Answer to referee 2 comments for "In-flight emission measurements with an autonomous payload behind a turboprop aircraft"

We would like to thank the reviewer for the suggestions to improve the manuscript. Below, you will find our responses to their comments. The reviewer's comments are written in normal font, and our answers are in italics.

General Comments

The paper describes a new in-flight measurement capability for measuring a range gaseous and particle emissions from aircraft engines. The work describes the first application of the method measuring behind a turboprop. Data from turboprops, both in-flight and on the ground, are rare and this provides an extremely useful dataset to the community. The application of two inlets, allowing for simultaneous in-plume and background, is a really nice feature.

The paper is very well written, with particular attention paid to the detail. The uncertainty in the measurements is well treated. I only have a few minor comments, which are listed below.

We thank the reviewer for this positive assessment and will address the specific comments below.

Specific Comments

In the abstract, define cruise altitude (Line 5)
The definition of cruise altitude is given two sentences later. We moved it now to the first time it was mentioned.

Line 14: do you mean lower than expected NOx based on predicted? Please clarify (based on previous, based on ground-based predictions?). Could it be caused by using H2O instead of CO2 for the calculation (see below)?
As there is no other publicly available TP emission data, the NOx emission index is low compared to known measurements of typical emission~~s~~ indices of modern higher-thrust jet engines, as later referenced in Harlass et al. 2024. It is, however, not an unrealistic value considering that the engine is an older turboprop engine with lower pressure ratios than modern engines.
It is unlikely that the use of H_2O as a tracer caused this, as the relative humidity during the measurement was low enough (Table 1) to consider water vapor an inert tracer. Moreover, an underestimation of water vapor would cause the calculation in Eq. 6 to yield a higher emission index for NO_x.

Line 14, tPM, nvPM or vPM size distributions?
Our setup measured the tPM size distribution. Thank you for pointing it out. We added it to the description here.

Line 29, delete comma after Both

We removed this.

Line 41, do you have a reference to a roadmap or similar to the potential for these new technologies?

We agree that this is needed. We included an FAA and an IATA roadmap as a reference to this statement.

https://www.faa.gov/aircraft/air_cert/step/disciplines/propulsion_systems/hydrogen-fueled_aircraft_roadmap ("Even at projected 3kW/kg by 2035 fuel cells may be best suited for aircraft carrying fewer than 75 passengers and short-haul flights").

https://www.iata.org/contentassets/8d19e716636a47c184e7221c77563c93/aircraft-technology-net-zero-roadmap.pdf

("ZeroAvia plans to deliver a 9-19 seater hydrogen fuel cell powered aircraft in 2025, and a 40-80 seater by 2027.")

Figure 2 is referenced before figure 1 (line 87).

True, we swapped figure 1 and figure 2.

There needs to be a bit more detail on the operational details of the Egrett (operating altitudes, range, science speed range)

The information is added: "The chase aircraft, a Grob G 520 Egrett (Fig. 1), is a high-altitude and long-endurance turboprop aircraft with a certified maximum operating altitude of 13,716 m (45,000 ft) (Grob aircraft SE), a maximum airspeed of 463 km/h (250 kn), and a range of 4260 km (2,300 Nmi) with an endurance of 8.0 hours dependent on payload and weather (NASA Airborne Science Program). Operated by AV Experts LLC, the Egrett was suited to test the instruments and to perform measurements in the near-field exhaust plume (100 - 1200 m) and background atmosphere."

Figure 5b. The model is outside of the error bars of the data exactly at pressures where most of the data is collected. There either needs to be more points added to constrain the fit better or this needs to be incorporated into the error analysis.

Thank you for this comment. We agree that additional measurement points between 250 and 350 hPa would have better constrained the fit in this range. We also acknowledge that the model may not capture the Mixing Condensation Particle Counter (MCPC) behavior perfectly here. The figures below show the subset of the MCMC results with the model median and the 1 σ credible interval. While the model median lies slightly outside the 1 σ range of the experimental data at 300 hPa, it remains within the credible interval from the MCMC posterior. Furthermore, the model considers all data points and their associated uncertainties. Adding more measurements there would not substantially change the fit curve but certainly reduce the uncertainties. We therefore retained the model together with its associated uncertainty, which is propagated in our error analysis to account for this local *deviation.*

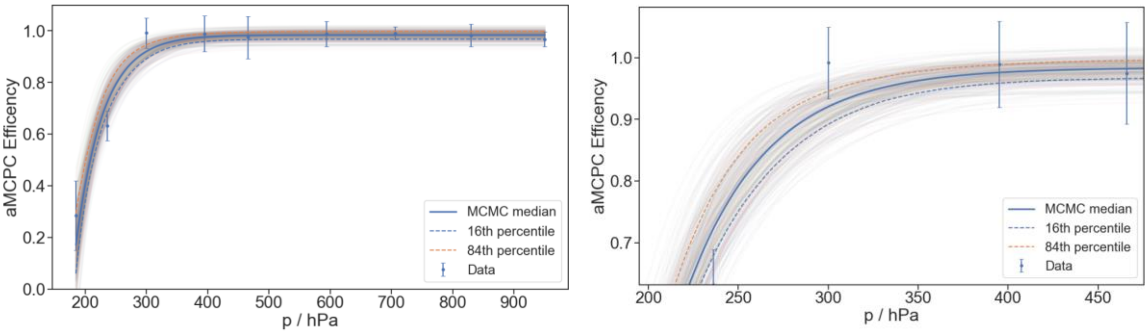

Figure 5d, It is not clear what the data is. The Y axis is labelled as aMCPC, but the figure legend is inlet system losses. Can this be explained more clearly, and which curve or curves are used in the loss correction section?
The Y-label can indeed be a bit confusing. It represents the inlet line losses in addition to the MCPC's cut-off losses. It therefore represents the combined effect of particle-size-related losses, as described in Section 2.1.6. It is now described more specific in the figure caption.

Line 267 – why is there a range of sheath flows? Have you verified this is not changing during a scan? Has this been incorporated into the error analysis (changing the sheath to aerosol ratio changing the resolution of the DMA etc)?
The range given here refers solely to the instrument's capability. Since the transfer function (the probability that an aerosol particle entering the SEMS will exit through the detector outlet) depends on the sheath flow, and a higher sheath flow narrows the sizing window, thereby improving resolution. We selected a sheath flow of 3 L min^{-1} for our measurements, which was constantly used throughout the measurements. This flow is maintained by an internal pump and was verified to be constant in the recorded data, although those data are not shown here.

Line 464 – where does the value of 10% undercounting and subsequent correction come from? That needs clarification.
This is calculated by assuming that the particles follow the unimodal lognormal size distribution that we have discovered in the plume. Applying the losses due to size (aMCPC cut off and inlet line diffusion losses) to this distribution, the difference in the number concentration following from the integrated distribution is ~10%.

The EINOx using the water vapor is an interesting approach. I can see not having CO2 co-located with the Nox at short distances might be an issue as one inlet may be in the plume and the other not. What I would like to see is the EInumber calculated with CO2 and H2O(CR2) from the box A inlet as these should give the same value and give confidence in the EINOx approach based on the author's claim that the WARAN and CR2 agree well.
We agree that the approach of calculating aerosol particle EI with water vapor is an interesting method, which we will apply for later missions. However, for this mission it is not possible to calculate the aerosol particle EI with H_2O as the CR2 is a temperature-based measurement system with a resolution similar to that of the WARAN, but with a slow response to changes in water vapor. It primarily captures limited changes in the H_2O mixing ratio typical of ambient conditions, rather than the rapid variations during in-plume measurements (We refered to these measurement capabilities in section 2.4 "However, it

must be noted that the equilibration time of the frost point measurement at high tropospheric altitudes and low dew points is on the order of tens of seconds.”). The CR2, however, provides more reliable detection of low ambient water concentrations. The statement of good agreement between WARAN and CR2 refers to the ambient mixing ratio.

Section 2.1.8, 4.4 and figure 9 – the paper goes into great detail on the uncertainties in the measurement system, but I do not see that same detail for the mSEMS or the OPC. For the mSEMS, the extremely low charging probabilities make quantification challenging. 100 particles at 10nm in dN/dlogDp space over a 5 second scan, corrected for charging efficiency, is an incredibly small number of particles getting to the MCPC detector. Is the variability in the data as shown by the red shaded area in figure 9 really larger than the uncertainty associated with the instrument and conditions (short scan time (smearing), low numbers, charging probabilities, possible changing DMA resolution)? Are the smallest sizes in the distribution a true representation of the PSD?

Thank you very much for this comment.

- We have now included the efficiency losses for small sizes in Figure 9. We also included the efficiency calibration of the SEMS + MCPC setup into the instrument analysis in Fig. 5e with a description in section 2.1.8. This illustrates the combined effect of SEMS and MCPC on the particle cut-off. The calibration uncertainties propagate into the corrected data and show now a realistic estimation for the error at <10 nm particle diameters. While the correction did not change the general appearance of the distribution, it did shift the mode from Dg = 34.7 ± 1.9 nm to Dg = 27.5 ± 2.0 nm.

- The OPC uncertainties were left out in this scope as they are discussed in great detail in Walser et al., 2017, https://doi.org/10.5194/amt-10-4341-2017, or in the PhD Thesis https://edoc.ub.uni-muenchen.de/21664/

Purely out of curiosity, given the relatively simple equations being used, would Monte Carlo simulations be a simpler and more accurate method of calculating the uncertainty rather than the full error equation?

Thank you for this comment. We agree that a Monte Carlo approach could provide a more general treatment, particularly when error terms are not strictly independent or normally distributed. In our case, we chose the analytical formulation because it makes the contribution of each term to the total uncertainty explicitly visible, allowing the reader to see which factors dominate the error budget.