


# Reply to: Uncertainty and bias in Liggio et al. (2019) on CO<sub>2</sub> emissions from oil sands operations

Received: 11 March 2020

John Liggio   & Shao-Meng Li <sup>1,2</sup>

Accepted: 9 August 2023

Published online: 06 September 2023

 Check for updatesREPLYING TO L. Fu & A. H. Legge *Nature Communications* <https://doi.org/10.1038/s41467-023-40818-5> (2023)

In our paper<sup>1</sup>, we presented an aircraft measurement-based assessment of the CO<sub>2</sub> emissions from the oil sands surface mining sector in Alberta, Canada, and demonstrated that overall CO<sub>2</sub> emissions were 64% higher than reported by the industry. Subsequently, Fu and Legge provided several comments, indicating disagreements with the algorithm utilized (TERRA) and with the approaches used for upscaling emissions to annual values while also identifying a legitimate technical error and a typo in the scale of one figure. In our reply to their comments, we provide evidence that their arguments and comments were a result of incorrect scientific analysis and/or incorrect assumptions.

A starting point of their rebuttal, which appears in their title and throughout their discussion, is that the results of the work<sup>1</sup> are biased. Bias, by definition, implies that the “truth,” or, in this case, the actual amount of CO<sub>2</sub> emissions, is already known, which is not the case. For the oil sands sector, industries report bottom-up estimates, while we have provided aircraft-based top-down emissions estimates. It is not a surprise that the two methods provide different results, but neither method should be tagged as being biased because the truth is unknown. Similar comparisons between top-down and bottom-up approaches have been reported in many scientific publications, and there is now significant evidence that emissions from oil and gas operations globally are underreported in comparison to actual emissions. Further, a recent paper<sup>2</sup>, which used different upscaling methodologies and 15+ years of satellite measurements, shows the same discrepancy between measured and reported oil sands CO<sub>2</sub> emissions as was shown in the previous work<sup>1</sup>. Regardless, we will address each of the arguments of Fu and Legge in the order in which they appear in their rebuttal.

## Issue 1. Technical errors

Fu and Legge assert that a single calculation error (not using the correct CO<sub>2</sub>/SO<sub>2</sub> molecular weight ratio) invalidates the results from the emissions estimation algorithm (TERRA). This is incorrect. The molecular weight ratios of CO<sub>2</sub>/SO<sub>2</sub> were not used in the calculation of the TERRA-derived emissions, nor were they used to directly validate TERRA-derived emissions. TERRA results were validated via direct

comparison of SO<sub>2</sub> to CEMS data as discussed previously<sup>1,3</sup>. The CO<sub>2</sub>/SO<sub>2</sub> molecular weight ratio was used to provide a second and independent method for estimating/scaling up the CO<sub>2</sub> emissions from stacks to compare with the results using the TERRA algorithm. In Liggio et al.<sup>1</sup>, we concluded that the two independent results were comparable within their respective uncertainties.

Further, it is important to realize that this molecular weight ratio error has no effect on Fig. 2 of the paper. Figure 2 is the primary figure in the paper, and it reflects the major conclusions of the paper (which are unrelated to the use of SO<sub>2</sub> upscaling or the CO<sub>2</sub>/SO<sub>2</sub> ratio). The error does affect Fig. 3. Specifically, in Fig. 3d, the green columns will be reduced by 31% so that the green column for SML will change from approximately 14 to 10, while the green column for SUN will change from approximately 7 to 5. The comparison that we were attempting to highlight in Fig. 3d was between the green (stack emissions based on ratios of molecular weight) and orange (stack emissions based on TERRA) columns. We concluded in the paper that they were comparable. With a 31% change (reduction) to the green column, they are certainly less comparable than they were for the SML facility. However, as we have already noted above, the overall CO<sub>2</sub> emissions from the facilities (ground plus upgrading stack) are not largely impacted. Notably, recent satellite-based emissions data<sup>4</sup> have demonstrated that the satellite-observed SO<sub>2</sub> emissions for SML and SUN in 2013 (when the study was done) were approximately 30% higher than reported. This would cause the green bars in Fig. 3b to increase by 30% so that the results shown in the figure would essentially remain unchanged.

We have published a correction to the paper that identifies the molecular weight error and included a revised Fig. 3. However, as noted above, had we taken into account the results from McLinden et al.<sup>4</sup>, the figure would have remained essentially unchanged. Fu and Legge further point to differences between Figs. 2 and 3 as being symptomatic of an overall issue with the use of TERRA since the sum of estimates in Fig. 3 does not exactly add up to that of Fig. 2. The differences between Figs. 2 and 3 are expected, as the application of TERRA to different parts of a given flight will indeed result in slightly

<sup>1</sup>Air Quality Research Division, Environment and Climate Change Canada, 4905 Dufferin St, Toronto, Ontario M3H 5T4, Canada. <sup>2</sup>College of Environmental Sciences and Engineering, Peking University, 100871 Beijing, China. ✉e-mail: [john.liggio@ec.gc.ca](mailto:john.liggio@ec.gc.ca)

different estimates. This is accounted for in the error bars of both Figs. 2 and 3 and will not impact the main conclusion of the paper.

## Issue 2. Upscaling approach using CEMS NO<sub>x</sub> data

Fu and Legge suggest that there is an issue with the approach to upscaling of derived emissions. Their argument is that continuous emissions monitoring system (CEMS) NO<sub>x</sub> data should not be used as a basis for upscaling (similar to using SO<sub>2</sub> data) because they are incompletely reported. They then use the argument of incomplete reporting to suggest that the previous approach<sup>1</sup> was faulty and biased.

However, the original paper<sup>1</sup> did not use CEMS NO<sub>x</sub> data in the manner they are suggesting. This is because CEMS only incorporates NO<sub>x</sub> from a few stacks for each facility while excluding some others and entirely excludes any mobile combustion emissions (of which there are many). That is precisely why we did not use a CO<sub>2</sub> (measured)/NO<sub>x</sub> (CEMS) emission factor approach, as Fu and Legge are suggesting. We agree with Fu and Legge that NO<sub>x</sub> CEMS data are likely to be incomplete. However, this is not relevant, as we are not using it directly to scale emissions of CO<sub>2</sub>. The majority of the analysis provided by Fu and Legge in their rebuttal uses a correlation between NO<sub>x</sub> as measured by CEMS and reported to inventories, versus CO<sub>2</sub> as measured by the research aircraft. We believe that this comparison is invalid considering the limitations associated with CEMS NO<sub>x</sub> reports noted above (i.e., not including all sources of NO<sub>x</sub>), while CO<sub>2</sub> from the aircraft measures the complete CO<sub>2</sub> from all sources within the facility.

Supplementary Figure S4b shows a correlation between NO<sub>x</sub> as measured by the aircraft versus CO<sub>2</sub> as measured by the aircraft (measured at the same times and in the same plumes). The purpose of Supplementary Fig. S4b was to show that there is a high degree of correlation between facility integrated CO<sub>2</sub> and NO<sub>x</sub>. It was not intended to be used as an upscaling approach, as Fu and Legge suggest. The relationship between measured CO<sub>2</sub> and measured NO<sub>x</sub> was important since it was assumed that NO<sub>x</sub> is correlated with the production of synthetic crude oil (SCO). If NO<sub>x</sub> and CO<sub>2</sub> are well correlated, then we can infer that CO<sub>2</sub> is also correlated with the production of synthetic crude oil. Since synthetic crude oil production is reported monthly in inventories, we can scale flight measurements of CO<sub>2</sub> emissions (taken over several hours) to monthly estimates of CO<sub>2</sub> emissions and then correlate estimates of monthly CO<sub>2</sub> emissions to monthly synthetic crude oil production. This is precisely how scaling was done in the paper, and this is well described in the text, including the noted uncertainties of the underlying relationship with NO<sub>x</sub> and SCO production.

Fu and Legge identified a typo in the units of the *x*-axis of Supplementary Fig. S4b (ppm listed, but it should have been ppb). The actual values for the data in the figure are correct (they are ppb), and we acknowledge the *x*-axis caption contains a simple typo, although it has no impact on any results reported.

While there may be different approaches to upscaling measurements, what Fu and Legge are suggesting is not credible because the TERRA emissions and the CEMS emissions are not fully comparing the same sources. Supplementary Figure S4b provides the groundwork for a more rigorous analysis of measured CO<sub>2</sub> compared to measured NO<sub>x</sub>, followed by upscaling by total oil sands NO<sub>x</sub> emissions (from all sources). This is beyond the scope of this paper.

## Issue 3. Perceived background procedure deficiencies and inconsistent application

Fu and Legge assert that the background subtraction performed by Liggio et al. was not consistently applied. They seem to be confusing “constant extrapolation” below the lowest flight altitude with “background subtraction” in general. The constant extrapolation was used below the lowest flight track to account for any CO<sub>2</sub> that was not directly measured. Previous work has already demonstrated that this extrapolation is the largest source of uncertainty in TERRA, but the

differences between extrapolation methods are small and, nonetheless, are included in the uncertainty estimates. Background CO<sub>2</sub> was determined differently and was not a constant extrapolation. The background levels of CO<sub>2</sub> for any given flight did exhibit minimal variability (0.5–3 ppm, as noted in the paper) relative to the enhancement caused by oil sands emissions (>80 ppm in some flights). Further, the premise of the developed background subtraction algorithm used in the paper was that it was free of bias and, in fact, applied consistently for every flight, contrary to their assertion. In addition, the impact of the application of a conservative varying of the background was conducted ( $\pm 1\sigma$ ) so that the impact of the background subtraction on the final emissions could be included in the uncertainty analysis. It has already been shown in the paper that significantly varying the baseline CO<sub>2</sub> had only a small effect on the uncertainty of the final results<sup>1</sup>.

Additionally, Fu and Legge chose data from a specific flight (F6) and then suggested that it shows inconsistent background levels for CO<sub>2</sub>, issues with meteorology, and emission below the lowest flight track. Their argument in this regard is unclear, as the plumes for SUN and SML were detected at different downwind location, and on different screens, such that they have minimal influence on each other. Further, the enhancement in CO<sub>2</sub> is well above any nearby background CO<sub>2</sub> determined. Fu and Legge also state that 2/3 of the SML emissions within this flight (F6) are below the lowest flight track. This is of no consequence, as the SI figure shows concentrations, not emissions through each screen, which are dependent on wind speeds (i.e., 2/3 of the concentration may remain a small emission flux). Overall, below-flight track concentrations/emissions are already a well-known uncertainty in TERRA (and all other aircraft mass balance methods), and it is clearly accounted for and described in our paper. As such, their text of section 3 of their “Matters arising” does not provide any additional insight into this well-known issue.

Fu and Legge also suggest that pre-buildup (i.e., “storage and release events”) of emissions adversely impacted the flights in this work. This is a significant oversimplification. The process of “pre-buildup” has recently been investigated<sup>2</sup>. That paper demonstrated that the contribution of “storage and release” events is likely to be minimal (<2–17%) for certain meteorological conditions and within the overall uncertainty of the TERRA approach. Fu and Legge suggest that emissions from SML adversely impacted the emissions for flights F14 and F7 and cite the Fathi et al. paper as a reason to reject this flight, but they provide no supporting evidence. The Fathi et al. paper<sup>3</sup> is a modeling study using model meteorology that is designed to demonstrate the types of conditions required for such events to occur. It does not imply they have actually occurred. Regardless, emissions on those days are not significantly different than emissions on other days for the same facilities, and hence there is no clear justification for removing any flights. Regardless, the exclusion of one flight (with a similar emission value) does not impact the results of this work.

Any top-down measurement of emissions comes with a number of associated uncertainties which is well established in the literature on this topic. In our paper, we have specifically discussed several quantifiable and unquantifiable uncertainties, including those associated with: (1) emissions below the lowest flight track, (2) background levels of CO<sub>2</sub>, (3) the proportionality of NO<sub>x</sub> to oil production, (4) the variability across flights, (5) use of scaling factors, (6) oil production volumes, (7) the inclusion of methane into the GHG estimates, (8) the impact of local vehicle emissions, and (9) emission seasonality. These uncertainties, and various others, are already described in the original paper under “Other considerations and assumptions” in “Methods.” While we did not mention meteorology at the time, we acknowledge that the impact of meteorological conditions is an important factor in designing aircraft studies for the purpose of emissions estimation. Fu and Legge also suggest that the TERRA code was not available to them to verify our results. This is incorrect, as it is clearly stated in the published paper, although it has not been requested by Fu et al.

Overall, the rebuttal of Fu and Legge is problematic, as many of their assumptions are not supported, and they provide numbers that are not backed up by calculations or references, and some information is taken out of context. The one legitimate item that they identified is related to a calculation error that affects Fig. 3d in Liggio et al.<sup>1</sup>, and for which we have already published a correction.

### Data availability

No new data was used for this reply. The authors declare that the data supporting the findings of the original study are available within the original paper and its Supplementary Information files. Should any raw data files be needed in another format, they are available from the corresponding author upon reasonable request.

### References

1. Liggio, J. et al. Measured Canadian oil sands CO<sub>2</sub> emissions are higher than estimates made using internationally recommended methods. *Nat. Commun.* **10**, 1863 (2019).
2. Wren, S. N. et al. Aircraft and satellite observations reveal historical gap between top-down and bottom-up CO<sub>2</sub> emissions from Canadian oil sands. *PNAS Nexus* **2**, pgad140 (2023).
3. Gordon, M. et al. Determining air pollutant emission rates based on mass balance using airborne measurement data over the Alberta oil sands operations. *Atmos. Meas. Tech.* **8**, 3745–3765 (2015).
4. McLinden, C. A. et al. Inconsistencies in sulfur dioxide emissions from the Canadian oil sands and potential implications. *Environ. Res. Lett.* **16**, 014012 (2021).
5. Fathi, S. et al. Evaluating the impact of storage-and-release on aircraft-based mass-balance methodology using a regional air-quality model. *Atmos. Chem. Phys.* **21**, 15461–15491 (2021).

### Author contributions

J.L. wrote the original draft manuscript. J.L. and S.-M.L. edited the manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

**Correspondence** and requests for materials should be addressed to John Liggio.

**Peer review information** *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© Crown 2023