General comments:
The manuscript “Comparing Airborne Algorithms for Greenhouse Gas Flux Measurements over the Alberta Oil Sands” written by Erland et al., provided a technical aspect of estimating the emissions based on airborne measurements. The authors derived the emissions estimates using two algorithms with and without surface extrapolation and found that the two Algorithms agreed well if the ideal conditions over flights met their assumptions. It is very nice to see that the emissions estimates were derived with several surface extrapolation options and the potential reasons why the emissions derived from two algorithms disagree with each other were also discussed. However, I do have a few concerns, which are 1) why are the spectral imaging methods included in the ms and how do they connect to the comparison of two airborne algorithms? It is not clearly explained in the ms. The authors spent a lot of pages through the whole ms, introducing the spectral imaging methods and comparing the emissions derived from them with those derived from mass balance approach. But I do not see the point. That part is more like another independent story, and 2) the specific background information of two algorithms is not sufficiently provided in the introduction. The dis/ad vantages and assumptions of the mass balance approach and remote spectral imaging method are claimed, but when it comes to the two algorithms, the authors simply say the two algorithms are developed and their comparison is not conducted. It is better to list a few studies using the two algorithms and gave a simple summary about the dis/ad-vantages or explanation of them. Overall, the ms is interesting and in a good quality, and therefore, it is suitable to be published in AMT with a minor review.

We thank the reviewer for their time on this manuscript. The concerns outline in the general comments have been addressed as follows:

1) Spectral imaging methods were included as useful complementary data to highlight the limitations of using only mass-balance box-flight methods to estimate annual emissions from oil sands facilities. We have updated our secondary objective to make this clearer to the reader and included transitional sentences to further incorporate this part.
2) More background information on the two algorithms has been added to the manuscript. Specifically, a paragraph has been added to the main manuscript before Table 1 to better develop concepts discussed. Two sections have been added to the supplement to expand on the table summarizing the two methods. Equations are included and diagrams have been created to contrast how the two methods derive emissions estimates. The introduction also now includes examples of studies utilizing the methods.

Specific comments:
Line 54-56: It is very nice to summarize methods for sampling anthropogenic GHG emissions into two major categories and list the literature as support. Could you separate the literature for each category as well? It will be more clear and handier for readers.

This has been updated to:

“While some airborne methods utilize eddy covariance measurements (Yuan et al., 2015; Wolfe et al., 2018), methods for sampling anthropogenic GHG emissions from the air tend to fall into two major categories: i) mass-balance methods (O’Shea et al., 2014; Gordon et al., 2015; Conley et al., 2017; France et al., 2021; Foulds et al., 2022) and ii) spectral imaging methods (Duren et al., 2019; Tyner and Johnson, 2021; Krautwurst et al., 2021; Cusworth et al., 2022).”

Line 69: “the location of emission sources must be known”. The expression is too absolute to justify. Most mass-balance flights are used to quantify the emissions from the oil and gas exploring facilities, and to quantify urban emissions from a city. The location of emission sources is very clear for the facilities, but in the case of city emissions, sometimes the location of emission sources is not clear. Before designing the flight, it is expected that the emissions would come from the citywide range (containing multiple source types), but the exact location is unknown, depending which species is of interests. The mass balance flights can detect some missing sources that may not be in the inventories.

This line has been removed.

Line 75: “Extrapolation to the ground is often the largest error source” referred to two published studies. The authors should mention in which cases the extrapolation is the largest error source since other studies have also pointed out the selection of background and the evolution of the boundary layer were also two major uncertainties. For example, in the paper “Assessment of uncertainties of an aircraft-based mass balance approach” written by Cambaliza et al., (2014), they did sensitivity analysis of estimated fluxes by changing several factors, including two methods for extrapolating the lowest altitude measurements to the surface measurements. They found the extrapolation influenced less than the background and the CBL.

We agree that there are certainly other large sources of error calculated for the various mass-balance methods. As we are focusing on mass-balance box-flights we have updated this passage to reflect a narrowed focus on that and to outline the fundamental assumptions of the methods more. The update now reads: “For mass-balance box-flights, extrapolation to the ground is often the largest error source, nearing ~30% when the bottom of the plume is not captured (Gordon et al., 2015; Conley et al., 2017). Airborne mass-balance box-flight Mass-balance airborne methods depend on the assumption of a stable boundary layer, and that the emission plume is captured at the top of the box and does not change during sampling (i.e., that conditions are stationary) (Fathi et al., 2021).”

Line 82: delete one of the two verbs? In the sentence “provide two approaches to evaluate calculate mass fluxes”

This has been changed to: “provide two approaches to evaluate mass fluxes”

Line 84-85: “If algorithm comparisons indicate agreement, then emission estimates from multiple campaigns using mass-balance and spectral imaging can be aggregated”, but I do not see the logic here. The two algorithms proposed by Gordon et al., (2015) and Conley et al., (2017) both calculate the emissions estimates using mass balance approach, right? I do not understand why spectral imaging is related to this.

We agree and have removed the inclusion of spectral imaging in this paragraph. While we feel multiple methods will improve the estimation of emissions budgets the inclusion of spectral imaging in this paragraph does not fit.

Line 117: the authors indicated the second objective. It is interesting but is not in line with the title. Maybe the authors could modify the title to contain the information of the second objective. From my perspective, the first and the second objective do not have necessary connections if the scope of the manuscript is what the title conveys.

The second objective has been modified to provide a better connection for the reader as to why the information is included in the manuscript and how it relates to the title. It now reads: "Since mass-balance flights are typically flown with the knowledge of and permission from facilities operators, these methods, while they may be accurate, may not necessarily reflect typical operating conditions or GHG emissions. Consequently, a secondary research objective was to examine the potential of utilizing complementary spectral imaging methods, such as AVIRIS-NG, to supplement mass-balance box-flights by providing contextual information to capture the spatial and temporal variability of oil sands GHG emissions." This independent data provides an important check on the representativeness of the mass-balance sampling.

Line 125-130: change "divergence" to "flux divergence", just to make it clear to the readers? For the description of "Conceptual Algorithm Steps" of SciAv, you mentioned "divergence" for several times and also in the following texts. The unit of divergence profile shows that the divergence indicates flux divergence. To be honest, I was a bit confused the first time I read it, and after reading the paper by Conley et al., (2017) and the following texts, I understand what it indicated exactly.

This has been changed.

Line 220: the subtitle should be 2.2.2

This has been changed.

Line 407: "The SciAv and TERRA estimates were also compared when no surface extrapolation was applied". I did not find the results without surface extrapolations. Provide the results of estimates using two algorithms without surface extrapolation.

This has been changed to: "The SciAv and TERRA estimates were also compared when no surface extrapolation was applied (the background surface extrapolation scenario)."

When applying this to the two different methods applying a background extrapolation meant applying no extrapolation below the lowest profile point which means there is only background mixing ratios for Scientific Aviation and applying a background extrapolation to TERRA. This difference is discussed further in Section 1.4 L160 of the supplement where further testing of the "no surface extrapolation" scenario is explored.

Supplement paragraph L160 reads: "To remove the effect of the surface extrapolation the two algorithms were compared when using each method's "background" surface extrapolation fit. For TERRA this meant fitting the chosen background mixing ratio value below the lowest flight lap, and for SciAv calculating zero divergence below the lowest flight lap".

This was removed from the main text to keep the manuscript as concise as possible.