4   Discussion

Analysing the results obtained in the comparison we see that the uptime of both instruments is comparable. At 80-90%, the miniDOAS uptime is adequate for an instrument in an automated monitoring network. We expect that the uptime of the miniDOAS will further improve in 2016, as from then on the instrument will benefit from the regular monitoring of the network performance.

4.1   Timescale of the intercomparison

Both instruments provide minute values of ambient ammonia concentrations. When looking at short timescales (minutes, hours) we see relatively large differences between the datasets. The differences get smaller as the timescale gets longer. When we look at the fits in the scatterplots of hourly, daily and monthly averages (Fig. 11, Fig. 12 and Fig. 16, respectively) we see that the slopes approach unity: 1.54 for hourly averages, 1.27 for daily and 1.03 for monthly averages. The offsets approach zero, from -7.34 μg m^{-3} for hourly averages, via -3.06 μg m^{-3} for daily to 0.65 μg m^{-3} for monthly averages.

We conclude that, on a timescale of minutes or even hours, the instruments do not compare well. This is caused by a distinct difference in temporal resolution: the typical integration time of the miniDOAS is 1 minute, its minute-measurements are delay-free and mutually independent. The AMOR has a much larger response time, despite its claimed temporal resolution of 3 minutes. Its response to abrupt changes shows a delay (order of 30-60 minutes) and a spread-out and flattening of short peaks. In general, the integral over time of the AMOR-observed ammonia seems to remain conserved, as is reflected by the good comparison of longer timescale averages discussed above. This means that (virtually) no ammonia is lost in the AMOR, but it will be recorded at a different moment in time than its actual appearance at the AMOR inlet.

On timescales of hours, e.g. when looking at daily cycles, we consider the miniDOAS concentrations to be more representative for the actual ambient ammonia concentrations than the AMOR measurements.

We therefore focused our comparison on longer timescales: daily and monthly values. Daily value pairs showed good agreement in a direct comparison, i.e. when the concentration values are plotted in the same graph (see e.g. Fig. 13). The smoothing and delay effects that are apparent in the minute and hourly values have largely disappeared. However, scatter plots (see e.g. Fig. 14) show still some deviations from y = x, indicating that some delay effects are still not smoothed out. This is to be expected, a high peak just before the transition to a new day will cause differences in two consecutive days. In the monthly averages, because there are far fewer transitions, such extremes have disappeared. This makes monthly averages the timescale of choice for the intercomparison.

The monthly averaged concentrations show a linear relationship, as indicated above. We conclude that for monthly averages the instruments compare well. Over the whole comparison period there is an average offset of 0.65 ± 0.28 μg m^{-3} and a slope of 1.034 ± 0.028 between the techniques. Thus, the miniDOAS measures on average slightly higher than the AMOR, over all concentration ranges.

From a scientific point of view this correspondence is excellent, especially since two completely different measurement techniques are used. As a reference, we refer to a study by Von Bobrutzki et al. (2010) that shows much larger discrepancies between different techniques. The systematic difference found between AMOR and miniDOAS amounts to roughly 10% of the typical ammonia background concentrations in the Netherlands of around 5 µg m^{-3}. Fortunately, not so much the absolute difference between the techniques is politically relevant but any jumps in ammonia trends, see e.g. van Zanten et al. (2017) or Wichink Kruit et al. (2017) for two studies in which this data is being used. We will discuss some possible explanations for the difference between the techniques.

## 4.2   Possible explanations for the difference between the techniques

A possible effect of the difference in measurement altitude (the AMOR measured at 3.5 m, the miniDOAS at 2.2 m) was studied using passive samplers. As discussed in Sect. 3.2, the results show no significant difference between AMOR inlet and miniDOAS path, so they offer no explanation for the observed positive bias between miniDOAS and AMOR. Further research with more precise equipment would be needed to reduce the statistical error in these measurements and study the effects of the altitude difference between 2.2 and 3.5 m, as well as the possible influence of the station housing.

The AMOR validation procedure may influence the intercomparison as well. AMOR data is validated based on concentration data only. No AMOR instrument parameters are included in the validation procedure. Closer inspection of validated AMOR values show periods after maintenance where values approved in the validation procedure may - in hindsight - be considered too low and erroneous. This conclusion can only be drawn when the miniDOAS dataset is used as an additional validation tool. Removing these data from the comparison (dubious as it would be from a scientific point of view) would however improve the comparison only slightly. Therefore, the validation procedure can be ruled out as a major source of the offset.

Another effect to be looked at is a possible loss of ammonia in the AMOR air inlet system, as this is a known effect in ammonia inlet lines (Yokelson et al., 2003). However, the AMOR air inlet system has been designed to minimise such effects. Especially the relatively high airflow through the instrument, of 25 L min^{-1} rather than the mL min^{-1} flows found in other instruments, should be effective in minimising these effects. As discussed in Sect. 3, no indication for ammonia loss was found in the measurement data. It seems therefore unlikely that ammonia loss is a major contributor to the bias found.

The AMOR calibration procedure should come up for scrutiny as well. AMOR calibrations are performed using calibration fluids, and thus only pertain to the 'liquid' part of the instrument, after ammonia has been absorbed in the denuder. Any losses in the airborne phase, e.g. in the inlet system, are not included in the calibration procedure. As stated previously, the reason for omitting this part in the calibration procedure is that it is virtually impossible to generate an adequate calibrated gas flow, as the AMOR tries to minimize inlet effects by using a very high airflow of 250 m^3 h^{-1}, from which a further 25 L min^{-1} is sampled by the instrument. We have not been able to study this aspect further in the framework of this comparison.

Shifting our attention to the miniDOAS, we note that the miniDOAS zero is determined by comparison to a DOAS reference instrument. Any offset in the reference instrument will show up as a similar offset in the reported miniDOAS values. The zero of the reference instrument is determined by study of a long time series, looking for periods of lowest values and assuming these occur at constant zero ammonia levels. If this assumption is incorrect, it results in the reference instrument underestimating the real concentrations. This would therefore lead to a negative bias in the concentrations reported by the miniDOAS, never to a positive one. There is no evidence for this in the dataset.

5

17
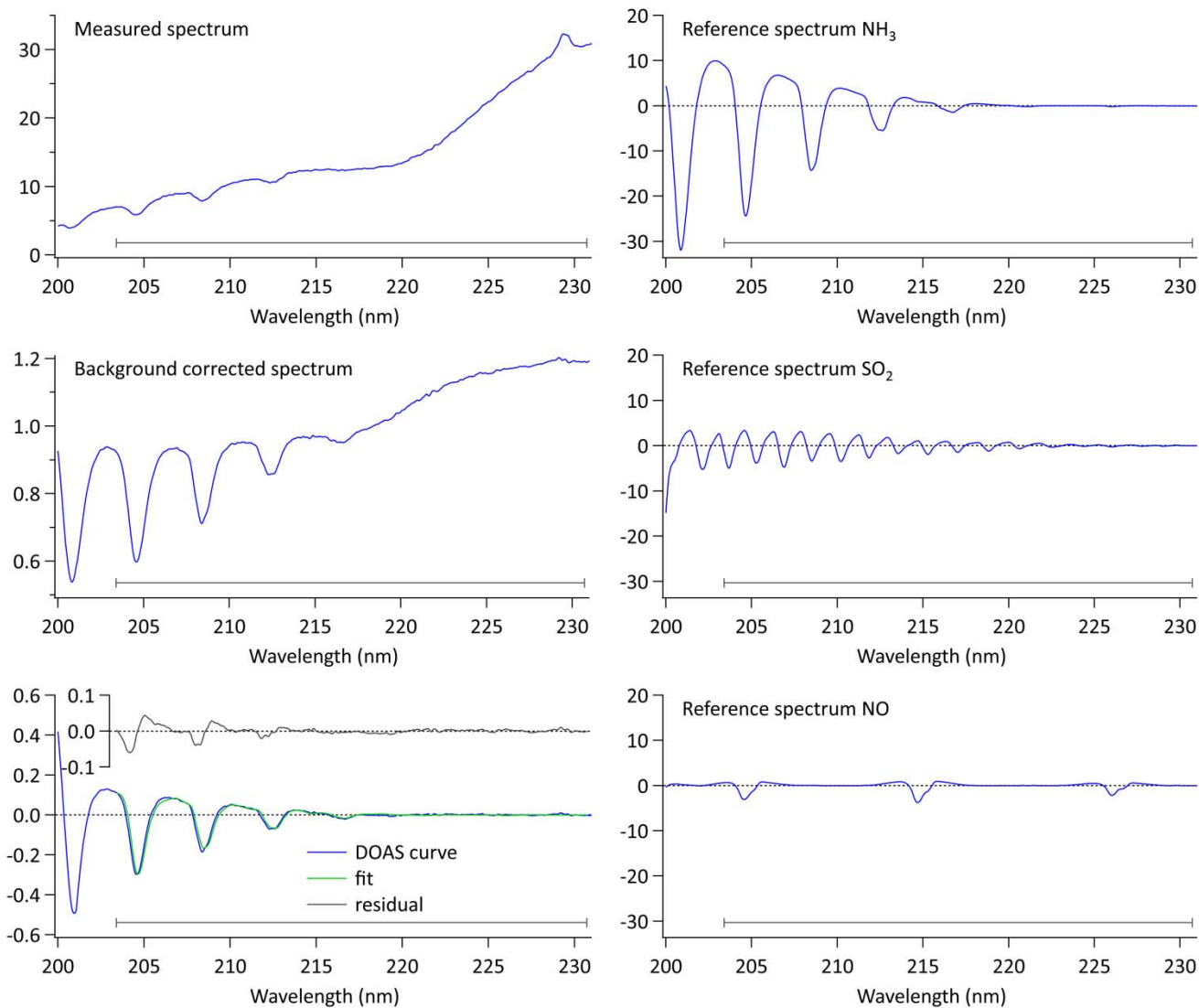
Figure 5: Spectral fit and concentration retrieval shown for a typical measurement. Left from top to bottom: the measured spectrum $I_{meas}(\lambda)$ (1 minute average); the background corrected spectrum $I_{bgc}(\lambda) = (I_{meas}(\lambda) - I_{dark}(\lambda)) / I_{background}(\lambda)$ (see Eq. 7); the DOAS curve $DC(\lambda) = \ln(I_{bgc}(\lambda) / [I_{bgc}(\lambda)]_{moving\ average})$ (see Eq. 4), the fit to the DOAS curve $fit(\lambda) = \alpha \cdot X(\lambda) + \beta \cdot Y(\lambda) + \gamma \cdot Z(\lambda)$ (see Eq. 5) and the residual $residual(\lambda) = DC(\lambda) - fit(\lambda)$. Right from top to bottom: NH_3, SO_2 and NO reference spectra used in the fitting procedure. Scales for the y-axes are arbitrary. Units for the y-axes of the reference spectra are the same. The residual is a tool that helps to identify potential flaws in the measurements, such as interference problems. The dark grey lines in the bottoms of the graphs show the analysis interval used. The measurement shown here was taken at Vredepeel on 23 January 2015, 6:36 UTC.