The manuscript addresses an important scientific issue and presents an innovative approach to assess CO2 emissions from urban areas. However, many parts in the preprint suggest that it is work in progress. The study holds potential but necessitates substantial revisions to strengthen its clarity as well as scientific rigor.

General comments:

The manuscript should make a clear distinction that its primary objective is not the inversion process itself, but rather the assessment of the confidence in the inversion results via predictable and diagnostic variables.

In the methodology section, more emphasis could be put on how the two-step procedure via predictable and diagnostic variables contributes to the accuracy assessment of emission estimates.

The manuscript heavily relies on Danjou 20xx, which is not accessible for verification. I think this is problematic and impedes a thorough evaluation in parts of this study.

The manuscript lacks clarity on what constitutes the error. The reference data used for determining the error is not explicitly defined.

The manuscript does not provide validation of its results with actual satellite data (e.g., OCO-3 SAA) or ground truth measurements. The absence of this validation makes the reliability of the simulated results vague to some extent.

The variability of the error distribution remains significant across different cities. What are the implications for estimations from real satellite images?

The study should discuss the detection limits of current and future satellite missions and how those might impact the results.

Is the purely random noise model imposed on XCO2 data in the study representative of real-world atmospheric and environmental conditions? In actual scenarios, factors like surface reflectivity of different land types and the presence of aerosols can introduce more structured or systematic errors rather than purely random ones. Would incorporating more realistic, structured errors enhance the model's applicability and accuracy in real-world urban CO2 monitoring scenarios? Given this potential limitation, what implications does this assumption have for future research?

The selection criteria for the size of the target emission zone radius, are not comprehensively described. I can't find a clear rationale for the chosen size of the emission zone radius. Any potential to estimate this radius through an inversion approach?

The influence of cloud coverage on satellite observations is a significant factor that the manuscript should address.

The authors of the manuscript should revise Section 3 to summarize only the key aspects of the Danjou 20xx study that are directly relevant to their current research. As mentioned above, it is problematic that the Danjou 20xx paper is not yet available and that the study heavily relies on it.

The description of the OLAM model and its simulations should be more comprehensive in order to better understand the simulation results.

The manuscript acknowledges that the method is not universally applicable to all cities and is still a work in progress. However, it should also describe how the proposed method would be applied to real satellite measurements and critically assess, whether its applicable at all within the error budget.

The conclusion should reiterate the study's focus on developing a procedure for accuracy estimation, summarizing how the study advances this goal and its implications for future research and practical applications.

Specific comments

The manuscript frequently uses abbreviations and technical terms without defining them (e.g., XCO_2, UNFCC, OCO, ppm, GOSAT-GW, WRF, OLAM, IQR).

3: The phrase "selecting images to be processed" should be more clearly defined.

Fig 1: Identifying the factors behind the peak XCO_2 values in the simulation domain deyond the city boundaries.

46: The Danjou et al. (20xx) should be made available to the reviewers due to its significant relevance to the research.

53: Clarify local background signal (not clear to me, what is meant by this term).

56, 58: "This study" can be misleading. Write "The study".

48-50: Consider moving the sentence up to line 25, before you start describing the OCOs.

261: Set by the user? Do all produce the same value?

405-406: Please clarify why Paris exhibits such a low emission rate? Is this due to the absence of significant point sources?

418: What are the implications or consequences for the study if the dependencies of errors are not completely comprehended, even when using synthetic data.

Fig. 5: Typo in y-label (thrue).

316: The GP2 inversion method is presumed to be a variation of the Gauss Plume approach. However, at this point it is unclear for me what the '2' in 'GP2' represents. Please provide further details or clarification.

520: Was the 12° threshold for wind variability found empirically or is there a rationale behind choosing it? Is this higher variability threshold, compared to what's mentioned in Danjou 20xx, a result of the increased resolution in the model?

578: The error appears to be highly sensitive to the city's radius. Could you clarify what "pseudo-image filtering" specifically entails? I don't have access to the Danjou 20xx source for reference. Additionally, referring to line 550 and following, I guess it implies filtering out scenes with variability above a certain threshold?

58: What are pseudo-images? Synthetic 2D CO2 concentration images, I guess?

229: What is the primary source of error in the emission estimates? Does it originate from the tree model or the GP2 inversion method?

47: Improve the English language in this sentence. It encapsulates the primary motivation of the study and thus should be distinctly emphasized and articulated.

76: The frequent citation of Danjou 20xx for all the "light" methods seems inappropriate. It would be more suitable to refer to the original papers specific to each approach.

74-79: The logic presented in this paragraph is unclear. It would be beneficial to revise and clarify the content for better understanding.

Sec. 2: Mention that ECMWF ERA-5 product is used for meteorological data. It is only mentioned in line 446.

130: What is the methodological rationale behind assuming constant emission rates? Do you expect that incorporating dynamic emission rates would significantly alter the study's results or conclusions?

217-221: Does the observation that all methods yield similar results suggest that the assumptions in the model used to generate the synthetic data might be overly simplistic?

Sec. 3: Does this paragraph solely describe the work done in Danjou 20xx? Certain statements, like in line 160 "The method used here...", are ambiguous in the current context. Does "here" refer to this study or to Danjou 20xx? It may be beneficial to condense the paragraph preceding Eq. (1) for clarity.

217-221: Consider clarifying this point earlier in the document, perhaps where the Gaussian plume inversion (Eq. (1)) is introduced, for better coherence and understanding.