Review of "In situ observations of greenhouse gases over Europe during the CoMet 1.0 campaign aboard the HALO aircraft" – Galkowski et al.

General Comments

The manuscript presents a broad overview and analysis of results of aircraft in situ measurements (continuous greenhouse and trace gases) and atmospheric trace gas and isotopic composition measurements in onboard flask samples conducted during the CoMet 1.0 campaign. It details these two measurement systems and inter-comparisons in the context of WMO compatibility goals and presents model-data comparisons to two commonly used global modelling systems with different resolutions and meteorological drivers. The article is well-written, concise, and scientifically sound, but I feel that the focus of this manuscript stretches the limits of the scope of the AMT journal. As such, and since the article is part of a special issue collection, the majority of my specific comments aim to provide more detail for important measurement techniques below.

Specific Comments

Section 2.2.1 describes the CRDS measurements, but there is some key information and clarification lacking from the description of this system and its operation that should be included here. For example, air sample drying and/or water vapour corrections to these greenhouse gas measurements should be discussed as these measurements are compared to dried flask sample measurements. Furthermore, 'mixing ratios' measured should be reported as dry air mole fractions: please replace 'mixing ratio' with 'mole fraction' throughout the text. It would also be useful to also specify how many in-flight offset-corrections occurred on average per flight as these were performed manually, and specify whether this is a single-point correction or if two calibration tanks were used.

For flask samples, please clarify the level to which air samples are dried with the magnesium perchlorate and the pressure and volume of air that is sampled in each glass flask (is this 1 L at ambient pressures?). On which type of analyser were the flasks measured? L169 states that there was a drift in the CO flask measurements between collection and measurement; for clarity, please elaborate on how you have determined this using the in situ data. L194: Please also clarify how flasks were sampled during vertical profiles; if the aircraft was only either ascending/descending over the flask fill time, the reader can infer that the air samples collected during these profiles represents an integrative mole fraction between potentially several layers of the atmosphere. It might therefore also be useful to state, on average, how long it takes to fill a flask.

Throughout the manuscript, the Jena CarboScope model is referred to as TM3, but in L235, it is mentioned that it will be referred to as "CarboScope" – please choose one or the other for consistency.

It is mentioned in Section 4 that the flask measurements are compatible, by WMO standards, with the G2401 measurements but this does not seem to be the case. The in situ-flask differences, given uncertainties, fall outside of WMO surface compatibility goals for CO2 (0.1 ppm) and CH4 (2 ppb). As compatibility between measurement systems is mentioned as one of the scientific goals of the CoMet 1.0 mission, it seems necessary to state these compatibility goals for CO2 and CH4 as defined by WMO.

Figure 3 does not define the differences shown (i.e. are flask – in situ values shown)? I also wonder why these results are shown by flight number rather than something more informative for understanding differences between the two systems (e.g. difference vs. altitude or difference vs. mole fraction), separating flights 1-7 and 8-9. I would suggest a figure showing these differences as a function of either of these parameters (perhaps in SI) to shed light on why differences are seen between in situ and flask greenhouse gases.

L290: Please describe how it was evidenced that a stratospheric intrusion was crossed.

Figure 4, at times (L292, and elsewhere) is described in kilometers, but the figure itself is denoted in pressure altitude. The text would be more consistent with this figure if it were describing events in pressure levels as well.

Figure 5 shows modeled vertical profiles, which seems redundant with Figure 6. In addition, these model vertical profiles are not ever discussed. I would suggest removing these panels.

Technical Corrections

- L14: Uncertainties are given in parentheses and not quotes, please rephrase.
- L83: Please define 'HALO' if this is an acronym
- L101: G4201-m should be G2401-m, I believe.
- L101: 'fulfil' is missing an 'l'
- L115: Please change 'tollerance' to 'tolerance'
- L200: Please change to 'boundary'
- L222: Please eliminate "when possible" and "when not"
- L244: Change 'these' to 'and'
- L253: Change 'brackets' to 'parentheses'

Figure 2 caption: please define what red/blue shading means. Black crosses are hard to see – perhaps a thicker line would suffice. Altitude should be specified as km ASL.

L273-275 might be more easily understood if (a)-(d) were noted on Figure 4.

Figure 5 denotes CH4/CO/CO2 mole fractions as 'X[]', which is somewhat mis-leading as XCH4/XCO/XCO2 typically denote total-column mole fractions.

L369-370: The offsets in the 3-10 km range are actually responsible for the tail.

Figure 8 caption "For flight no. 7 on the four flasks..." (plural)

Figure S2: Please increase line width for purple crosses, as these are difficult to see.