Dear Dr. Rayner:

We thank you for your thorough review of our manuscript. We provide our responses to your comments in italicized text. We partition the text in the general comment section of your review to extract questions raised by you and then answer them in sequential order. We also answer the specific comments in sequential order but we restart the numbering to match with the numbering of the specific comments. The changes made in light of your comments are identified by the label Action and follow the Response label.

General Comments

1. **Reviewer Comment:** This paper presents a series of quantities that can be derived from linear inverse theory. Put roughly they are the similarity of footprints (nearly the independence of rows of the Jacobian), the local sensitivity of the result to various inputs and finally a global sensitivity using a first-order Taylor expansion with respect to all inputs.

   Response: *This comment summarizes the study.*

   Action: *No action required.*

2. **Reviewer Comment**: The metrics are potentially useful and some are, to my knowledge, novel.

   Response: *We thank the reviewer in appreciating our effort and the contribution of our work.*

   Action: *No action required.*

3. **Reviewer Comment**: The paper is potentially in scope for AMT though I think it needs more work to make it more relevant to likely readers.

   Response: *We think our manuscript is appropriate for GMD as:*

   - *GMD Journal has previously published papers on atmospheric inverse modeling covering a wide variety of topics, and*
   - *Our manuscript falls within the scope identified by the editorial board of the Journal. Please see our response below:*

   List of a few papers among many that have been published in GMD that deal with atmospheric inverse modeling. Some of these are specifically focused on method development

   - Cho, T., Chung, J., Miller, S. M., and Saibaba, A. K.: Computationally efficient methods for large-scale atmospheric inverse modeling, Geosci. Model Dev., 15, 5547–5565, https://doi.org/10.5194/gmd-15-5547-2022, 2022.
   - Vojta, M., Plach, A., Thompson, R. L., and Stohl, A.: A comprehensive evaluation of the use of Lagrangian particle dispersion models for inverse modeling of greenhouse gas emissions, EGUsphere [preprint], https://doi.org/10.5194/egusphere-2022-275, 2022.
   - Liu, X., Weinbren, A. L., Chang, H., Tadić, J. M., Mountain, M. E., Trudeau, M. E., Andrews, A. E., Chen, Z., and Miller, S. M.: Data reduction for inverse

modeling: an adaptive approach v1.0, Geosci. Model Dev., 14, 4683–4696, https://doi.org/10.5194/gmd-14-4683-2021, 2021.

- Miller, S. M., Saibaba, A. K., Trudeau, M. E., Mountain, M. E., and Andrews, A. E.: Geostatistical inverse modeling with very large datasets: an example from the Orbiting Carbon Observatory 2 (OCO-2) satellite, Geosci. Model Dev., 13, 1771–1785, https://doi.org/10.5194/gmd-13-1771-2020, 2020.
- Hase, N., Miller, S. M., Maaß, P., Notholt, J., Palm, M., and Warneke, T.: Atmospheric inverse modeling via sparse reconstruction, Geosci. Model Dev., 10, 3695–3713, https://doi.org/10.5194/gmd-10-3695-2017, 2017.
- Miller, S. M., Michalak, A. M., and Levi, P. J.: Atmospheric inverse modeling with known physical bounds: an example from trace gas emissions, Geosci. Model Dev., 7, 303–315, https://doi.org/10.5194/gmd-7-303-2014, 2014.
- Chai, T., Stein, A., and Ngan, F.: Weak-constraint inverse modeling using HYSPLIT-4 Lagrangian dispersion model and Cross-Appalachian Tracer Experiment (CAPTEX) observations – effect of including model uncertainties on source term estimation, Geosci. Model Dev., 11, 5135–5148, https://doi.org/10.5194/gmd-11-5135-2018, 2018.

Is the submitted manuscript within the scope of the journal?

Our paper covers two focus areas identified by the Journal

- *Development and technical papers, describing developments such as new parameterizations or* ***technical aspects of running models*** *such as the reproducibility of results;*
- *New methods for assessment of models,* ***including work on developing new metrics for assessing model performance*** *and novel ways of comparing model results with observational data*

Does it contribute anything new to the field or atmospheric inverse modeling?

In this study we provide:

- *Analytical expressions to conduct post hoc (that is after an inversion has been performed) local sensitivity analysis by computing partial derivatives*
- *Demonstrate a scientifically interpretable framework for ranking thousands of spatio-temporally correlated input parameters with same or different units.*
- *A mathematical schema for global sensitivity analysis but it remains considerably harder to perform in the absence of the knowledge about uncertainties associated with all the inputs that go in an inversion.*
- *Develop methods to assess spatio-temporal correlation between forward operators of two or multiple observations. This is tied to overall diagnostics of the estimated fluxes as fluxes remain highly sensitive to the forward operator and improvement in understanding the representation of atmospheric transport through spatio-temporal association in the forward operator can lead to significant improvement in designing the components of a suitable inversion framework.*

Even though we have a comprehensive awareness of the literature associated with atmospheric inverse modeling and the methods used to assess them but here we would like to refer a paper by Michalak et al., 2017 (see below) that does not discuss sensitivity analysis in the context of linear atmospheric inverse problems

Michalak, A. M., Randazzo, N. A., and Chevallier, F.: Diagnostic methods for atmospheric inversions of long-lived greenhouse gases, Atmos. Chem. Phys., 17, 7405–7421, https://doi.org/10.5194/acp-17-7405-2017, 2017.

<u>Action:</u> *We have modified our manuscript in light of your and second reviewer's comments.*

4. **<u>Reviewer Comment</u>***:* My first problem with the paper is its title. The word "assessment" suggests some comment on the quality or robustness of an inversion. The authors don't do that and it's not clear from the paper that the generated metrics can do it.

   <u>Response</u>: *We agree with the reviewer's interpretation about the capability of the metrics that they do not provide any information regarding the robustness of inversion. The choice of the word assessment in the previous title of the manuscript was based on our understanding that providing information about the sensitivity of various inputs that goes into linear atmospheric inverse modeling in governing inverse solution can be considered as an assessment of inversions.*

   <u>Action</u>: *After considering reviewers' comments we have now changed the title of the paper as:* **Metrics for evaluating the "quality" in linear atmospheric inverse problems: a case study of a trace gas inversion**

5. **<u>Reviewer Comment</u>***:* I'm unclear, for example, what new information is provided by the overlap of footprints. It might well mean that parts of the control space are under-sampled by the observations but the posterior uncertainty already tells us this.

   <u>Response</u>: *We surmise that this comment is with respect to the IOAMI and JSD metrics. Yes, it is true that the both IOAMI and JSD provides information about what parts of the control space is under sampled and we agree that this information can also be obtained from posterior uncertainty, model resolution matrix and the plot of the forward operator or footprints themselves. However, IOAMI and JSD metrics proposed in our paper are more comprehensive as they allow researchers to:*

   - *Assess linear or non-linear correlation between footprints in space and time while also accounting for intensity of footprints. This correlation can also be expressed in the units of footprints, which to authors' knowledge is novel in the domain of atmospheric inverse problems.*
   - *Build a stable non-stationary covariance model for model-error covariance with diagonal and off-diagonal terms that can be incorporated in the inverse problems. Implementation details about these metrics have been submitted in the code and included in the supplementary material*
   - *Note that, although posterior uncertainty can indicate areas of low and high uncertainty, this uncertainty is also conflated with prior uncertainty. Deconvoluting or apportioning this correctly to find out uncertainty contribution*

from model-data is a challenging problem. In fact, one of the main objectives of the global sensitivity method that we present in our manuscript is to understand this apportionment of posterior uncertainty.

- *The technique proposed here does not suffer from this problem and provides an easy way to incorporate this into model-data error covariance. So, to reiterate, here our goal is to not exactly find low/high uncertainty areas but to provide meaningful correlation structure in model-data error covariance.*

Action: *No action required.*

6. **Reviewer Comment**: For the linear Gaussian case the posterior uncertainty can be calculated without measurements.

   Response*: Yes, in the linear Gaussian case, we can compute this even if the model-data error covariance is 0. We envision that the reduction term would be very close to prior error covariance with posterior uncertainty being close to 0.*

   Action: *No action required.*

7. **Reviewer Comment**: Likewise, the sensitivity of the posterior estimate to the value of a given measurement is potentially useful as a warning flag for measurements that might have undue control on the outcome but it's not really developed.

   Response: *The question of the undue importance of observation in governing the estimate of emissions or state vector depends on the goal of the study. In the case study presented in this work, we are interested in knowing the observations that influenced the estimate of emissions at the site of the Aliso Canyon gas leak. In other applications, the goal can be entirely different so providing a method for flagging observation is not desirable. However, the proposed method can rank the importance of observations in governing the estimate of emissions or state vector and section 4.2.1 shows how to evaluate these rankings.*

   Action: *Measurement influence via local sensitivity analysis and ranking on posterior estimate is discussed in detail in section 3.4 and 4.2.1*

8. **Reviewer Comment*:*** The global sensitivity analysis, which allows consideration of all inputs to the linear inverse problem, is potentially more interesting but again is not developed beyond generation of the first-order expansion. The example presents a good opportunity to demonstrate application of these methods but this is not taken far beyond calculation of the diagnostics.

   Response: *We thank the reviewer for the comment. The global sensitivity analysis (GSA) presented here uses local sensitivities but actually belongs to the class of variance-based methods. However, regardless of the methods we choose, full GSA is very complicated since it requires knowledge of the first-order joint dependence (aka their covariance) of the parameters. To exemplify, in an atmospheric inversion this would mean knowing the joint dependences of all the parameters including Q, R, and the transport model input parameters. Developing this method beyond first order requires second-order joint dependence (aka third-order joint moments). This is essentially knowing the entire joint distribution of the parameters. Such knowledge is very difficult to obtain in real applications and thus we do not develop it further.*

As mentioned in the answer of comment 5 of the reviewer, one of the other objectives of the GSA method adopted here is to be able to apportion the posterior uncertainty into Q and R component contributions. We can achieve that via first-order expansion.

Action: *This section (section 3.3) has been modified to reflect the caveats associated of the proposed method. It has been substantially rewritten in light of your and second reviewer's comments.*

Reviewer Comment: I see two possibilities for the paper (1) Repackage it as a technical note focusing on the calculation of the diagnostics or (2) Extend the work to generate diagnostics of overall inversion performance, probably focusing on robustness.

Response: *Our study is a technical note and focuses on the calculation of the proposed diagnostics referred by the reviewer as the first possibility.*

Action: *We have clarified the contribution of our paper in terms of a technical note in the last paragraph of the introductory section of the manuscript. Note as part of the technical note we have provided two detailed MATLAB Livescripts that implements all the diagnostics we provide in the manuscript and this has been mentioned in the last paragraph of the beginning of the section 2 (line 98 to 102).*

Specific Comments:

1. **Reviewer Comment:** What is a footprint-induced probability distribution?

   Response: *When we have many realizations of a random variable (scalar or vector-valued), we can compute a probability distribution based on the values of the random variable. This is often called probability distribution induced by the random variable. Likewise, probabilistically if we consider a set of all footprints from a transport and dispersion model as realizations of any underlying random vector, we can also come up with a probability distribution obtained by the footprints also known as a footprint-induced probability distribution.*

   Action: *We acknowledge that this is too pedantic and therefore we have modified this sentence in the revised manuscript (see lines 170-171).*

   Reviewer Comment: L190: The definition of the averaging kernel is true but this is an odd motivation for it, much better below when contemplating sensitivity of result to prior

   Response: *We agree that interpreting averaging kernel via local sensitivity of the estimated fluxes with respect to observations is not traditionally done. However, here our goal is simply to establish the link between local sensitivity with respect to observations and ubiquitous averaging kernel and show that averaging kernel and DOFS are just two subcases from the whole spectrum.*

   Action: *No action required.*

   Reviewer Comment: Eq. 19: it's worth noting that this sensitivity is very close to the proportional uncertainty reduction A P^-1 and hence the averaging kernel. By the way I thank the authors for making me think hard enough about the relationship between AK, DOFS and uncertainty reduction to finally get an intuitive sense of it

Response: *We convey sincere thanks to the reviewer for this comment. Yes, this sensitivity can also be thought as the proportion of posterior uncertainty to that of the prior uncertainty (i.e. VQ^-1) which intuitively makes sense. Whereas proportional uncertainty reduction is nothing but the averaging kernel. Thus sensitivity w.r.t. prior is negatively correlated with unknown fluxes or averaging kernel.*

Action: *In light of your comment we have added text the manuscript on this subject. Please see lines 254-256 in the revised manuscript.*

2. **Reviewer Comment:** L275: When commenting on covariances between H, Q etc we should also note that constraints like conservation of mass introduce strong covariances within the parameters of H. Covariances can only occur on physically plausible manifolds. This is a profoundly under-studied problem in transport modeling and there is probably great insight to be borrowed from Numerical Weather Prediction.

   Response: *We thank the reviewer for this insight. Yes, covariance within H parameters can be significantly high when conservation of mass is a constraint and in certain conditions (i.e. physically plausible manifolds) Q and H parameters can exhibit high correlation. Unless we have good knowledge of these cases, it is not possible to compute these dependencies.*

   Action: *No action required.*

   Reviewer Comment: L431: not sure what the authors mean by aggregation error here. If they're truly commenting on temporal aggregation error they should cite DOI:10.5194/acp-11-3443-2011.

   Response: *Yes, we are commenting on the aggregation error.*

   Action: *We have now included the reference mentioned by the reviewer.*