison. This protocol does not allow for communication of the intercomparison results during the campaign, other than through comparison plots where data owners are not explicitly named. For more detail, please consult Roscoe et al. 2011, where the protocol of the semi-blind exercise is explained. For the case of the HCHO exercise the comparison was fully open. The expression "semi-blind intercomparison" only refers to the time period between 15 and 30 June (when the semi-bind NO2 and O4 intercomparison was conducted). No further explanation has been added in the manuscript.

Page 6: If I well understand, only the 30° and zenith measurements are kept, because measurements with SZA less than 75° are kept. But later on the paper, measurements at lower elevation are presented. There is confusion somewhere. Can you better explain the geometry of the measurements?

Reply: During the whole campaign, data were measured at different elevation angles (2°, 4°, 8°, 15°, 30° and the zenith), all day long. For our HCHO intercomparison, the off-axis elevation angles are compared, but only for SZA<75°. The main motivation for doing so is to gather data from all groups (not all groups have been measuring during twilight) and to avoid considering error-prone twilight measurements not relevant for HCHO. For Figure 3, only measurements at 30° elevation and at zenith have been used, and the calculation of the vertical column has been performed using a simple geometrical approximation.

Page 7: Can you explain what a "solar I0 correction" is?

Reply: Basically, the I_0 effect arises because sky-light spectra and laboratory absorption cross sections are usually measured at different spectral resolutions and using different light sources. To match the resolution of sky-light instruments, high-resolution laboratory data are traditionally filtered using the known instrumental slit function, whereby an error is introduced because in actual measurements the spectra (and not their logarithm) have been filtered by the slit function. A detailed description of this effect and the method used in the present study to account for it are described in Aliwell et al. (2002).

Page 8 and Figure 3: Are the same errors bars for Figure 3a than for Figure 2b? If yes, can you indicate this on the figure caption? I am afraid with such large errors bars, which seems larger that the HCHO diurnal variation. It is said that the errors bars are for individual measurements. Then, it is confusing to plot on the same figure the individual errors and the mean values. It could be better to plot the errors bars of the mean values, and to write somewhere the typical error values for the individual measurements. Nevertheless, this is confusing. Is it errors or real measurements dispersion?

Reply: We assume that referring to "Figure 2b" in this question is a typing error for "Figure 3b" as no error bars are displayed in figure 2. In figure 3b the dots are the hourly averages of the black dots of figure 3a, while the error bars are the corresponding standard deviations, presenting the measurement dispersion.

Figures 5 and 6: The dashed line is quite confusing. It seems that the lines correspond to "y = x", but can be confused with a real fit following the values given above the figures. Can you indicate this, or even better, put the real fits?

Reply: Done. It has been added in the figure caption that the dashed line is the y = x line.

Page 11: Can you describe in more details the "geometrical approximation"?

Reply: The geometrical approximation makes the assumption that the trace gas under investigation is entirely confined in a thin layer above the surface, located below the last scattering point. In this idealised case, the tropospheric AMF can be derived from simple geometrical considerations: $AMF_{\alpha} = 1/\sin(\alpha)$, where α is the elevation angle. In order to remove the stratospheric content, the zenith measurement closer in time is subtracted to get the tropospheric VCD (not an issue for HCHO):

$VCD_{tropo} = SCD_{tropo} / AMF_{tropo}$

$= (DSCD(\alpha) – DSCD(\text{zen of the scan})) / (1/\sin(\alpha) - 1/\sin(90°))$

$= dDSCD(\alpha) / (1/\sin(\alpha) -1)$

For $\alpha=30°$, $VCD_{tropo} = dDSCD(30°)$.

We refer to Hönninger et al. (2004) where the geometrical approximation was first introduced. We also add a reference to the more recent Ma et al. (2012) paper where all the formulas are detailed in the text. (Ma, J.Z., Beirle, S., Jin, J.L., Shaiganfar, R., Yan P. and T. Wagner, Tropospheric NO2 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. Discuss., 12, 26719–26781, 2012, http://www.atmos-chem-phys-discuss.net/12/26719/2012/acpd-12-26719-2012.html).

Page19: Can you estimate more precisely the effect of "atmospheric noise", in particular by providing an estimate of the error produced by this effect?

Reply: The term atmospheric noise was used to design the atmospheric variability of HCHO in time and space, occurring during the time of measurement. However, its estimation is difficult, due to the simultaneous instrumental measurement uncertainty, and out of the scope of the slant columns error estimation performed here. The sentence has thus been removed.