

Response to Reviewers

Electrochemical sensors onboard a Zeppelin NT: In-flight evaluation of low-cost trace gas measurements

Referee comments are in **black** and authors responses are in **blue**

Reviewer #2:

I think it is a really good platform for low-cost sensors to measure the boundary layer. Please find some comments and questions below about the manuscript and methods that were used.

We thank the reviewer for the positive feedback. Response to each comment is provided below.

Time synchronization

What is LTE?

We added the abbreviation in section 2.1 as follows: “Long Term Evolution (LTE) – a standard for wireless broadband communication for mobile devices – equipment for remote access, [...]”.

Did you apply the Savitzky-Golay filter before or after you had applied your temperature correction? How noisy were the raw data versus after this filter had been applied. Did the analyte peaks stand out prominently from the noise?

We applied the Savitzky-Golay filter before the temperature correction. We now added two new figures in the supplement to address your comments. Fig. S2 shows the noise distributions of the background signals before and after applying the Savitzky-Golay filter. Fig. S3 indicates the range of values that is reduced. For that, we calculated the differences of the sensor signals before and after applying the filter. The reduction is within 10 ppb (2σ , corresponding to around 95 % of the data), i.e. similar to the noise levels shown in Fig. S2. Therefore, larger peaks above the detection limit of the sensors, for example as shown in Fig. 7, are not affected by the filter. This is now also highlighted in Sect. 2.3.1:

“Fig. S3 shows the difference between the sensor signal before and after applying the filter at different NO and NO₂ concentrations. The signal reduction is within 10 ppb (2σ , corresponding to around 95 % of the data), i.e., similar to the noise levels of the sensors (Fig. S2), independent of the NO and NO₂ concentrations, which highlights the minor influence of the Savitzky-Golay filter at larger peaks.”

Determination of amount fractions

Not quite sure how you reached a window size of 31 seconds for the t90 value. Did you run the EC sensors in the laboratory under controlled conditions and then change the [NO] and [NO₂] to concentrations that you'd expect to see on the Zeppelin and calculate it from there?

Was Mead et al 2013 also using these sensors on a Zeppelin or aircraft? Was the concentration range in that paper the same as you'd expect for your deployment?

You use the fact that Mead et al 2013 found a t_{90} value of 21 sec for NO₂. What was Mead et al t_{90} value for NO? The manufacturer provides vastly different t_{90} for NO₂ and NO. Also do we expect that the EC sensor technology hasn't changed in ten years so you and Mead et al have the same versions of EC?

That is a great point. We now added our laboratory results in Table S2 that show t_{90} of $34.6 \text{ s} \pm 2.0 \text{ s}$ and $28.1 \text{ s} \pm 1.6 \text{ s}$ for NO and NO₂, respectively, in agreement with the manufacturer and published measurements. The following is also added in Sect. 2.3.2:

“This value is derived from the combined information given by the laboratory measurements (Table S2), in agreement with the manufacturer and by published measurements (Mead et al., 2013): [...]”

Sensor signal dependences

So you assume that an increase in NO or NO₂ from 0 ppb to 2 ppb does not cause any increase of EC. Did you check this in the lab?

The limit of detection (LOD) defined as three times the standard deviation of the baseline ($3 \times \sigma$) is 6.0 ppb and 10.5 ppb for NO and NO₂, respectively. Therefore, 2 ppb is below the detection limit of the ECS and we consider it a valid approximation here.

Why did you not find the median NO signal (of the six NO EC) and subtract that from each NO sensor to get the ‘signal primarily influenced by other factors’? I think that would be more appropriate rather than assuming a step change of the analyte concentration doesn't impact the EC output at all.

We now rephrase this sentence to better clarify the steps followed here: “This procedure reduces the data to only background concentration periods and provides the signal that is influenced by cross interferences.”.

Why did you correlate WE to the change in humidity and not the humidity? Or if you wanted to use the differential, why not use that for temperature also, dT/dt ? Changing the way these are plotted makes it difficult to see quickly the difference between the temperature and the humidity dependence.

We added Fig. S7 in the supplement and highlight that we could not detect any impact from dT/dt in Sect. 3.1.

Validation of ECSs performance

I think there needs to be a timeseries of the 6 NO/NO₂ sensors showing how they all behaved over the duration of the flights. The slopes look good but did they start out correlated and then drift apart as they all respond differently to each interfering variable? Or were they all slightly offset and this stayed consistent throughout? Maybe add temperature and humidity onto this too, with a shared time axis.

We added multiple time series of amount fractions, T and RH for different dates to the supplement (Fig. S11) and a discussion about the stability in the main text (Sect. 3.2.3). We found a constant offset for setup #5 that was also shown in Fig. S10.

Comparison of ECSs with the MIRO MGA

Accuracy plots are nice but please add the x-axis. Is this over time? In which case the accuracy decreases with the T,AUX and dRH/dt correction towards the end of the experiment. Or is it a shared axis with the

plot below – so accuracy decreases at higher NO_x (MIRO) concentrations? That seems a little odd. Why would the accuracy decrease if the reference instrument is measuring more than 125 ppb NO_x? I think if this is the case a little description of why that might happen is needed.

It is a shared x-axis. We now added the spline and ticks for the accuracy plot axis and colored the scatterplot by the number of flight measurements performed per bin. Because of the low number of data points above 125 ppb the variability is relatively high which is reflected by the precision (error bars) and also leads to the lower accuracies observed.

Both the EC NO and NO₂ concentrations are overestimated (even at the higher concentrations you mention a 30 % overestimation). Was this a consistent overestimation throughout? Did you zero the EC amount fractions to the MIRO at the beginning of the experiment? It would be interesting to see if this overestimation changed over a 20 hour flight or over the entire campaign. I was under the impression the EC drift apart over a matter of hours – days.

In Sect. 3.2.3, we state that NO₂ was underestimated for measurements below 25 ppb and NO overestimated for measurements below 40 ppb (see Fig. S9). We did not zero the amount fractions to the MIRO at the beginning of the experiment, because we developed a correction method that is independent of a reference device. As mentioned in one of the previous comments we added a short discussion about ECS stability in section 3.2.3 and multiple time series for different flight days in Fig. S11. We now reference Wei et al. (2018) in the main text who found a drift below 2 ppb/month.

I know you mention that the AUX electrode is supposed to remove all impacts of variables other than temperature but I think you could write a few sentences about the NO_x EC response to other gases. Did you look at how the NO_x EC responded to cross interferences? Was the signal influenced by any of the other compounds measured by the MIRO? i.e. did changes in the ozone concentration cause a response in NO_x EC?

We added a Figure in the supplement (Fig. S12) to evaluate a possible cross interference of the corrected sensor signals to the other compounds measured by the reference device. We could not find any significant cross interference for the atmospheric concentrations that were present. This is now also highlighted in section 3.2.3:

“In this study, measurements of gases are limited to CO, CO₂, CH₄, N₂O, H₂O, and O₃ performed by the MIRO MGA (Tillmann et al., 2022). No significant cross interference was observed under the present atmospheric concentrations, as shown in Fig. S12.”

Conclusions

Deviations of the sensor values to MIRO decreased from an average of 27.3 to 3.5 ppb. What is the accuracy of the MIRO? The concentrations of NO_x observed by this reference instrument seem to be pretty low; the Weisweiler timeseries shows that for this flight the majority of the NO_x was around 5 ppb. A deviation of 3.5 ppb is quite a lot for concentrations in this range.

The precision of the MIRO MGA provided by the company is 0.3 ppb for NO and 0.08 ppb for NO₂. We agree that 3.5 ppb is high in the mentioned background concentration range below 5 ppb. We highlight limitations of the sensors in section 3.3: "These case examples, highlight the potential of the ECS to detect emission sources with concentrations down to 20 ppb. Although, there are larger relative uncertainties due

to the deviations of the individual NO and NO₂ sensors (see Sect. 3) the ECSs show better agreement with the MIRO at higher concentrations"

I think you successfully show that you can improve the accuracy of the NO NO₂ EC sensors by correcting for temperature and humidity but I am not sure you could reliably deploy EC sensors in the Zeppelin with no reference device and trust the output. It does seem to qualitatively pick up the NO_x emissions from the power plant and industrial regions but there is still large amounts of uncertainty (below 5 ppb deviations between MIRO and EC increased up to 300 % and -600 %; the MIRO reported ambient NO_x mixing ratios of 5 ppb for the majority of the time).

We agree with the reviewer that low-cost sensors aboard the Zeppelin can only be used for the detection of significant emission plumes. We further highlight this by adding a colorscale in Figure 6 that shows that the majority of ambient NO_x concentrations is in the range below 20 ppb and further discuss it in the conclusions:

"However, at lower concentrations below 40 ppb and 25 ppb average deviations of 34.4 % and -31.3 % for NO and NO₂, respectively, were evident. Below 5 ppb, these deviations increased up to 300 % and -600 %, because of the small absolute values, indicating the limitations of ECSs for accurate quantification at lower amount fractions. "

On page 11, you also say that you only ended up using sensor set up #2 for the noise plots and intercomparison with MIRO. Really, this means that 1 out of 6 EC sensors can be used on a UAV platform without a reference instrument. Or would you still advise having 6 NO_x EC and then taking the best performing one?

We now include Figure S10 to highlight that the majority of EC sensors showed agreement within 10 % with the reference instrument. The main reason we chose setup #2 for this analysis was the high data coverage during all of the flights that we now highlight in section 2.3.1:

"The clocks on each of the six sensor setups were manually pre-set, therefore time synchronization was not ensured. In the first step we chose a master setup, here setup #2 that was operational throughout all flights with a data coverage of 99 %."