

### Reply to comments from Referee #3

*“Evaluating different methods for elevation calibration of MAX-DOAS instruments during the CINDI-2 campaign” by Sebastian Donner et al. 2019 addresses a very important topic of pointing accuracy of the MAX-DOAS instruments. The authors describe and compare four different methods to measure offset of the actual pointing from the expected pointing. This topic is very important for reliability and consistency of MAX-DOAS observations and is within the scope of the Atmospheric Measurements Technics. I strongly believe this paper can become a great resource for MAX-DOAS community. I recommend publishing it after some changes.*

We would like to thank the referee for this positive assessment and several useful comments.

#### Major comments:

*1. While the described methods are intended as calibration of instrumental pointing accuracy no discussion of the instrument performance evaluation after calibration was applied is presented.*

Many thanks for this valuable comment. The instrument performance was assessed in the overview paper by Kreher et al. (2019). Our study focuses on the details of the elevation calibration procedure. This information was added to the paper. Further, a statement was added in the conclusions that in upcoming campaigns strong emphasis should be put on the monitoring and possible correction of the elevation calibration.

*2. The paper mainly focuses on determination of the instrument specific apparent elevation angles of the target and inadequately addresses the errors associated with determination of the absolute position of the calibration targets.*

Thanks for this good comment. The estimates of the errors of the absolute positions of the targets and the corresponding uncertainties of the elevation angles are given as follows:

-White stripe:  $\pm 5 \text{ mm} \Rightarrow \pm 0.08^\circ$

-near lamp:  $\pm 2 \text{ mm}$  and  $0.1^\circ$  (provided by IUP-HD)  $\Rightarrow \pm 0.11^\circ$

-far lamp:  $\pm 50 \text{ cm} \Rightarrow \pm 0.02^\circ$

-horizon:  $\pm 5 \text{ m} \Rightarrow \pm 0.22^\circ$

The uncertainties of the angles were calculated from the uncertainties using geometry. We added these accuracy estimates for the positions of the targets to the revised version of the paper.

*3. The authors often use method precision to describe accuracy. To evaluate usefulness of the presented methods both are needed.*

An appendix was added mainly consisting of two tables describing and evaluating different systematic and statistical error sources. The main findings are:

- The systematic uncertainties are dominated by the errors introduced by the determination of the target positions and the fit errors
- The statistical errors are dominated by the precision of the used motors/steppers

Based on these findings, different passages in the paper were modified and new ones were added to address this comment. Further, the error assumptions were motivated more clearly.

*4. Measurements of distances, heights, estimation of water levels have no associated with them measurement accuracy and precision reported. Sometimes details how these measurements or estimations were conducted are missing completely.*

The uncertainties of the measurements of horizontal distances are not critical and can be neglected since the absolute horizontal distances are much larger than the vertical distances. However, error estimates were added to the revised version of the paper (see answer to major comment 2). Further, the descriptions of the determination of vertical (and partly horizontal) distances were revised.

*5. Fits of Gaussian functions to data have no fitting errors reported.*

(Standard) fit errors were added to all plots showing Gaussian fits. Additionally, typical values for the fit errors were added to the revised version of the paper, especially in tables A1 and A2 which summarise the error sources.

*6. Five Pandora instruments (1 KNMI, 2 LuftBlick and 2 NASA) during CINDI-2 were performing sun scans on a regular basis (once per hour) to actively calibrate their azimuth and zenith pointing. This method should also be described for comparison with the other methods.*

Thanks for this valuable and important hint. A new section describing the sun scans was added to the revised version of the paper. Further, this method is discussed in the conclusion section of the paper and mentioned at different sections of the revised paper.

*7. More emphasis should be placed on the quality of the positioners.*

Thanks for this interesting comment. However, Figures 20 and 21 show that for most (better performing) instruments the daily horizon elevations can be reproduced quite well and the values scatter rather closely around their median value. The reproducibility of the horizon elevations is typically significantly better than  $\pm 0.1^\circ$ . Given the uncertainties in the determination of the visible horizon this is rather small. The same is found for the far lamp measurements (Figure 18). Here, small error bars are found for the standard error of the mean lamp elevation which also indicates good reproducibility. Therefore, we conclude that the quality of the positioners does not dominate the overall uncertainties and does not put doubt on the findings of this study for most of the instruments. However, we added detailed information on the statistical and systematic error sources to the paper (especially in the two new tables A1 and A2 in the Appendix) for the different methods as described in this answer to the revised version of the paper.

*8. More explanation needs to be provide on how exactly the horizon scans can be used as a calibration tool (considering dependence on the FOV, scattering conditions, uncertainty in underlying surfaces, light incident angles, and true horizon).*

The description of the horizon scans is already quite detailed and also precautions and prerequisites are discussed at several passages of the paper (e.g. P.8 L.8-12 or P.14 L.26-27). However, we agree that this information is quite scattered over the whole paper. Therefore, we added and stressed some precautions and prerequisites which have to be fulfilled to use the horizon scans as a useful calibration tool. Namely:

- Days with good visibility should be used
- Days with rapidly varying cloud cover and/or low-lying clouds should be avoided

- Another interesting aspect is that with increasing FOV the slope of the horizon scans becomes weaker. However, in spite of this weaker contrast this is in general no problem.

9. Paper can be reorganized to be more concise. Some of the tables and figures can be merged (e.g. Table 1 and 2) and some eliminated at all. Text has some redundancy and needs proofreading. I recommend creating a table with a summary of each method including: (1) setup and “absolute” prerequisites; (2) measurements needed, their typical accuracy and precision, data analysis involved; (3) advantages; (4) disadvantages; (5) overall expected accuracy and precision of zero-elevation calibration based on CINDI-2 data for different types of instruments.

Tables 1 and 2 were merged as well as Figures 4 and 5. Further, Figures 12 to 15 were combined. Figure 23 was removed from the paper and the respective paragraph was shortened in the manuscript. Regarding the proofreading, we agree that both grammar and spelling is not perfect since we are no native speakers, however, Copernicus will provide a proofreading procedure during typesetting. An overview table (new table 5) summarising the different methods, listing advantages and disadvantages and giving uncertainty estimates was added.

### **Minor comments:**

*P2, L31-32: Do any of the instruments have laboratory done FOV scans? It will be interesting to compare field of view between the lab and in the field.*

Thanks for this interesting question. The values for the FOVs listed in Table 2 were provided by the groups and were determined in the laboratory. Despite the fact that the determination of the FOV is not the main aim of the paper, we added a new section “FOV determination” to the paper, where a comparison between the retrieved FOVs (from horizon and lamp scans) and the reference FOV is provided. In general reasonable agreement is found. However, the determination of the FOV seems to be less stable as compared to the determination of the target positions.

*P3, L13-14: This sentence is unnecessary.*

The sentence was removed from the paper.

*P3, L28-29: Figure 19 has data for 28 instruments. Why Table 1 lists only 12? Tables 1 and 2 should be combined.*

Thanks for pointing this lack of clarity out. The horizon scans were performed by all instruments during the campaign which followed the standardised measurement protocol and reported them to the referee of the campaign (28 instruments). However, only 12 instruments (from 11 groups) participated in the far lamp measurements. This was made clearer in the abstract and the text. Further, we removed the number of instruments from the abstract (P.1 L.5) and revised section 2.2. As already mentioned above Tables 1 and 2 were merged and the instrument ID was removed.

*P3, L31: There is no need to cite the URL when Kreher et al is already cited.*

The URL was removed from the paper.

*P4, L2: This sentence is redundant.*

This sentence is meant to stress the variety of the different instruments and is kept in the revised version of the paper but in a modified form since Tables 1 and 2 were merged.

*P4, L15: Five Pandoras participated in CINDI-2. Each of them performed sun scans as part of routine operation that served as azimuth and elevation calibration. This method should be also presented for comparison.*

Many thanks for this hint, to which we fully agree. See also answer to major comment 6.

*P4, L16: What is "horizontal line of the telescope"? Is it the optical axis of the telescope/ fiber setup when the instrument points at zero-degree elevation angle? How do you determine it?*

The "horizontal line of the telescope" is defined as the line of sight of the instrument at 0° elevation. This is the property which is actually calibrated by all the methods. This information was added to the paper.

*P4, L30: This is another reason why sun scanning by Pandora instruments should be discussed in this paper.*

See answer to major comment 6.

*P5, L6: How were the distances measured from the lamp to the instruments? How were the vertical distances measured? The land and the canal banks were covered with grass and are not perfectly flat. What is the uncertainty in all distance measurements?*

The horizontal distance was measured using Google Maps, where the location of both the lamp and the measurement site can be clearly identified (see Figure 3). Further, it should be mentioned that the accuracy of the horizontal distance is not critical, since the distance is quite large (more than 1 km). Regarding the vertical distances: they were measured manually using a laser level which was projected onto a folding rule and the channels located next to the lamp and the measurement site. In that way, first the height difference between lamp and the channel's water surface could be determined. Since all channels were connected to each other (except one step which was determined in the same way), the lamp position could be marked on the containers as indicated in figure 3 and described in section 3.2.1. Of course the banks and the land are not perfectly flat, however, the error introduced by that is very small and the overall uncertainty is dominated by measurement errors. The total error in the determination of the lamp mark is estimated to be around 0.2 m which translates to an uncertainty of 0.01° in the lamp position. Additionally, the height differences ( $\Delta h$ ) between the lamp mark and the telescope units of the instruments have to be determined. For this difference an error of around 30 cm is estimated. Therefore the total uncertainty of the lamp position relative to the telescopes is estimated to be 0.5 m which translates to an uncertainty of 0.02° (see also answer to major comment 2). This information as well as revised description of the estimation of lamp position was added to revised version of the paper.

*P5, L8: Information from Fig 2 can be communicated in Fig 3 and Fig 2 removed.*

Thanks for this idea, however, we do not really agree here. Figure 2 explains the general idea of the elevation calibration procedure which is the same for all methods. Figure 3 sketches the specific application of this general idea to the specific case of the far lamp measurements and gives more details on the setup which was actually used during the campaign. Therefore, we

think figure 3 would be overloaded if we would add the information given in figure 2, and thus we decided to keep both figures in the paper.

*P5, L7: The word "compare" next to most Fig and Table references is unnecessary.*

Thanks for this hint. For the revised version of the paper we checked for those phrases and removed redundant ones.

*P5, L8: The lamp light was collimated and then "directed". How exactly was this achieved? What was the accuracy of the lamp pointing? How uniform was the resulting beam that was visible on the container?*

As described in the paper the lamp was directed towards the instruments using a large aperture lens. Here, the lamp was put in the focal point of the lens which was achieved by minimising the size of the beam (this was already done prior to the campaign). Then the lamp was manually directed towards the campaign site by eye. Here, it should be noted that the exact pointing is not critical as long as the instruments are located within the light cone. We assumed that the diameter of the lens is homogeneously bright. Nevertheless, also this assumption is not a critical point, because the angle under which the full lens is seen from the campaign site is  $<0.01^\circ$ . We added this description to the revised version of this section.

*P5, L19: Which earlier night, 8 Sep 2016, or before that? Is 0.16 deg the offset from the initial "a priori" calibration done in the lab or from the earlier night?*

Thanks for pointing this like of clarity out. The pre-calibration was done using a water level during the setup of the instrument. Then the finer adjustment was done using the results of the far lamp scans from 7<sup>th</sup> (in this night the lamp measurements were tested by our group with an scan resolution of  $0.1^\circ$  but the scanning was done manually), 8<sup>th</sup> and 10<sup>th</sup> September. The other two nights are then somehow tests of reproducibility. All values in this paper are given relative to the elevation calibration which was obtained by these finer adjustments and which was finally used for the campaign. This information was added to the text.

*P5, L21: Fig 4 is unnecessary and should be removed.*

The exact spectrum of the lamp depends on several properties, e.g. pressure in the lamp and optical filters. Thus, we think it is quite interesting and useful to show a real measured spectrum of the lamp. However, we combined figures 4 and 5 to reduce the numbers of figures in the paper.

*P6, L14: How did you decide that one direction was better than the other? What was positioning error of the lamp? Please include characteristics of instrument positioners (manufacturer accuracy and precision, and used resolution) in Table 2.*

Since also the elevation sequences defined in the synchronised measurement protocol had ascending elevations, we decided on an ascending scanning scheme for all other calibration exercises in order to be consistent. Regarding the positioning error of the lamp, see the answers to major comments 2 and 4. As pointed out above (answer to major comment 7) the positioner accuracy is not the limiting factor of this study, which is dominated by uncertainties related with the determination of the target positions.

*P6, L18: Is 0.16deg the initial calibration? Or is this the effect of the positioner resolution?*

Regarding part 1 of the question, see answer to comment “P5, L19”. Regarding part 2: Here we conclude that the  $0.16^\circ$  is an effect of the scan resolution. The positioner resolution is  $0.01^\circ$  for the MPIC instrument.

*P7, L9-10: What is the leveling accuracy of the laser level? What light source is used? How is uniformity of the beam achieved? How accurate is determination of the light source center? What are the requirements of light source installation? It is also assumed that the optical axis of the telescope/fiber setup co-align with the mechanical center of the telescope (e.g. the fiber however can be slightly higher or lower than this estimation). What is the final error in determining this beta (= zero) offset angle between the center of the telescope/fiber and the lamp?*

The accuracy of the laser level has been tested in the lab and amounts to approximately  $0.1^\circ$ . A Hg-Lamp was used. Using a cylinder lens, the laser beam leaves the laser level as a horizontal stripe with a thickness of approximately 2 mm, which is the limiting factor in the determination of the light beam centre (which leads to about  $0.04^\circ$  additional uncertainty in the determination of beta). The stripe has a certain curvature (smile), which is accounted for by only considering the centre of the stripe. This commercial laser level comes with a tripod with adjustable height that can be placed onto any suitable surface. The laser beam creates parallel light that hits the entrance optics approximately at its centre, as controlled by eye. Any vertical displacement of the laser beam is not of importance, since the incoming parallel light is projected onto the focal point of the entrance lens where the fibre entrance is located, independent of any displacement of the beam from the optical axis. The information on the accuracy of the laser level and the error introduced by the beam thickness was added to the paper.

*P7, L18-19: Were all the scans done from the same direction (upwards or downwards). Looking at the intensities for scan 2 and 3 they might be an indicator of the positioner backlash or pointing issues.*

The elevation pointing is continuously regulated by comparison of orientation of the telescope measured by the built-in tilt sensor with the nominal angle. Therefore, there are no backlash effects to be expected. This information was added to the paper.

*P9, L5: It is not quite clear how this FOV determination eliminates dependencies on the scattering conditions (wavelength), underlying surfaces and their albedo, as well as solar position.*

This method determines the “effective” FOV for that specific measurement. Many measurements under different sky conditions give good statistics. The mentioned effects are not critical as long as the horizon is clearly visible and only measurements with favourable conditions (high visibility, no variable and/or low-lying clouds) are used. As described above (answer to major comment 8), this information is provided several times in the manuscript. However, it is stressed more in the revised version of the paper.

*P10, L23. Should “horizon” be replaced with “horizontal”?*

Thanks for this hint. “Horizon” is indeed not correct. Actually, “target positions” fits better. The text was adjusted accordingly.

*P10, L26: How was “visible horizon” determined? How was “closeness” to the visible horizon determined? Figure 3 suggests that the lamp was at 3.3 m above ground. Is “ground level” referring to 3.3 m above ground?*

As “visible horizon” we defined the transition of tree tops to the open sky. We replaced the sentence “..., the xenon lamp was placed close to the visible horizon but at ground level...” by “..., the xenon lamp was placed directly in front of a row of trees which mark the visible horizon (the transition of the tree tops to the open sky)...”. Further, the following statement was added to the paper: “The vertical position of the lamp was 3.5 m above the water level in the water channel which was located next to the measurement site (see Fig. 3), and there a few meters below the tree tops.” Additionally, the whole paragraph was slightly modified (see also answer to the next comment).

*P10, L29. How was 6.5 to 8.0 m estimated?*

These heights were estimated from the differences between lamp scan and horizon scan results using simple geometry.  $0.3^\circ$  and  $0.37^\circ$  lead to heights of 6.7 m and 8.27 m (at lamp distance), respectively. Since the lamp is in average seen at  $0.01^\circ$  (corresponding to 0.22 m at lamp distance) for the MPIC instrument this leads to heights of 6.5 m and 8.0 m, respectively. These heights are consistent with the (estimated by eye) heights of trees. This is a little bit hand waving but overall consistent and the differences are understandable. A summary of this information was added to the paper.

*P11, L13: I recommend combining Fig 12-15 into one figure. These figures give a good sense of the apparent FOV, as the lamp is scanned, impacted by both optics and precision of the positioner. Selecting the azimuth with the maximum intensity is somewhat arbitrary for some of the instruments with asymmetric FOV (e.g. Fig 13).*

Thanks for this good suggestion. Figures 12-15 were combined. The reviewer is right that the selection of the azimuth with the maximum intensity is somehow arbitrary. However, this choice is not critical for the interpretation of the results.

*P11, L20: Differences in positioner pointing precision is also an important parameter of the apparent FOV.*

Thanks for this good comment. This information was added to the text.

*P11, L21: Fig 12 suggests that FWHM is about 0.65 deg for the BIRA instrument not 1 deg.*

Thanks for this hint. As mentioned at the first minor comment, we added a section on the comparison between the retrieved FOVs (from horizon and lamp scans) and the reference FOV. There, smaller retrieved FOVs were found compared to the ones provided by the groups. However, as explained in the answer of the next comment the mentioned passage was removed from the paper.

*P11, L22: I would say good alignment, center “spot” fiber arrangement and good positioning are the reasons for a relatively uniform FOV. Unless the lamb beam was not uniform.*

Thanks for this good comment. We already mentioned the fibre arrangement in the paper, however, did not stress this enough. The text was adjusted accordingly. Further, the quality of the positioning was added to the text. After discussion with the co-authors, the influence of the size of the FOV on the smoothness of the intensity distribution was removed.

*P12, L23: Isn't the spread expected due to differences in prior reference calibration of 0 elevation angle?*

Thanks for the comment. We think it refers to P.13 L.3. The reviewer is right, the spread is expected due to differences in the prior reference calibration of the 0° angle as explained in the text (P.13 L.12-17).

*P13, L9: Precision (repeatability) will be more appropriate here instead of "accuracy".*

Thanks for this hint. The text was adjusted accordingly.

*P13, L16: I would replace "measure of consistency" with: "measure of variability".*

We agree, the text was adjusted accordingly.

*P13, L22-23: How were 6.5 and 8 m as a function of wavelength (340 and 440 nm) estimated?*

They were estimated using the horizon elevations of the MPIC instrument as described on P.10 L.29-32. A more detailed description can be found in the answer to the comment P10. L29 and was added to the text in section 3.6 (3.7 in the revised version). Further, a cross reference to this section was added in the revised version of the paper.

*P14, L32: Reading the text that follows you assert this difference is due to surface reflectivity. I also will add effects of FOV and wavelength dependent scattering. For a system with a 0.6 deg field of view placed directly on the ground pointing at 0 deg elevation angle (assuming no obstacles): half of the FOV will receive photons scattered in the atmosphere and half reflected from the underlying surfaces. Since Rayleigh scattering is wavelength dependent more photons at longer wavelength would be scattered in the "above ground" half of FOV. So for the instruments with FOV ~0.6 nm FWHM the telescopes should point at least 0.3 deg below horizon to minimize the effect of FOV size. Considering that the instruments during CINDI-2 were located ~4 m and 7 m above ground this angle should be even larger and depends on the wavelength and distance to obstacles and solar position (SZA and RAA).*

Thanks for this comment and the detailed explanation. However, radiative transfer simulations confirmed that the wavelength dependence of scattering cannot explain the observed wavelength dependence of the elevations scans. This information was added to the paper and the whole part was shortened.

*P14, L29: Figure 19 does not support this statement. BIRA\_4 instrument horizon position is about the same for both wavelengths.*

Many thanks for this hint. We corrected the text accordingly.

*P15, L25-31: It is unclear how apparent horizon measurements can be suitable for pointing accuracy calibration.*

Indeed the horizon scans have to be considered with care. But as long as the apparent position of the horizon is known, horizon scans are a valuable tool for elevation calibration. However, only measurements under favourable conditions (high visibility, no rapidly varying and/or low-lying clouds) should be used. As mentioned in the answer to major comment 8, we added these

precaution measures which should be fulfilled to the text (horizon scans and conclusions sections). Further, in this text passage “apparent” was changed to “visible”.

*P16, L5: The conclusion maybe applies to the better performing instruments, while the rest of them mostly excluded from the analysis.*

This conclusion is derived from the regression analyses which included all instruments which performed lamp measurements and horizon scans. Naturally, some perform better than others. Reasons for deviations were assessed in sections 4.1, 4.2, 4.3, however, these descriptions were slightly modified for the revised version of the paper. Nevertheless, the results of most of the instruments (for which both methods were applied) are rather close to the fitted line.

*P16, L12-13: I do not agree that all the instruments showed consistent results. TLS mainly derived dependence for the better performing instruments... Also the authors have not demonstrated that the instruments improved their pointing performance as a result of any of these calibration methods.*

The reviewer is right, the TLS fit weights the different data points with respect to their individual errors. However, Figures 24 and 25 show that the results from many of the instruments scatter around the regression line, except 1-2 worse performing instruments (for which the lower weighting in the regression is justified). In summary, given the uncertainties (which are discussed in the paper) and the different instrument performances, we think that this statement is still justified. We replaced the word “all” by the word “most”. Regarding the second point: This is true. However, as mentioned above, the focus of this paper was on the details of the elevation calibration. The assessment of the performance during the measurement campaign is described in Kreher et al. (2019).

*P16, L33: I would not call this method accurate since some of the instruments clearly showed asymmetric FOV and different functions other than Gaussian could describe the intensity distribution potentially leading to larger errors.*

Fit errors were added to all plots showing Gaussian fits. They indicate that the standard fitting errors are rather small despite the fact that some intensity distributions do not strictly follow a Gaussian shaped curve. As one conclusion of this study it was found that the centre of the Gaussian is a rather robust measure which might be different for the other parameters which are fitted (e.g. amplitude of the Gaussian). This becomes clearer when comparing the fitted centres with the centres retrieved from a centre of mass approach (see section 3.2.2). Nevertheless, the sentence was changed to “Furthermore, this method is very accurate and precise as long as the instrument has a mostly symmetric FOV”.

*P17, L1: Precision should not be confused with accuracy. In my opinion the authors have not accounted for all the uncertainties to claim accuracy of  $\pm 0.05$  deg.*

As mentioned previously (answers to major comments 3 and 7) the systematic errors are dominated by the errors of determination of the lamp position and the fit error. However, for the setup and the measurements used in this study this estimate is justified. But the reviewer is right that for different locations the errors might be larger and therefore, the method might be less accurate.

*I recommend combining Table 1 and 2 and adding 2 columns with positioner maker, accuracy and precision data.*

The tables were merged in the revised version of the paper.

*I do not think Tables 3 and 4 are needed.*

We prefer to keep these two tables in the paper to provide a clear summary of the results. Note that these tables are small and don't consume much space.

*Table 5: I think replacing "row" with "container level" might be clearer.*

We agree and adjusted the table accordingly.

*Figure 2 and 3 should be combined.*

We prefer to keep these two figures separately as explained above.

*Figure 3: Upper and lower "panels" instead of "parts".*

The caption and the text were adjusted accordingly.

*Figure 4 is unnecessary*

We prefer to keep this figure as explained above.

*Figure 10: It might be good to raise MAX-DOAS instrument in line with the stripe*

Thanks for this good suggestion. The figure was modified accordingly.

*Figures 12, 13, 14, 15 should be combined to give better idea about apparent FOV for different instruments.*

We agree and combined the figures.

*Figure 16: intensity distributions in (a, b) clearly do not have Gaussian distributions, please add fitting errors.*

Standard fitting errors were added to all plots showing Gaussian fits.

*Figures 24 and 25: Y-axis = Apparent horizon elevation; X-axis = apparent lamp elevation. Add fitting errors to slope and intercept.*

The two figures were modified accordingly. Further, the standard error of the mean lamp and horizon elevations are now shown instead of the standard deviations.