https://doi.org/10.5194/gmd-2022-89
Preprint. Discussion started: 23 May 2022

Geoscientific Model Development Discussions

Metrics for assessing Linear Inverse Problems: a case study of a Trace Gas Inversion

Vineet Yadav[1], Subhomoy Ghosh[2,3], and Charles E. Miller[1]

[1]Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA, USA
[2]University of Notre Dame, Notre Dame, IN, USA
[3]National Institute of Standards and Technology, Gaithersburg, MD, USA

Correspondence: Subhomoy Ghosh (sghosh4@nd.edu)

Abstract. Multiple metrics have been proposed and utilized to assess the performance of linear Bayesian and geostatistical inverse problems. These metrics are mostly related to assessing reduction in prior uncertainties, comparing modeled observations to true observations, and checking distributional assumptions. These metrics though important should be augmented with sensitivity analysis to obtain a comprehensive understanding of the performance of inversions and critically improve confidence in the estimated fluxes. With this motivation, we derive analytical forms of the local sensitivities with respect to the number of inputs such as measurements, covariance parameters, covariates, and forward operator or jacobian. In addition to local sensitivity, we develop a framework for global sensitivity analysis that shows the apportionment of the uncertainty of different inputs to an inverse problem. The proposed framework is applicable to any other domain that employs linear Bayesian and geostatistical inverse methods. We show the application of our methodology in the context of an atmospheric inverse problem for estimating urban GHG emissions in Los Angeles. Within its context, we also propose a mathematical framework to construct correlation functions and components of uncertainty matrices from a pre-computed jacobian that encompasses non-stationary structures.

1 Introduction

Inverse models within the context of atmospheric applications are often used for constraining global to regional scale fluxes of trace gases (for discussion see, Enting, 2002). At global scale, data assimilation (for further details on data assimilation, see Wikle and Berliner, 2007) that sequentially assimilates observations and updates the prior estimates of fluxes by utilizing an atmospheric model coupled with chemistry remains the primary inverse modeling framework. This framework at regional scale is complimented by inversions that assimilates all observations simultaneously by utilizing a pre-computed forward operator or jacobian (Lin et al., 2003) that describes relationship between observations and fluxes (for details, see Enting, 2002). This work focuses on these latter class of inverse methods. It specifically addresses sensitivity analysis and correlation in the jacobian in the context of Bayesian (for e.g., see Lauvaux et al., 2016) and geostatistical inverse methods (see Kitanidis, 1996).

The senstivity analysis in context of this study is covered under local and global themes. Primarily, we focus on local sensitivity analysis (LSA) that computes measure of the effect of a given input on a given output. This is obtained by computing partial derivatives of an output of interest with respect to an input factor (See Rabitz, 1989, and Turányi, 1990). Within global

https://doi.org/10.5194/gmd-2022-89
Preprint. Discussion started: 23 May 2022


Geoscientific Model Development Discussions

theme we focus on how uncertainty in the model output can be apportioned to different sources of uncertainty with respect to corresponding model input (Saltelli et al., 2008).

Overall, in atmospheric trace gas inversions mostly LSA is performed. Within this context, LSA asseses how sensitive the posterior estimates of fluxes are with reference to the underlying choices or assumptions, like (1) observations included, (2) model-data error covariance, (3) the input prior information and its error, and (4) the jacobian (for discussion see, Michalak et al., 2017). This task is sometimes performed to arrive at a robust estimate of fluxes and their uncertainties. It is achieved by running an inverse model multiple times by varying the inputs and assessing their impact on the estimated fluxes and uncertainties. Another complimentary way to do LSA is by computing local partial derivatives with respect to these quantities down to an individual entry that go in an inversion.

LSA can be grouped with standard information content approaches such as averaging kernel or model resolution matrix and degrees of freedom for signal (DOFS; for details see section 3.2.1 of this manuscript, Rodgers, 2000, and Brasseur and Jacob, 2017). Averaging kernel matrix shows how the estimated fluxes are related or sensitive to the true fluxes.Thus it belongs to the LSA category. However, LSA is more informative than DOFS and averaging kernel alone as it goes after individual components (see section 3.2) that determine DOFS. Furthermore, DOFS is a measure that provides an estimate of information resolved by an inversion. In comparison, LSA focuses on quantifying the impact and the relative importance of various components of an inversion in governing the estimates of fluxes.

In this work, we provide analytical expressions to conduct post hoc (that is after an inversion has been performed) LSA through local partial derivatives. In order to provide a complete framework, a mathematical schema for global sensitivity analysis (GSA) is also discussed 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.

We also develop methods to assess spatio-temporal correlation between jacobians of two or multiple observations. This is tied to overall diagnostics of the estimated fluxes as fluxes remain highly sensitive to the jacobian and improvement in understanding the representation of atmospheric transport through spatio-temporal association in the jacobian can lead to significant improvement in designing the components of a suitable inversion framework.

2 Organization of the study

We have divided this work in to two parts. In the first part (Section 3.1), we provide diagnostics associated with the jacobian. This is done as these derivations can be deployed for assessment of the jacobian prior to or separately from an inversion. In the second part (Section 3.2), we cover local and global sensitivity analysis that can only be conducted after inverse estimation of the fluxes. These two parts are followed by a methane (CH_4) case study that demonstrates the applicability of our methods

https://doi.org/10.5194/gmd-2022-89
Preprint. Discussion started: 23 May 2022


Geoscientific Model Development Discussions

(i.e., Section 4). To maintain maximum transparency, facilitate assessment, and show applicability of our methods in Section 3 we also provide two well documented interactive MATLAB Live scripts (for details on Live script see MatlabLivescript), one for each methodological part that contains equations, code, and visualizations as it relates to the real-data case study described in Section 4. All of these are included as supplementary material. Separate pdfs of these Live scripts are also included for the readers who do not have access to MATLAB.

3 Methods and derivation

3.1 Analysis of the jacobian

In inversions that assimilates all observations simultaneously, first a jacobian for each observation that would be included in an inversion is obtained from a transport model. These observations of trace gases can be obtained from multiple platforms that include in-situ network of fixed locations on the surface, intermittent aircraft flights and satellites. In most situations, the spatio-temporal coverages of these jacobians are visually assessed by plotting an aggregated sum or mean of their values over a map of the spatial domain of the study. However, standard quantitative metrics to assess their coverage and intensity in space and time remains completely absent. In this study, we present two metrics for this assessment and these are defined below. These metrics conform to triangular inequality and therefore can be defined as distance function in their respective metric spaces.

Note sometimes in the published literature on trace gas inversions the jacobian obtained from a transport model is referred to as a sensitivity matrix or footprint. Henceforth, to avoid misinterpretation, we always refer to jacobian as footprint. We show our application through footprints constructed by running a Lagrangian transport model. However, our methods can also be applied in analytical Eulerian framework (see Brasseur and Jacob, 2017 for details).

3.1.1 Integrated area overlap measurement index (IAOMI)

The Integrated Area Overlap Measurement Index (IAOMI) summarizes the shared information content between two footprints and hence indirectly between two observations under the assumption that error in dispersion and transport is absent and the sources of emissions are uniformly distributed. It is therefore a measure of the uniqueness of the flux signal associated with an observation in comparison to other observations.

Intuitively, IAOMI can be understood in terms of functions on sets. For e.g., given two footprints we can consider their entries or intensities as two sets indexed by spatio-temporal coordinates. Within such a context of sets, IOAMI is just the proportion of the common contribution of the two footprints from the intersected area with respect to the overall contribution of the two footprints. This is demonstrated through a Venn diagram in Figure 1.

$$\nu_{F,G} = \frac{\Sigma_{A_F \cap A_G} H_1(F,G)}{\Sigma_{A_F \cup A_G} H_2(F,G)} \tag{1}$$

https://doi.org/10.5194/gmd-2022-89
Preprint. Discussion started: 23 May 2022

Geoscientific Model Development Discussions
Open Access
EGU

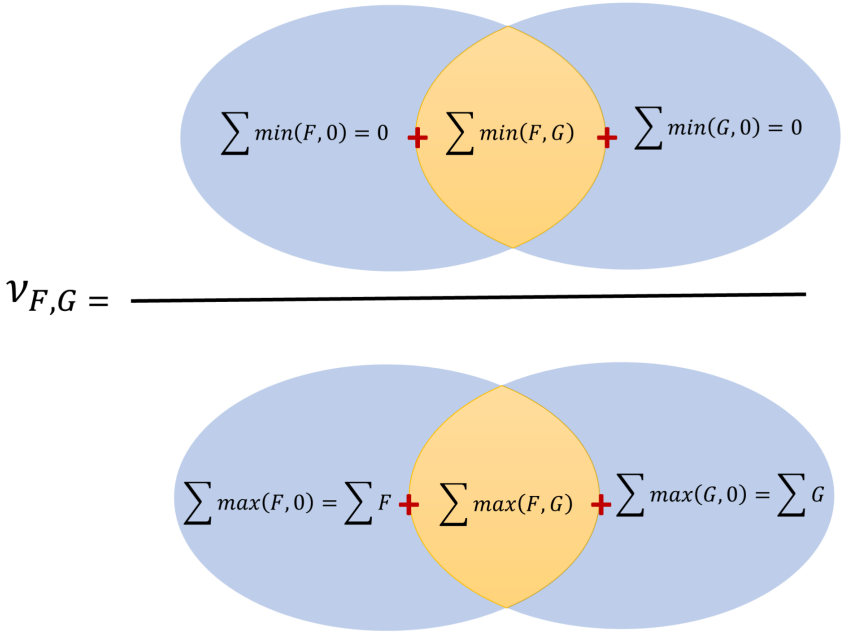


Figure 1. Venn diagram that defines IOAMI in terms of two hypothetical footprints F and G

Where for any footprint S, the corresponding set A_S on which footprints are always positive, is defined as $A_S = \{(i,j,t) : (i,j) \in \mathcal{D}, t \in T, S(i,j,t) > 0\}$. The functionals H_1 and H_2 are functions of the two footprints under comparison and are defined as:

$$H_1(F,G) = \begin{cases} min(F,G) & \text{on } A_F \cap A_G \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad H_2(F,G) = \begin{cases} max(F,G) & \text{on } A_F \cap A_G \\ F & \text{on } A_F \cap A_G^c \\ G & \text{on } A_F^c \cap A_G \end{cases} \tag{2}$$

Note that, the $\nu_{F,G}$ defined above can also be written as a simple ratio of the sum over minimums over sum of the maximums.

$$\nu_{F,G} = \frac{\Sigma_{A_F \cup A_G} min(F,G)}{\Sigma_{A_F \cup A_G} max(F,G)} \tag{3}$$

Thus, IAOMI ν can also be thought as a measure of similarity between two footprints. It is evident from equation 3 that this is a weighted Jaccard similarity index or Ruzicka index (Cha, 2007). It follows that ν is closed and bounded in $[0,1]$ and accounts for both spatio-temporal spread and intensity of the footprint. A stronger ν implies larger overlap of intensity in space and time and is analogous to finding common area within two curves. The corresponding measure of dissimilarity can be defined by $1-\nu$. The smaller the overlap or the larger the value of $1-\nu$, the larger is the dissimilarity. Note the ν metric

https://doi.org/10.5194/gmd-2022-89
Preprint. Discussion started: 23 May 2022


Geoscientific Model Development Discussions

is only indicative of the overlap in the spatio-temporal intensity between two footprints. To measure how much of the shared intensity has come from either footprint, we use a metric I_{ν_F} defined as:

$$I_{\nu_F} = \frac{\Sigma_{A_F \cap A_G} H_1(F,G)}{\Sigma_{A_F} H_3(F)} \tag{4}$$

Where $H_3(F) = F$ on A_F and 0 everywhere else. Likewise, we can define I_{ν_G} which shows proportional contribution of the footprint G on the shared intensity. Both ν and $I_{\nu_.}$ can be computed from observations taken from same or different platforms, at same or different time or for two different in-situ measurement sites over a specified time-interval.

3.1.2 Spatio-temporal Area of Dominance (STAD)

The notion of the spatio-temporal area of dominance (STAD) stems naturally from IAOMI. For any two footprints F, and G, we can find out the left-over dominant contribution of F and G by computing quantities $F - G$ and $G - F$ that leads to determination of the area where F or G is dominant.

Mathematically, for two footprints F and G, STAD of F with respect to G is defined as:

$$\text{STAD}_F(F,G) = \begin{cases} F - min(F,G) & \text{on } A_F \cap A_G \\ F & \text{otherwise} \end{cases}$$

IAOMI and STAD of any footprint F with respect to the footprints F and G are linked by the following equation:

$$\nu_{F,G} \Sigma H_2(F,G) + \Sigma \text{STAD}_F(F,G) = \Sigma F \quad \text{on } A_F \cup A_G \tag{5}$$

Given a set of footprints $\{F, G_1, \cdots, G_K\}$, STAD for any particular footprint F with respect to all other footprints can be generalized from equation 5 as $F_{\text{STAD}}(F, H_{-F})$ where $H_{-F} = \max_k G_k$ on A_G; $A_G = \cup_k A_{G_k}$. A_{G_k} is the set on which footprint G_k is always positive (see section 3.1.1 for it's definition). STAD can be aggregated over any time-periods. Intuitively, STAD determines areas in space-time where one footprint dominates over other footprints. This is especially useful in locating the primary sources of emissions that influences an observation.

3.1.3 Jensen-Shannon distance (JSD) for footprints

Dissimilarity between footprints can also be measured via entropy (for definition, see MacKay et al., 2003) based distances. Entropy distances are sensitive in capturing differences between two distributions that are similar in 1st order (e.g. mean, or median) and second order moments (e.g. variance, or quartile deviation) but differ in higher order moments (e.g. Kurtosis) or modes (e.g. unimodal vs. multimodal). Entropy based distance metrics that adhere to triangular inequality can also be combined with spatio-temporal coverage to measure the probabilistic divergence between two footprints. One such metric is

https://doi.org/10.5194/gmd-2022-89
Preprint. Discussion started: 23 May 2022


Geoscientific Model Development Discussions

Jensen-Shanon distance (JSD) (Nielsen, 2019) which can be used to compute distance between two distributions generated by the footprints. Normalized footprints can be seen as samples from an underlying high-dimensional probability distribution such that total sum is one. For any particular time point t, this normalization by the total sum, for a vectorized footprint ($F(i,j,t)$ can be written as $F(k,t)$ where k spans over all combinations of i and j) of an observation can be given as:

$$P_{F(k,t)} = \frac{F(k,t)}{\sum_{k=1}^{g} F(k,t)} \tag{6}$$

where P denotes a probability measure. We can then use JSD to compute distance between two footprint induced probability distributions. After normalization, JSD can be computed as:

$$JSD\left(P_F||P_G\right) = \sqrt{\frac{1}{2}D\left(P_F||M\right) + \frac{1}{2}D\left(P_G||M\right)} \tag{7}$$

where D stands for Kulback-Leibler (KL) divergence (see MacKay et al., 2003 for details). D of any probability measure p with respect to another probability measure q is defined as: $D(p||q) = \sum p\log(p/q)$. M is defined as: $M = \frac{1}{2}\left(P_F + P_G\right)$. The symbol $||$ is used to indicate that $D\left(P_F||M\right)$ and $D\left(P_G||M\right)$ are not conditional entropies (see MacKay et al., 2003). JSD is closed and bounded in $[0,1]$ when KL divergence is computed with base 2 logarithm. Intuitively, JSD and $1-\nu$ (i.e. 1-IAOMI) are comparable since both of them are measures of dissimilarity.

Note that one can use JSD or 1-IAOMI matrix of all pairwise footprints as a representative distance matrix for describing correlations in model-data errors (i.e., $\mathbf{R}$ in equation 8). This correlation matrices need to be at least positive semi-definite. Since JSD or 1-IAOMI matrices are real, symmetric, and admit orthogonal decomposition, element-wise exponential of such symmetric diagonalizable matrices would be positive-semidefinite. Thus, they can be incorporated in $\mathbf{R}$ via the commonly adopted exponential kernel of the distance matrix (see Ghosh et al., 2021). Furthermore, the IAOMI matrix itself is a positive semidefinite (Bouchard et al., 2013) matrix and can also be directly incorporated in $\mathbf{R}$ as a measure of correlation. However, we do not explore this area of research in this manuscript.

3.2 Local sensitivity analysis in inversions

For linear Bayesian and geostatistical inverse problem, the solutions (see, Tarantola, 2005 for the batch Bayesian and Kitanidis, 1996 for the geostatistical case) can be obtained by minimizing their respective objective functions. These objective functions can be given by equations 8 and 9 as:

$$L(\mathbf{s}|\mathbf{y},\mathbf{s}_{\text{prior}},\mathbf{H},\mathbf{Q},\mathbf{R}) = \frac{1}{2}(\mathbf{z}-\mathbf{H}\mathbf{s})^T\mathbf{R}^{-1}(\mathbf{z}-\mathbf{H}\mathbf{s}) + \frac{1}{2}(\mathbf{s}-\mathbf{s}_{\text{prior}})^T\mathbf{Q}^{-1}(\mathbf{s}-\mathbf{s}_{\text{prior}}) \tag{8}$$

$$L(\mathbf{s}|\mathbf{y},\mathbf{H},\mathbf{Q},\mathbf{R},\boldsymbol{\beta}) = \frac{1}{2}(\mathbf{z}-\mathbf{H}\mathbf{s})^T\mathbf{R}^{-1}(\mathbf{z}-\mathbf{H}\mathbf{s}) + \frac{1}{2}(\mathbf{s}-\mathbf{X}\boldsymbol{\beta})^T\mathbf{Q}^{-1}(\mathbf{s}-\mathbf{X}\boldsymbol{\beta}) \tag{9}$$

https://doi.org/10.5194/gmd-2022-89
Preprint. Discussion started: 23 May 2022

Geoscientific Model Development Discussions
Open Access
EGU

where lower case symbols represent vectors, the uppercase symbols represent matrices, and this same approach of representation is adopted throughout the manuscript. In equation 8 and 9, $\mathbf{z}_{(n,1)}$ with units ppm are available measurements. $\mathbf{H}_{(n,m)}$ with units ppm μmoles^{-1}m^{2}sec is a footprint, obtained from a transport model that describes the relationship between measurements and unknown fluxes. $\mathbf{s}_{(m,1)}$ are unknown fluxes that have units μmoles m^{-2}sec^{-1}. $\mathbf{R}_{(n,n)}$ with units ppm^2 is the covariance of the model-data errors. $\mathbf{X}_{(m,p)}$ are known covariates related to $\mathbf{s}$. The unit of each of the covariate in $\mathbf{X}$ is the unit of it's measurement or if it is standardized (e.g. subtract a covariate by its mean and divide by its standard deviation) then it is unitless (for further discussion on standardization and normalization see Gelman and Hill, 2006). The units of $\beta_{(p,1)}$ are such that $\mathbf{X}\beta$ and $\mathbf{s}$ have the same units. $\mathbf{Q}_{(m,m)}$ with units $\left(\mu\text{moles m}^{-2}\text{sec}^{-1}\right)^2$ describes errors between unknown $\mathbf{s}$ and $\mathbf{X}\beta$.

The analytical solutions for the unknown fluxes $\mathbf{s}$ in the Bayesian case (denoted by subscript B) and the geostatistical case (denoted by subscript G) can be obtained from equations 10 and 11 as given below.

$$\hat{\mathbf{s}}_B = \mathbf{s}_{\text{prior}} + \mathbf{Q}\mathbf{H}^T\left(\mathbf{H}\mathbf{Q}\mathbf{H}^T + \mathbf{R}\right)^{-1}\left(\mathbf{z} - \mathbf{H}\mathbf{s}_{\text{prior}}\right) \tag{10}$$

$$\hat{\mathbf{s}}_G = \mathbf{X}\boldsymbol{\beta} + \mathbf{Q}\mathbf{H}^T\left(\mathbf{H}\mathbf{Q}\mathbf{H}^T + \mathbf{R}\right)^{-1}\left(\mathbf{z} - \mathbf{H}\mathbf{X}\beta\right) \tag{11}$$

Equation 11 is often expressed as $\mathbf{s}_G = \mathbf{X}\beta + \epsilon$ where $\mathbf{X}\beta$ is the mean and $\epsilon = \mathbf{Q}\mathbf{H}^T\left(\mathbf{H}\mathbf{Q}\mathbf{H}^T + \mathbf{R}\right)^{-1}\left(\mathbf{z} - \mathbf{H}\mathbf{X}\beta\right)$ is the stochastic part of the estimated fluxes. As the estimate of $\mathbf{s}_G$ in equation 11 depends on the unknown β, it needs to be estimated prior to obtaining $\hat{\mathbf{s}}_G$. The solution for the $\hat{\beta}$ can be obtained from pre-determined quantities as described earlier in the context of equation 9 and can be given as:

$$\hat{\boldsymbol{\beta}} = \left((\mathbf{H}\mathbf{X})^T\left(\mathbf{H}\mathbf{Q}\mathbf{H}^T + \mathbf{R}\right)^{-1}\mathbf{H}\mathbf{X}\right)^{-1}(\mathbf{H}\mathbf{X})^T\left(\mathbf{H}\mathbf{Q}\mathbf{H}^T + \mathbf{R}\right)^{-1}\mathbf{z} \tag{12}$$

Plugging in $\hat{\boldsymbol{\beta}}$ in equation 11 leads to equation 13 where all symbols are defined previously or in equation 14.

$$\hat{\mathbf{s}}_G = \mathbf{X}\boldsymbol{\Omega}^{-1}\mathbf{A}^T\boldsymbol{\Psi}^{-1}\mathbf{z} + \mathbf{Q}\mathbf{H}^T\boldsymbol{\Psi}^{-1}\left(\mathbf{z} - \mathbf{A}\boldsymbol{\Omega}^{-1}\mathbf{A}^T\boldsymbol{\Psi}^{-1}\mathbf{z}\right) \quad \text{where} \tag{13}$$

$$\mathbf{A} = \mathbf{H}\mathbf{X}, \boldsymbol{\Psi} = \left(\mathbf{H}\mathbf{Q}\mathbf{H}^T + \mathbf{R}\right), \boldsymbol{\Omega} = (\mathbf{H}\mathbf{X})^T\left(\mathbf{H}\mathbf{Q}\mathbf{H}^T + \mathbf{R}\right)^{-1}\mathbf{H}\mathbf{X} \tag{14}$$

We differentiate equation 10 with respect to $\mathbf{s}_{\text{prior}}$, $\mathbf{R}$, $\mathbf{Q}$, $\mathbf{z}$ and equation 13 with respect to $\mathbf{X}$, $\mathbf{R}$, $\mathbf{Q}$, $\mathbf{z}$ to obtain the local sensitivities. There are two ways to differentiate $\hat{\mathbf{s}}$ with respect to $\mathbf{z}$, $\mathbf{X}$, $\mathbf{H}$, $\mathbf{Q}$, and $\mathbf{R}$. In the first case, every entry in $\mathbf{z}$, $\mathbf{X}$, $\mathbf{H}$, $\mathbf{Q}$, and $\mathbf{R}$ can be considered as a parameter that results in differentiation of $\hat{\mathbf{s}}$ with respect to these quantities. On the other hand, if the structures of the covariance matrices $\mathbf{Q}$ and $\mathbf{R}$ are determined by parameters then $\hat{\mathbf{s}}$ can be differentiated just with respect to these parameters. In the former case, equations 10 and 13 are used to differentiate $\hat{\mathbf{s}}$ with respect to an entry at a time in $\mathbf{z}$, $\mathbf{X}$, $\mathbf{H}$, $\mathbf{Q}$, and $\mathbf{R}$. Such an approach of entry-by-entry differentiation that is one at a time differentiation is useful if the computational cost in terms of memory constraint is important or if we would like to know the influence of a single entry on $\hat{\mathbf{s}}$. We provide both sets of equations in this work.

https://doi.org/10.5194/gmd-2022-89
Preprint. Discussion started: 23 May 2022

Geoscientific Model Development Discussions

3.2.1 Local sensitivity analysis with respect to observations, priors, scaling factors, and footprints

Local sensitivity of $\hat{\mathbf{s}}$ with respect to observations ($\mathbf{z}$) can be given as

$$\frac{\partial \hat{\mathbf{s}}_B}{\partial \mathbf{z}} = \mathbf{Q}\mathbf{H}^T\mathbf{\Psi}^{-1} \tag{15}$$

$$\frac{\partial \hat{\mathbf{s}}_G}{\partial \mathbf{z}} = \mathbf{X}\mathbf{\Omega}^{-1}\mathbf{A}^{\mathbf{T}}\mathbf{\Psi}^{-1} + \mathbf{Q}\mathbf{H}^T\mathbf{\Psi}^{-1} - \mathbf{Q}\mathbf{H}^T\mathbf{\Psi}^{-1}\mathbf{A}\mathbf{\Omega}^{-1}\mathbf{A}^{\mathbf{T}}\mathbf{\Psi}^{-1} = \mathbf{\Lambda} \tag{16}$$

where all quantities are as defined earlier. In the geostatistical case, equation 16 is also referred as $\mathbf{\Lambda}$ (for e.g. see Michalak et al., 2004 and Gourdji et al., 2010). The units of $\frac{\partial \hat{\mathbf{s}}}{\partial \mathbf{z}}$ are μmoles^{-1}m^{2}sec^{-1}ppm^{-1}, which is inverse of the units of $\mathbf{H}$. Local sensitivities with respect to an observation z_i for both the Bayesian and geostatistical case can written as vector of sensitivities times an indicator for the i$^{\text{th}}$ entry i.e. $\frac{\partial \hat{\mathbf{s}}}{\partial \mathbf{z}} e_i$ where $e_i = \frac{\partial \mathbf{z}}{\partial z_i}$ is a vector of zeros with the i$^{\text{th}}$ entry equals to 1.

Note by utilizing $\frac{\partial \hat{\mathbf{s}}}{\partial \mathbf{z}}$, we can also obtain an averaging kernel (or model resolution matrix) and DOFS (see Rodgers, 2000). The averaging kernel matrix for any linear inverse model can be written as:

$$\mathbf{Avk} = \frac{\partial \hat{\mathbf{s}}}{\partial \mathbf{z}} \times \mathbf{H} \tag{17}$$

$$\text{DOFS} = \text{Tr}\left(\mathbf{Avk}\right) \tag{18}$$

where **Avk** is the local sensitivity of $\hat{\mathbf{s}}$ with respect to the true unknown fluxes and DOFS represents the amount of information resolved by an inverse model when a set of observations have been assimilated (for a detailed discussion, see Rodgers, 2000 and Brasseur and Jacob, 2017). Theoretically, the value of DOFS cannot exceed number of observations (n) in case of an underdetermined system and number of fluxes (m) in case of an overdetermined system.

We can directly compute local sensitivity of $\hat{\mathbf{s}}$ with respect to the prior mean flux $\mathbf{s}_{\text{prior}}$ in the Bayesian case. In the geostatistical case, the prior mean is modeled by two quantities $\mathbf{X}$ and β. In this scenario, we need to find sensitivities with respect to $\mathbf{X}$ as well as β. These local sensitivities are given by:

$$\frac{\partial \hat{\mathbf{s}}_B}{\partial \mathbf{s}_{\text{prior}}} = \mathbf{I} - \mathbf{C}\mathbf{H} \tag{19}$$

$$\frac{\partial \hat{\mathbf{s}}_G}{\partial \mathbf{X}} = \mathbf{K}_z \otimes \left(\mathbf{I} + \left(\mathbf{M}\mathbf{A}^T - \mathbf{X}\mathbf{\Omega}^{-1}\mathbf{A}^T - \mathbf{Q}\mathbf{H}^T\right)\mathbf{\Psi}^{-1}\mathbf{H}\right) + \left(\mathbf{X}\mathbf{\Omega}^{-1} - \mathbf{M}\right) \otimes \left(\mathbf{F}_z - \mathbf{K}_z\mathbf{A}^T\Psi^{-1}\mathbf{H}\right) \tag{20}$$

$$\frac{\partial \hat{\mathbf{s}}_G}{\partial \hat{\beta}} = \mathbf{X} - \mathbf{C}\mathbf{A} \tag{21}$$

where $\mathbf{A} = \mathbf{H}\mathbf{X}$, $\mathbf{B} = \mathbf{Q}\mathbf{H}^T$, $\mathbf{C} = \mathbf{B}\mathbf{\Psi}^{-1}$, $\mathbf{\Omega} = \mathbf{A}^T\mathbf{\Psi}^{-1}\mathbf{A}$, $\mathbf{K}_z = \mathbf{z}^T\Psi^{-1}\mathbf{A}\Omega^{-1}$, $\mathbf{M} = \mathbf{C}\mathbf{A}\mathbf{\Omega}^{-1}$, and $\mathbf{F}_z = \mathbf{z}^T\mathbf{\Psi}^{-1}\mathbf{H}$. The symbol $\otimes$ represents the Kronecker product. The quantity $\frac{\partial \hat{\mathbf{s}}_B}{\partial \mathbf{s}_{\text{prior}}}$ is unitless whereas the units of each column of $\frac{\partial \hat{\mathbf{s}}_G}{\partial \hat{\beta}}$ is of

https://doi.org/10.5194/gmd-2022-89
Preprint. Discussion started: 23 May 2022

Geoscientific Model Development Discussions

the form $(\mu\text{moles}^{-1}\text{m}^2\text{sec}^{-1})(\text{unit of } \beta_i)^{-1}$. The sensitivity matrix $\frac{\partial \hat{\mathbf{s}}_G}{\partial \mathbf{X}}$ is of dimension $m \times mp$ where every i$^{\text{th}}$ block of m columns $((i-1)m + A : im)$ of $\frac{\partial \hat{\mathbf{s}}_G}{\partial \mathbf{X}}$ has units of the form $(\mu\text{moles}^{-1}\text{m}^2\text{sec}^{-1})(\text{unit of } \mathbf{X}_i)^{-1}$ where $\mathbf{X}_i$ is the i$^{\text{th}}$ column of $\mathbf{X}$.

Sometimes, it is important to know the influence of the prior of any particular grid point or an area consisting of few points on $\hat{\mathbf{s}}$. Local sensitivities of $\hat{\mathbf{s}}$ with respect to the i$^{\text{th}}$ entry in $\mathbf{s}_{\text{prior}}$ and $\hat{\beta}_i$ are straightforward and can be written as $\frac{\partial \hat{\mathbf{s}}_B}{\partial \mathbf{s}_{\text{prior}}} e_i$ and $\frac{\partial \hat{\mathbf{s}}_G}{\partial \hat{\beta}} e_i$ respectively. However, the entry-wise $\frac{\partial \hat{\mathbf{s}}_G}{\partial \mathbf{X}_{ij}}$ is more complex and can be given by:

$$\frac{\partial \hat{\mathbf{s}}_G}{\partial \mathbf{X}_{ij}} = (\mathbf{I} - \mathbf{CH}) \left(\left(\mathbf{I} - \mathbf{X}\boldsymbol{\Omega}^{-1}\mathbf{X}^T\mathbf{H}^T\Psi^{-1}\mathbf{H} \right) \frac{\partial \mathbf{X}}{\partial \mathbf{X}_{ij}} \boldsymbol{\Omega}^{-1}\mathbf{X}^T + \mathbf{X}\boldsymbol{\Omega}^{-1} \frac{\partial \mathbf{X}^T}{\partial \mathbf{X}_{ij}} \left(\mathbf{I} - \mathbf{H}^T\boldsymbol{\Psi}^{-1}\mathbf{H}\mathbf{X}\boldsymbol{\Omega}^{-1}\mathbf{X}^T \right) \right) \mathbf{F}^T \tag{22}$$

where $\frac{\partial \mathbf{X}^T}{\partial \mathbf{X}_{ij}} = E_{ij}$ is a single-entry matrix with a one for a X_{ij} for which differentiation is being performed and zero everywhere else. For $\mathbf{z}$, entry-by-entry differentiation can be easily performed, since both equations 10 and 13 result from linear models and are functions of the form $\boldsymbol{\Phi}\mathbf{z} + \mathbf{n}$ where $\boldsymbol{\Phi}$ and $\mathbf{n}$ are independent of $\mathbf{z}$. For example, $\boldsymbol{\Phi}$ and $\mathbf{n}$ for equation 10 are $\mathbf{QH}^T \left(\mathbf{HQH}^T + \mathbf{R} \right)^{-1}$ and $\mathbf{s}_{\text{prior}} - \mathbf{QH}^T \left(\mathbf{HQH}^T + \mathbf{R} \right)^{-1} \mathbf{Hs}_{\text{prior}}$ respectively and are independent of $\mathbf{z}$. In this case, $\frac{\partial \hat{\mathbf{s}}_B}{\partial \mathbf{z}_i}$ can be written as $\boldsymbol{\Phi}\mathbf{e_i}$ where $\boldsymbol{e_i}$ is a single-entry vector with a one for a $\boldsymbol{z}_i$ for which differentiation is being performed and zero everywhere else. $\frac{\partial \hat{\mathbf{s}}_G}{\partial \mathbf{z}_i}$ can similarly be defined for the respective $\boldsymbol{\Phi}$.

Local sensitivity of $\hat{\mathbf{s}}$ with respect to a forward operator has units of the form $\left(\mu\text{moles}^{-1}\text{m}^2\text{sec}^{-1}\right)^2 \text{ppm}^{-1}$. In the Bayesian case this sensitivity can written as:

$$\frac{\partial \hat{\mathbf{s}}_B}{\partial \mathbf{H}} = \mathbf{Q} \otimes \mathbf{P}_z - \mathbf{BP}_z \otimes \mathbf{C}^T - \mathbf{BC}^T \otimes \mathbf{P}_z - \mathbf{Q} \otimes \mathbf{D} + \mathbf{BD} \otimes \mathbf{C}^T + \mathbf{BC}^T \otimes \mathbf{D} - \mathbf{s}_{\text{prior}} \otimes \mathbf{C}^T \tag{23}$$

In the geostatistical case, this sensitivity can be partitioned into two components i.e., $\frac{\partial \hat{\beta}}{\partial \mathbf{H}}$ and $\frac{\partial \hat{\epsilon}}{\partial \mathbf{H}}$ as shown in equation 24 where $\frac{\partial \hat{\beta}}{\partial \mathbf{H}}$ and $\frac{\partial \hat{\epsilon}}{\partial \mathbf{H}}$ are obtained in an orderly sequence from equations 25 and 26.

$$\frac{\partial \hat{\mathbf{s}}_G}{\partial \mathbf{H}} = \mathbf{X}\frac{\partial \hat{\beta}}{\partial \mathbf{H}} + \frac{\partial \hat{\epsilon}}{\partial \mathbf{H}} \quad \text{where} \tag{24}$$

$$\frac{\partial \hat{\beta}}{\partial \mathbf{H}} = -\mathbf{L} \otimes \mathbf{G}_z - \mathbf{P}_z^T\mathbf{A}\boldsymbol{\Omega}^{-1}\mathbf{X}^T \otimes \mathbf{K}^T + \mathbf{G}_z\mathbf{HQ} \otimes \mathbf{K}^T + \mathbf{N} \otimes \mathbf{G}_z + \mathbf{L} \otimes \mathbf{P}_z^T - \mathbf{P}_z^T\mathbf{HQ} \otimes \mathbf{K}^T - \mathbf{N} \otimes \mathbf{P}_z^T \tag{25}$$

$$\frac{\partial \hat{\epsilon}}{\partial \mathbf{H}} = \mathbf{Q} \otimes \mathbf{P}_z - \mathbf{Cz} \otimes \mathbf{C}^T - \mathbf{CHQ} \otimes \mathbf{P}_z - \mathbf{XK}^T\mathbf{z} \otimes \mathbf{C}^T - \mathbf{CA}\frac{\partial \hat{\beta}}{\partial \mathbf{H}} \tag{26}$$

The expanded form of some of the symbols in equations 23 through 26, those have not been expanded yet can be written as $\mathbf{D} = \boldsymbol{\Psi}\mathbf{Hs}_{\text{prior}}$, $\mathbf{G}_z = \mathbf{z}^T\boldsymbol{\Psi}^{-1}\mathbf{A}\Omega^{-1}\mathbf{A}^T\boldsymbol{\Psi}^{-1}$, $\mathbf{L} = \boldsymbol{\Omega}^{-1}\mathbf{X}^T$, $\mathbf{N} = \boldsymbol{\Omega}^{-1}\mathbf{A}^T\boldsymbol{\Psi}^{-1}\mathbf{HQ}$, $\mathbf{P}_z = \boldsymbol{\Psi}^{-1}\mathbf{z}$, and $\mathbf{K} = \boldsymbol{\Psi}^{-1}\mathbf{A}\Omega^{-1}$. The unit of $\frac{\partial \hat{\mathbf{s}}}{\partial \mathbf{H}}$ is $(\mu\text{moles}^{-1}\text{m}^2\text{sec}^{-1})^2\text{ppm}^{-1}$.

https://doi.org/10.5194/gmd-2022-89
Preprint. Discussion started: 23 May 2022

Geoscientific Model Development Discussions

There might be times when we would like to know sensitivity of the transport ($\mathbf{H}$) with respect to certain source locations only. In this case, we can use ij form of equations 23 through 26 to obtain $\frac{\partial \hat{\mathbf{s}}_B}{\partial H_{ij}}$ in parts. In this formulation $\frac{\partial \hat{\mathbf{s}}_B}{\partial H_{ij}}$ can be given as:

$$\frac{\partial \hat{\mathbf{s}}_B}{\partial H_{ij}} = \mathbf{C}\frac{\partial \mathbf{H}}{\partial \mathbf{H}_{ij}}\left(\mathbf{C}(\mathbf{H}\mathbf{s}_{\text{prior}} - \mathbf{z}) - \mathbf{s}_{\text{prior}}\right) + (\mathbf{Q} - \mathbf{CHQ})\left(\frac{\partial \mathbf{H}}{\partial \mathbf{H}_{ij}}\right)^T \mathbf{\Psi}^{-1}(\mathbf{z} - \mathbf{H}\mathbf{s}_{\text{prior}}) \tag{27}$$

$$\frac{\partial \hat{\mathbf{s}}_G}{\partial H_{ij}} = \mathbf{X}\frac{\partial \hat{\beta}}{\partial H_{ij}} + \frac{\partial \hat{\epsilon}}{\partial H_{ij}}, \quad \text{where} \tag{28}$$

$$\frac{\partial \hat{\beta}}{\partial H_{ij}} = \left(-\mathbf{K}^T\frac{\partial \mathbf{H}}{\partial H_{ij}}\left(\mathbf{XN} - \mathbf{CAS} + \mathbf{QH}^T\right) + \mathbf{K}^T\mathbf{HQ}\frac{\partial \mathbf{H}^T}{\partial H_{ij}}\left(\mathbf{\Psi}^{-1}\mathbf{AS}^T - \mathbf{I}\right) + \mathbf{\Omega}^{-1}\mathbf{X}^T\frac{\partial \mathbf{H}^T}{\partial H_{ij}}\left(\mathbf{I} - \mathbf{\Psi}^{-1}\mathbf{AS}\right)\right)\mathbf{\Psi}^{-1}\mathbf{z} \tag{29}$$

$$\frac{\partial \hat{\epsilon}}{\partial H_{ij}} = \left(\mathbf{Q}\frac{\partial \mathbf{H}^T}{\partial H_{ij}} - \mathbf{C}\frac{\partial \mathbf{H}}{\partial H_{ij}}\mathbf{QH}^T - \mathbf{CHQ}\frac{\partial \mathbf{H}^T}{\partial H_{ij}}\right)\mathbf{\Psi}^{-1}\left(\mathbf{z} - \mathbf{A}\hat{\beta}\right) - \mathbf{C}\left(\frac{\partial \mathbf{H}}{\partial H_{ij}}\mathbf{X}\hat{\beta} + \mathbf{A}\frac{\partial \hat{\beta}}{\partial H_{ij}}\right) \tag{30}$$

where $\mathbf{S} = \mathbf{A\Omega}^{-1}$ and the matrix $\frac{\partial \mathbf{H}}{\partial H_{ij}}$ is a single-entry matrix with a one for a H_{ij} entry for which the differentiation is being performed and zero everywhere else. Units of $\frac{\partial \hat{\mathbf{s}}_B}{\partial H_{ij}}$ and $\frac{\partial \hat{\mathbf{s}}_G}{\partial H_{ij}}$ are the same as their kronecker product counterparts.

3.2.2 Local sensitivity analysis with respect to error covariance matrices and prior information

In order to compute the local sensitivities of $\hat{\mathbf{s}}$ with respect to $\mathbf{Q}$ and $\mathbf{R}$, consider that they are parametrized as $\mathbf{Q}(\theta_{\mathbf{Q}})$ and $\mathbf{R}(\theta_{\mathbf{R}})$ where $\theta_{\mathbf{Q}}$ and $\theta_{\mathbf{R}}$ are the parameter vectors. The differentiation with respect to error covariance parameters in $\mathbf{Q}$ and $\mathbf{R}$ can be accomplished through equations 31 through 34 where the subscript i indicates the i^{th} covariance parameter for which differentiation is being performed.

$$\frac{\partial \hat{\mathbf{s}}_B}{\partial \theta_{Q_i}} = (\mathbf{I} - \mathbf{CH})\frac{\partial \mathbf{Q}}{\partial \theta_{Q_i}}\mathbf{H}^T\mathbf{\Psi}^{-1}(\mathbf{z} - \mathbf{H}\mathbf{s}_{\text{prior}}) \tag{31}$$

$$\frac{\partial \hat{\mathbf{s}}_G}{\partial \theta_{Q_i}} = \left(-\mathbf{X\Omega}^{-1}\mathbf{A}^T\mathbf{\Psi}^{-1}\mathbf{H} + \mathbf{I} - \mathbf{QH}^T\mathbf{\Psi}^{-1}\mathbf{H} + \mathbf{QH}^T\mathbf{\Psi}^{-1}\mathbf{A\Omega}^{-1}\mathbf{A}^T\mathbf{\Psi}^{-1}\mathbf{H}\right)\frac{\partial \mathbf{Q}}{\partial \theta_{Q_i}}\mathbf{H}^T\mathbf{\Psi}^{-1}(\mathbf{z} - \mathbf{A\Omega}^{-1}\mathbf{A}^T\mathbf{\Psi}^{-1}\mathbf{z}) \tag{32}$$

$$\frac{\partial \hat{\mathbf{s}}_B}{\partial \theta_{R_i}} = -\mathbf{C}\frac{\partial \mathbf{R}}{\partial \theta_{R_i}}\mathbf{\Psi}^{-1}(\mathbf{z} - \mathbf{H}\mathbf{s}_{\text{prior}}) \tag{33}$$

$$\frac{\partial \hat{\mathbf{s}}_G}{\partial \theta_{R_i}} = (-\mathbf{X\Omega}^{-1}\mathbf{A}^T - \mathbf{B} + \mathbf{CA\Omega}^{-1}\mathbf{A}^T)\mathbf{\Psi}^{-1}\frac{\partial \mathbf{R}}{\partial \theta_{R_i}}\mathbf{\Psi}^{-1}(\mathbf{z} - \mathbf{A\Omega}^{-1}\mathbf{A}^T\mathbf{\Psi}^{-1}\mathbf{z}) \tag{34}$$

The units of $\frac{\partial \hat{\mathbf{s}}}{\partial \theta_{Q_i}}$ and $\frac{\partial \hat{\mathbf{s}}}{\partial \theta_{R_i}}$ are of the form $(\mu\text{moles}^{-1}\text{m}^2\text{sec}^{-1})(\text{unit of } \theta_{Q_i} \text{ or } \theta_{R_i})^{-1}$. It is also possible to find $\frac{\partial \hat{\mathbf{s}}}{\partial \mathbf{Q}}$ and $\frac{\partial \hat{\mathbf{s}}}{\partial \mathbf{R}}$ directly as shown in equations 35 through 38.

https://doi.org/10.5194/gmd-2022-89
Preprint. Discussion started: 23 May 2022

Geoscientific Model Development Discussions
Open Access
EGU

$$\frac{\partial \hat{\mathbf{s}}_B}{\partial \mathbf{Q}} = \mathbf{H}^T \boldsymbol{\Psi}^{-1}(\mathbf{z} - \mathbf{H}\mathbf{s}_{\text{prior}}) \otimes \left(\mathbf{I} - \mathbf{H}^T \boldsymbol{\Psi}^{-1} \mathbf{B}^T\right) \tag{35}$$

$$\frac{\partial \hat{\mathbf{s}}_G}{\partial \mathbf{Q}} = \left(\mathbf{G}_z - \mathbf{z}^T\right) \boldsymbol{\Psi}^{-1} \mathbf{H} \otimes \left(\left(\mathbf{B} - \mathbf{M}\mathbf{A}^T + \mathbf{L}^T \mathbf{A}^T\right) \boldsymbol{\Psi}^{-1} \mathbf{H} - \mathbf{I}\right) \tag{36}$$

$$\frac{\partial \hat{\mathbf{s}}_B}{\partial \mathbf{R}} = \boldsymbol{\Psi}^{-1}(\mathbf{z} - \mathbf{H}\mathbf{s}_{\text{prior}}) \otimes \boldsymbol{\Psi}^{-1} \mathbf{H}\mathbf{Q} \tag{37}$$

$$\frac{\partial \hat{\mathbf{s}}_G}{\partial \mathbf{R}} = \left(\mathbf{G}_z - \mathbf{z}^T\right) \boldsymbol{\Psi}^{-1} \otimes \left(\mathbf{B} - \mathbf{M}\mathbf{A}^T + \mathbf{L}^T \mathbf{A}^T\right) \boldsymbol{\Psi}^{-1} \tag{38}$$

Equations 35 through 38 are useful when $\mathbf{Q}$ and $\mathbf{R}$ are fully or partially non-parametric. However, dimensions of these matrices can be quite large and users needs to be careful in realizing the full matrix.

3.3 Global sensitivity analysis: a variance-based approach

Global sensitivity analysis is a process of apportioning the uncertainty in an output estimate to the uncertainty in each input parameters. The term "global" stems from the idea of accounting for the effect of all input parameters simultaneously. This is different from "local" sensitivity analysis where the effect of a small change in each parameter on the functional output is separately considered keeping everything else constant. Although quite important, a detailed global sensitivity analysis is challenging due to it's demand for the knowledge of probabilistic variations of all possible combinations (also known as covariance) of the input parameters.

Broadly, we can consider $\hat{\mathbf{s}}$ as a function of the independent variables $\mathbf{Q}, \mathbf{R}, \mathbf{H}, \mathbf{X}(\text{or } \mathbf{s}_{\text{prior}})$, and $\mathbf{z}$ i.e. $\hat{\mathbf{s}} = f(\mathbf{Q}, \mathbf{R}, \mathbf{H}, \mathbf{X}(\text{or } \mathbf{s}_{\text{prior}}), \mathbf{z})$. We can then see how uncertainties of the individual components of f are accounted in the overall uncertainty of $\hat{\mathbf{s}}$. We apply multivariate Taylor series expansion of $\hat{\mathbf{s}}$ about it's mean. Approximation up to first-order polynomial of the Taylor series expansion leads to the equation:

$$V_{\hat{\boldsymbol{s}}} = \left(\frac{\partial \hat{\mathbf{s}}}{\partial \boldsymbol{\theta}}^T \mathbf{W}_\theta \frac{\partial \hat{\boldsymbol{s}}}{\partial \boldsymbol{\theta}}\right)_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} + \text{Error}, \quad \text{where}$$

$\boldsymbol{\theta} = (\boldsymbol{\theta_Q}, \boldsymbol{\theta_R}, \boldsymbol{\theta_H}, \theta_{\mathbf{X}}(\text{or } \mathbf{s}_{\text{prior}}), \boldsymbol{\theta_z})$ is the vector of parameters and $\mathbf{W} = \text{Var}(\boldsymbol{\theta})$ is the covariance matrix of the parameters. It is however, challenging to estimate some of the individual covariance quantities such as the cross-covariance between $\boldsymbol{\theta_R}$ and $\boldsymbol{\theta_H}$ or between $\boldsymbol{\theta_H}$, and $\boldsymbol{\theta_Q}$ to get the best possible decomposition of the total uncertainty of $\hat{\mathbf{s}}$. Assuming no cross-covariance between $\mathbf{Q}$ and $\mathbf{R}$ and ignoring other parameters not related to the variance parameters, the diagonal of the variance of the posterior fluxes can be approximated as:

$$V_{\hat{s}_i} = \sum_{j=1}^{L} \left(\frac{\partial \hat{\boldsymbol{s}}}{\partial \boldsymbol{\theta_Q}(j)}\right)_i^2 V_{\boldsymbol{\theta_Q}}(j) + \sum_{k=1}^{M} \left(\frac{\partial \hat{\boldsymbol{s}}}{\partial \boldsymbol{\theta_R}(k)}\right)_i^2 V_{\boldsymbol{\theta_R}}(k) \Bigg|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \tag{39}$$

https://doi.org/10.5194/gmd-2022-89
Preprint. Discussion started: 23 May 2022


Geoscientific Model Development Discussions

Where the subscript i on the right-hand side of equation 39 refers to the i^{th} element of the derivative vector which is a scalar and parameters $\boldsymbol{\theta_Q}(j)$ and $\boldsymbol{\theta_R}(k)$ refer to the jth and kth parameters of the sets $\boldsymbol{\theta_Q}$ and $\boldsymbol{\theta_R}$ respectively. From equation 39, we can see how uncertainty in the flux estimate is apportioned into variance components of $\boldsymbol{\theta_Q}$ and $\boldsymbol{\theta_R}$ of an inversion frameworks. No normalization is necessary in such global sensitivity analysis since on the right hand side of equation 39, the variance components are naturally weighted in such a way that both sides have same units. Once the two components of $V_{\hat{s}_i}$ (i.e. equation 39) are computed, they can also be summed over the solution space (e.g. number of gridcells $\times$ number of time-periods) of $\hat{\mathbf{s}}$ and ranked to find the relative importance of the parameters.

Even after simplification, implementation of equation 39 is difficult as it requires knowledge of uncertainties associated with the parameters of $\mathbf{Q}$ and $\mathbf{R}$ that are generally not known. Note that, it is also possible to have a complete apportionment of $V_{\hat{\mathbf{s}}}$ for all the parameters of f (e.g., $\mathbf{Q}, \mathbf{R}, \mathbf{H}, \mathbf{X}$(or $\mathbf{s}_{\text{prior}}$), and $\mathbf{z}$) at least up to first-order polynomial in the Taylor's series. However, it's implementation is difficult since it requires knowledge about the covariances of all the parameters. We do not further discuss global sensitivity analysis in the context of the case study presented in this work, but we have shown it's application with respect to $\mathbf{Q}$ and $\mathbf{R}$ in the MATLAB Live script.

3.4 Local sensitivity based relative importance of covariates, covariance parameters and observations

It is easy to compare importance of observations, covariates and covariance parameters within themselves as they have same units. However, the notion of relative importance amongst them is harder for two reasons. First, we need to standardize or normalize the local sensitivities such that they are in same units and preferably bounded (for discussion see Link and Doherty Jr, 2002). All the derivatives in this work result in vectors or matrices that may consist of negative, positive and extreme values and can be in different units. Normalization of parameters with different units has only been applied for scalars (i.e. a single parameter in a simulation model; for details see Saltelli et al., 2008). Normalization of sensitivity matrices is not common in literature. Second, we also need to adopt a technique to compare these senstivity matrices and find their relative importance in influencing $\hat{\mathbf{s}}$.

One of the ways to address the first problem is via global sensitivity analysis that is described in subsection 3.3. Although, it is a direct approach, it is not tenable in most scenarios. In this work, sensitivity matrices are harmonized via a two-step normalization. For each sensitivity matrix with same units (like $\frac{\partial \hat{\mathbf{s}}_B}{\partial H_{ij}}$), first, a simple min-max normalization (see Vafaei et al., 2020) is applied to each column to normalize the values between 0 to 1. The normalized columns are then aggregated by groups to form columns representing combined sensitivity of the groups. For example, a group can be observations of a particular instrument or set of estimation points. Then a second step min-max normalization is applied to each of these aggregated columns to transform them to vary between 0 and 1.

Once the normalized sensitivity vectors are obtained for each quantity of interest, the objective is to compute a score for each of these vectors such that these scores can be ranked. To authors knowledge, no standard methods exist for such comparisons.

https://doi.org/10.5194/gmd-2022-89
Preprint. Discussion started: 23 May 2022


Geoscientific Model Development Discussions

An approach such as principal component analysis (PCA) (Jolliffe and Cadima, 2016) could have been used to answer this question. However, the underlying question here is to rank the sensitivities with respect to their importance in influencing the estimate of $\hat{\mathbf{s}}$ which cannot be achieved by PCA. Hence, we adopt a regression-based approach to assess the relative importance of covariates, covariance parameters and observations. In this work, we define the regression model as:

$$\mathbf{\Gamma} = \mathbf{E}\gamma + \xi \tag{40}$$

where $\mathbf{\Gamma}$ are fluxes obtained from an inversion, $\mathbf{E}_{(m,\text{number of derivatives})}$ is a matrix of the previously estimated normalized sensitivities or derivatives with dimensions, $\gamma_{(\text{number of derivatives},1)}$ are the unknown coefficients for the sensitivities or derivative vectors, and $\xi_{(m,1)}$ is the unobserved error associated with the regression model. To exemplify, $\mathbf{E}$ in equation 40 can be arranged as:

$$\mathbf{E} = \left[\frac{\partial \hat{\mathbf{s}}}{\partial \mathbf{z}} \; \frac{\partial \hat{\mathbf{s}}}{\partial \mathbf{Q}} \; \frac{\partial \hat{\mathbf{s}}}{\partial \mathbf{R}} \; \cdot \; \cdot \right] \tag{41}$$

where γ is the vector of relative importance weights. In a regression-based approach, as described in equation 40, multicollinearity between independent variables in $\mathbf{E}$ can pose a problem for determining the importance of independent variables in influencing $\mathbf{\Gamma}$. To avoid this problem, we computed relative importance weights by using the method outlined in Johnson, 2000. These weights are computed by first deriving uncorrelated orthogonal counterparts of the independent variables in $\mathbf{E}$ and then regressing dependent variable on them to get importance weights for each independent variable. The weights are standardized by the coefficient of determination i.e., R^2 such that they range between 0 to 1 with the sum of all the weights being 1. A detailed description of this method is given in Johnson, 2000 and the implementation of this method is included in the Live script submitted with this manuscript. Note that an approach of LASSO could also have been employed here to obtain the relative weights of the predictors under multicollinearity. However, “inference after selection" is ambiguous in linear regression which is the case for LASSO coefficients (see Berk et al., 2013 or chapter 6 of Hastie et al., 2015 for details). Consequently, interpreting the LASSO coefficients as relative importance scores may not be the best approach here. Thus, we do not use this technique here.

Note that most analytical inversions use DOFS to diagnose information content of an inversion. DOFS = 0 implies that no informational gain happened in an inversion. In this case, estimated flux reverts back to prior. In Equation 40, this means that the γ coefficient that corresponds to $\mathbf{Q}$ would be the largest. Likewise if DOFS is large, then γ coefficients for $\mathbf{z}$ and $\mathbf{R}$ should be larger (and likely correlated). We show this correspondence with standard approaches in section 4.

Finally, all different kinds of diagnostic methods that are applied in the context of any regression-based model can be used for understanding the relationship between dependent and independent variables. However, what independent variables to include in $\mathbf{E}$ depends on the specific case study under consideration.

https://doi.org/10.5194/gmd-2022-89
Preprint. Discussion started: 23 May 2022


Geoscientific Model Development Discussions

4 Results: Los Angeles methane inversion case study

To demonstrate the applicability of our methods we utilize data from our published work on CH_4 fluxes in the Los Angeles megacity (see Yadav et al., 2019). In this previous work, fluxes were estimated for South Coast Air Basin (SoCAB) region (Figure 2) at 0.03° spatial (1826 grid-cells) and 4-day temporal resolution from the Jan 27, 2015 through Dec 24, 2016. However, in the current work we utilize input data from Oct 23, 2015 through Oct 31, 2015 that is a single inversion period to contextualize the applicability of our methods. This period overlaps with the beginning of the now well-studied Aliso Canyon gas leak (Conley et al., 2016). We do not extend our analysis for the full duration of the previous study as this is not the objective of this work and all the details associated with the inverse flux estimates can be found in that work. Furthermore, we present our sensitivity based equations with respect to the geostatistical approach to inverse modeling as this was the approach adopted in the previous study.

For each observation included in the case study, a footprint was obtained by using Weather Research Forecasting-Stochastic Time Inverted Lagrangian Model (see Yadav et al., 2019). These footprints are used to demonstrate the application of the methodology for building IOAMI and JSD based correlation matrices in the MATLAB Live script. They are also used in conjunction with measurements, and prior information to estimate the fluxes and perform LSA.

4.1 Spatio-temporal area of dominance (STAD) from the footprints

In this work we identify STAD for the 4-day period for which inversion was performed. The spatial domain of the study over this time period is uniquely disaggregated by STAD as shown in Figure 2. The STAD for different sites are mostly spatially contiguous but for some sites we found isolated grid cells which were not within contiguous zones. We have manually combined these with STAD for the nearest site to create a spatially continuous map as shown in Figure 2. The discontinuous version of the STAD shown in Figure 2 is included in the Livescript. The discontinuities in the STAD mostly result from unequal number of observations across sites and indicates that aggregation over longer time-period is required to completely identify a noise free STAD. We do not investigate the time-period of this aggregation as this is beyond the scope of this work.

Overall, the STAD for each site indicates regions of emissions that contributes most to the observational (e.g. CH4 enhancement) signal. This in turn allows us to sub-divide the spatio-temporal variations in fluxes or enhancements by the STADs.

4.2 Sensitivity analysis

One of the main goals of the sensitivity analysis after performing inversions is to identify the observations that had most influence on the flux estimates. Overall assessment of the importance of different inputs to an inversion after observations is also important to explore. We describe the process of performing this analysis within the context of the case study mentioned in section 4. This section discusses the relative importance of the input quantities in influencing $\hat{\mathbf{s}}$ via the local sensitivities.

https://doi.org/10.5194/gmd-2022-89
Preprint. Discussion started: 23 May 2022

Geoscientific Model Development Discussions
Open Access
EGU

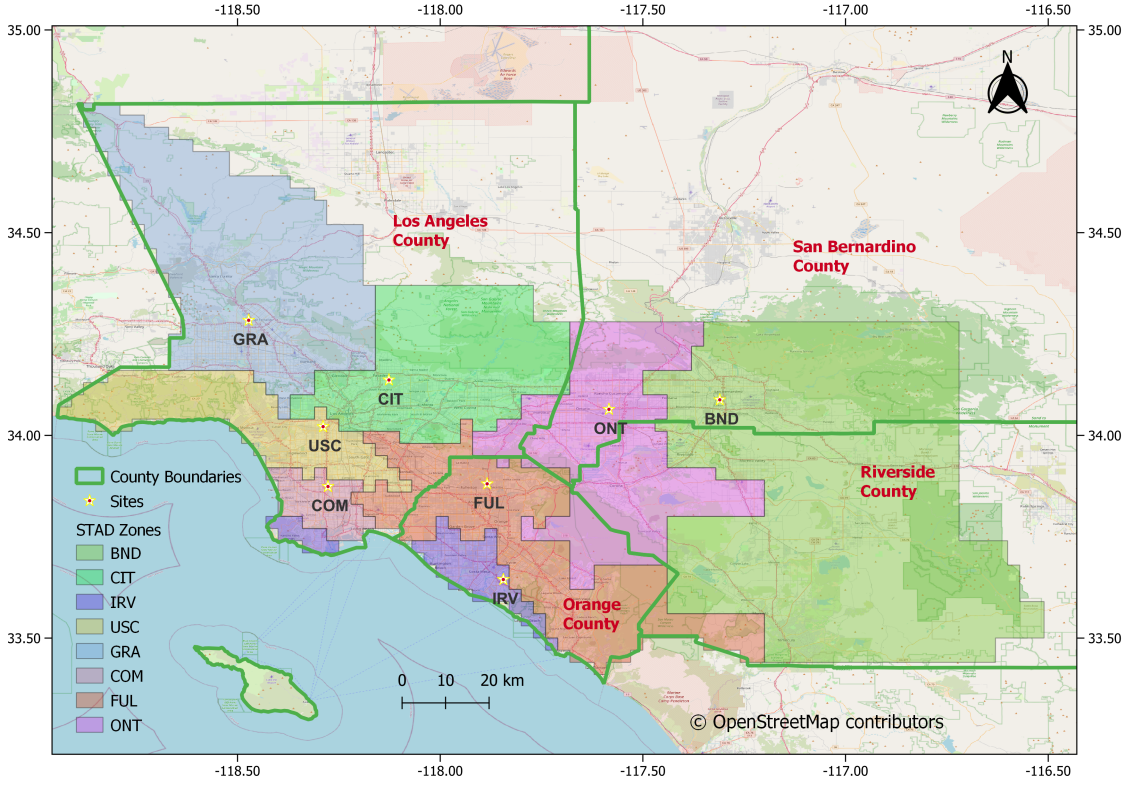
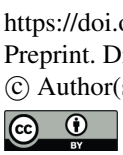


Figure 2. Study area with county boundaries, measurement locations and Spatio-Temporal Area of Dominance of measurement locations.

Site	Importance Score	Rank
GRA	0.26	1
ONT	0.24	2
COM	0.13	3
IRV	0.11	4
BND	0.10	5
CIT	0.07	6
FUL	0.07	7
USC	0.06	8

Table 1. The importance scores and ranking of 8 sites based on the sensitivity of the estimated fluxes ($\hat{\mathbf{s}}$) to observations ($\mathbf{z}$).

https://doi.org/10.5194/gmd-2022-89
Preprint. Discussion started: 23 May 2022



Geoscientific Model Development Discussions
Open Access
EGU

4.2.1 Comparison and ranking of the observations

Importance of individual measurements in influencing $\hat{\mathbf{s}}$ can be easily computed through relative importance methodology described in section 3.4. Although, all entries of $\frac{\partial \hat{\mathbf{s}}}{\partial \mathbf{z}}$ are in same units, direct ranking of observations or sites without employing relative importance technique can lead to misleading results. This happens due to presence of large negative and positive values in $\frac{\partial \hat{\mathbf{s}}}{\partial \mathbf{z}}$ that are governed by the overall spatio-temporal spread, intensity of footprints, and large observations.

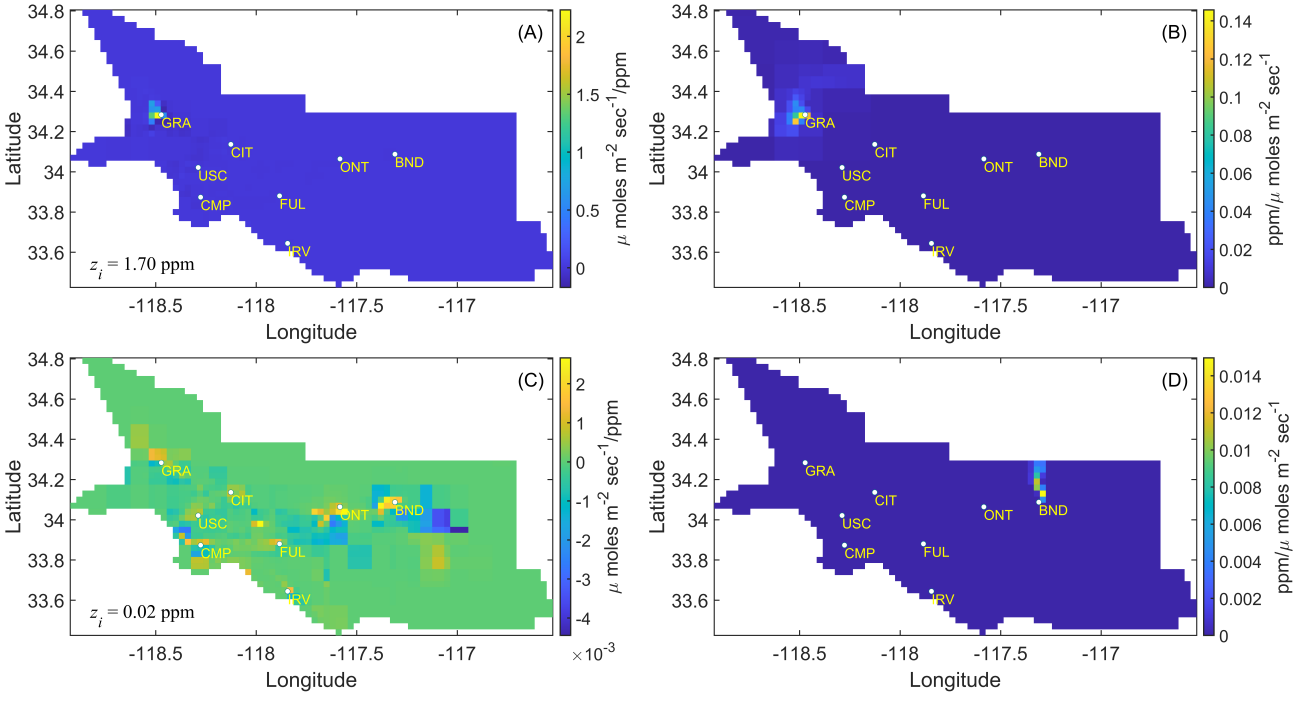


Figure 3. The sensitivities ($\frac{\partial \hat{\mathbf{s}}}{\partial \mathbf{z}}$) and footprints ($\mathbf{H}$) of the most and least important observation in inversions. Subplot A and C show the sensitivity of the estimated fluxes with respect to the most (A) and least important (C) observation. The CH_4 enhancement associated with these observations is shown in the bottom left corner of the subplots and identified by symbol z_i. The right subplots B and D show footprints associated with the sensitivities shown in subplots A and C respectively.

For the case study in this work, we find that observations collected at the GRA site that is located nearest to the source of Aliso Canyon gas leak are most influential in governing $\hat{\mathbf{s}}$ as shown by site-based rankings in Table 1. These rankings primarily show the importance of observations from a site in influencing the estimated fluxes for the time period in consideration. Observation based assessment of $\frac{\partial \hat{\mathbf{s}}}{\partial \mathbf{z}}$ resulted in ranking an observation with the largest enhancement of 1.7 ppm to be most important. Contrarily, an observation for the BND site that had an enhancement of 0.02 ppm is found to be least important in influencing $\hat{\mathbf{s}}$. Note this is not an observation with lowest enhancement but with lowest influence. The most and least important observation along with their corresponding footprints are shown in Figure 3.

https://doi.org/10.5194/gmd-2022-89
Preprint. Discussion started: 23 May 2022

Geoscientific Model Development Discussions
Open Access
EGU

4.2.2 Relative importance of $\mathbf{Q}, \mathbf{R}, \mathbf{X}, \beta$, and $\mathbf{z}$

After the two-step normalization of $\frac{\partial \hat{\mathbf{s}}}{\partial \mathbf{z}}, \frac{\partial \hat{\mathbf{s}}}{\partial \mathbf{X}}, \frac{\partial \hat{\mathbf{s}}}{\partial \mathbf{H}}, \frac{\partial \hat{\mathbf{s}}}{\partial \beta}, \frac{\partial \hat{\mathbf{s}}}{\partial \mathbf{Q}}$, and $\frac{\partial \hat{\mathbf{s}}}{\partial \mathbf{R}}$ as described in section 3.4, the spatial plots of all these quantities can be created to explore the regions of the low and high weights at the grid scale as shown in Figure 4.

Figure 4 shows that the weights of $\frac{\partial \hat{\mathbf{s}}}{\partial \mathbf{X}}$ is lower in the regions well constrained by the observations. However, the opposite is true in the case of $\frac{\partial \hat{\mathbf{s}}}{\partial \mathbf{Q}}$ and $\frac{\partial \hat{\mathbf{s}}}{\partial \mathbf{R}}$. This implies, that data constrained regions have lower posterior uncertainty thereby increasing the influence of prescribed or estimated uncertainty parameters. There is smoothness in the weights of $\frac{\partial \hat{\mathbf{s}}}{\partial \mathbf{Q}}$ in the domain except around some sites, which is an indication that the estimates of $\hat{\mathbf{s}}$ remain insensitive to the $\mathbf{Q}$ parameter in these regions. These relationships can be quantified by assessing correlation between local sensitivities and $\hat{\mathbf{s}}$ as shown in Figure 5.

There is strong evidence of multicollinearity among independent variables in explaining the dependent variable (e.g. see first column of the figure 5). The direction of the best fit line appears to be in sync with the expectation regarding CH_4 fluxes in the region during that time period. Thus, $\frac{\partial \hat{\mathbf{s}}}{\partial \mathbf{z}}$ is positively correlated with $\hat{\mathbf{s}}$, which implies that higher enhancement in $\mathbf{z}$ leads to an increase in the estimated fluxes. Similar to $\frac{\partial \hat{\mathbf{s}}}{\partial \beta}$ is also positively correlated with $\hat{\mathbf{s}}$ implying that any increase in the scaling factor increases the estimated fluxes. The negative relationship of $\frac{\partial \hat{\mathbf{s}}}{\partial \mathbf{X}}$ and $\hat{\mathbf{s}}$ just indicates that an increase in $\frac{\partial \hat{\mathbf{s}}}{\partial \mathbf{X}}$ inversely influences the magnitude of the estimated fluxes. This occurs as $\hat{\mathbf{s}}$ reverts to $\mathbf{X}$ in regions unconstrained by observations whereas opposite happens in areas constrained by observations that in the context of the case study includes sources of largest emissions.

https://doi.org/10.5194/gmd-2022-89
Preprint. Discussion started: 23 May 2022

Geoscientific Model Development
Discussions
Open Access
EGU

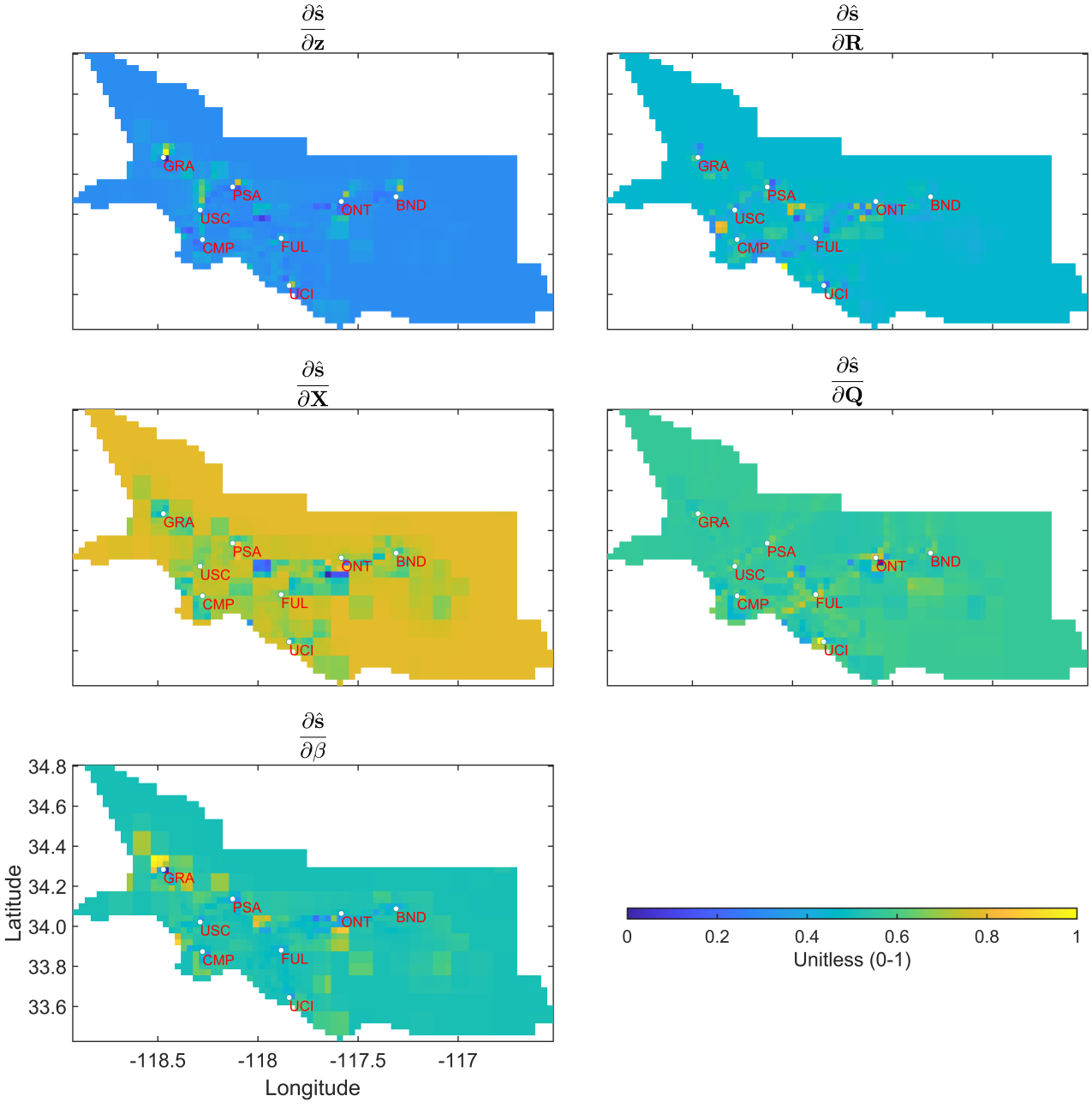


Figure 4. Local sensitivities of the estimated fluxes ($\hat{\mathbf{s}}$) with respect to $\mathbf{z}$, $\mathbf{R}$, $\mathbf{X}$, $\mathbf{Q}$, and β from top-left to bottom-right respectively. Note, in the case of $\frac{\partial \hat{\mathbf{s}}}{\partial \mathbf{z}}$, $\frac{\partial \hat{\mathbf{s}}}{\partial \mathbf{R}}$, and $\frac{\partial \hat{\mathbf{s}}}{\partial \mathbf{X}}$ two-step normalization is performed to generate subplots associated with these quantities. Derivatives with respect to: (1) observations in $\mathbf{z}$, (2) parameters in $\mathbf{R}$, and (3) entries in $\mathbf{X}$ are normalized between 0 and 1 and then after aggregating these for every grid-cell another Min-Max normalization is performed to limit their ranges between 0 and 1. Only single normalization is performed in case of $\frac{\partial \hat{\mathbf{s}}}{\partial \mathbf{Q}}$ and $\frac{\partial \hat{\mathbf{s}}}{\partial \beta}$ as they consist of only one parameter.

https://doi.org/10.5194/gmd-2022-89
Preprint. Discussion started: 23 May 2022

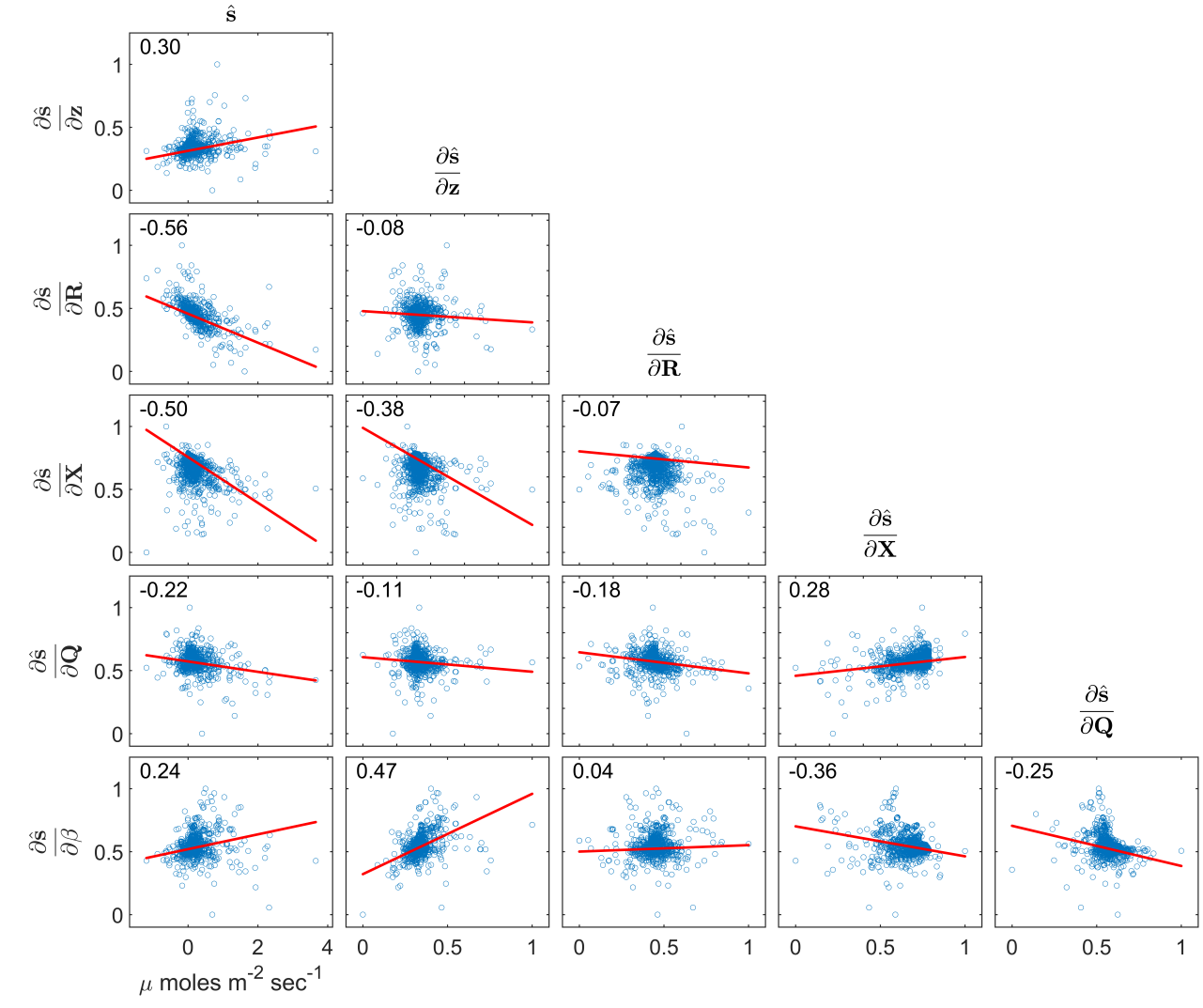


Figure 5. Scatterplots of relationships between $\hat{\mathbf{s}}$ and $\frac{\partial \hat{\mathbf{s}}}{\partial \mathbf{z}}$, $\frac{\partial \hat{\mathbf{s}}}{\partial \mathbf{R}}$, $\frac{\partial \hat{\mathbf{s}}}{\partial \mathbf{X}}$, $\frac{\partial \hat{\mathbf{s}}}{\partial \mathbf{Q}}$, $\frac{\partial \hat{\mathbf{s}}}{\partial \beta}$. Note as before in Figure 4 all the derivatives are normalized to limit their range between 0 and 1 whereas $\hat{\mathbf{s}}$ has units of μmoles^{-1}m^{2}sec^{-1}. The correlation coefficient of the relationships shown in each scatterplot is reported on the top right corner of the subplots. The least square line of best fit is shown in red color in every subplot.

https://doi.org/10.5194/gmd-2022-89
Preprint. Discussion started: 23 May 2022


Geoscientific Model Development Discussions

5 Discussion

This work provides diagnostic tools to investigate components of two inverse modeling framework as well as assesses quality of the inferred estimates. Sensitivity analysis is an important diagnostic tool to understand the impact of the choices made with respect to inputs on the estimated fluxes. However, it is not a recipe for selecting proper forms of $\mathbf{X}$ or structure of $\mathbf{Q}$ and $\mathbf{R}$ before performing an inversion. Other tools or methods such as Bayesian Information Criterion, Variance Inflation Factor should be used to perform this task.

The case study in this work is designed only to demonstrate the methodologies described in section 3. We do not impose non-negativity constraints to obtain positive CH_4 fluxes as was done in the original 2019 study (Yadav et al., 2019). This is done because posterior likelihood changes its functional form under non-negativity constraints and the analytical forms of senstivity equations presented in this work become invalid. Thus, some CH_4 fluxes obtained in this study have negative values as can be seen in the map of $\hat{\mathbf{s}}$ in the MATLAB Live script. However, even in these situations assessing sensitivity through an inversion without imposition of non-negativity is useful as it provides insights into the role of $\mathbf{z}$, $\mathbf{R}$, $\mathbf{Q}$, and $\mathbf{X}$ in governing estimates of non-negative $\hat{\mathbf{s}}$.

Like $\mathbf{z}$, the importance of $\mathbf{Q}$ and $\mathbf{R}$ parameters can be directly obtained when all parameters have the same units. This happens in the case study presented in this work. However, this is not guaranteed as $\mathbf{R}$ can be a function of variance parameters and spatio-temporal correlation length expressed in the distance units in space and time. Furthermore, a nonstationary error covariance $\mathbf{R}$ can have parameters that have even more complicated units. This situation is not limited to $\mathbf{R}$ and also applies to the prior error covariance $\mathbf{Q}$ and $\mathbf{X}$. Under these conditions, a comparison between the sensitivity matrices is only possible after normalization. Therefore, for comparative assessment we recommend use of a multiple linear regression based relative importance method to rank these quantities.

The importance of $\frac{\partial \hat{\mathbf{s}}}{\partial \mathbf{z}}$ is best explored by first performing column based normalization and then employing the relative importance method. Additionally, column based normalization can be augmented by row-based normalization to assess and rank influence of observations in governing gridscale estimates of $\hat{\mathbf{s}}$. Qualitatively, column and row-based assessment increase our understanding about the spatio-temporal estimates of $\hat{\mathbf{s}}$. This is especially important when point sources are the dominant sources of emissions. Moreover, it also provides an insight into temporal aggregation error as the information encoded in an instantaneous measurement can get lost over the coarser time-period of inversion. This aggregation error also manifests spatially and is determined by the resolution at which fluxes are obtained. Note in many situations these aggregation errors are unavoidable as the choice of the spatio-temporal resolution of inversions is governed by the density of observations in space and time.

https://doi.org/10.5194/gmd-2022-89
Preprint. Discussion started: 23 May 2022


Geoscientific Model Development Discussions

Other than aggregation error, the aggregation of the estimated fluxes also has profound implications as it affects the robustness of the estimated fluxes. It can be proved (see Appendix A) that aggregation of $\hat{\mathbf{s}}$ in space and time from an inversion conducted at finer resolution leads to reduction in uncertainty. However, even though ratio of observations to the estimated fluxes increases the number of fluxes uniquely resolved declines at coarser resolution (see Appendix B).

Computing the Kronecker form of the local sensitivities (equations 20, 23 through 26, and 35 though 38) is faster for small problems. However for large inverse problems storage costs associated with these equations can become prohibitive. In these situations, we propose the use of ij form of the equations (equations 22, 27 through 30, and 31 though 34) for assessment. Furthermore, computational problems can also arise in importance ranking if we have large number derivatives (e.g. greater than 10,000) as relative importance method uses eigen value decomposition that has $O(n^3)$ computational complexity. To overcome this problem we advise grouping of derivatives to reduce the dimension of the problem.

Finally, the estimation of STAD and importance of sites can be influenced by data gaps therefore is not advised in presence of vast differences in the number of observations between sites. Furthermore, if observations from different platforms (e.g. aircraft, satellites and in-situ sites) are used in an inversion then combined ranking of observational platforms is untenable as spatio-temporal densities of measurements are different across platforms.

6 Conclusions

Our work makes novel and major contributions that can significantly improve understanding of linear inverse problems. It provides: (1) a way to understand the correlations in the footprints or atmospheric transport model, and (2) a framework for post hoc analysis of the impact of inputs on the estimated fluxes. The authors are not aware of any work where local sensitivities with different units are compared to rank the importance of inputs to an inversion model.

With respect to footprints, we provide mathematical foundations for IOAMI, and Jensen-Shannon based metrics. These two metrics can be used to construct and accommodate a non-stationary error covariance for atmospheric transport component of the model-data mismatch matrix $\mathbf{R}$. Furthermore, IOAMI based assessments can be extended to identify STAD from footprints that help in disaggregating regions of influence of observations over a chosen temporal duration. This assists in understanding the connection between the sources of emissions and observations from a particular measurement location.

The IOAMI and JSD based metrics provide an important insight into two critical and only required components for an inversion that is observations and footprints (e.g., influence of an observation to sources of emissions through STAD). This task can be accomplished prior to conducting an inversion and should be complimented by post hoc LSA, which is a necessity for understanding the behavior of an inverse model. Overall, LSA can answer questions like for which locations and in what order of precedence was an observation important in influencing the estimated fluxes. This kind of analysis is entirely different

https://doi.org/10.5194/gmd-2022-89
Preprint. Discussion started: 23 May 2022

Geoscientific Model Development Discussions

from estimating uncertainty that tells us reduction in the prior uncertainty due to observations.

LSA is not a replacement for statistical tests that check the underlying assumptions and model specifications of inverse models. Neither is it a recipe for selecting inputs to an inverse model. However, it has an important role as explained above that can lead to an improved understanding of an inverse model.

Code and data availability. All the code and data utilized in this study are submitted as supplementary material.

Appendix

Here we show the proofs of two mathematical statements on the robustness and quality of the estimated fluxes as mentioned in section 5. First, we show why marginal variance of the estimated fluxes (which is the diagonal of covariance matrix of $\hat{\mathbf{s}}$) decrease when estimated fluxes are post aggregated to a coarser scale or upscaled (A). Second, we show why in such case the model resolution (also termed as, total information resolved by the observations) also decreases (B). Note that, the nomenclature used in the appendix should not be confused with the nomenclature introduced in section 3. The abbreviations and symbols used here are independent of what are used in the section 3.

Appendix A: Proof of the reduction of marginal variance of $\hat{\mathbf{s}}$ when upscaling is performed

Post inversion upscaling of any flux field $\mathbf{s}$ is equivalent to premultiplication by a weight matrix (in fact, a row stochastic matrix). This can be written as:

$$\tilde{\mathbf{s}} = \mathbf{J}\hat{\mathbf{s}} \tag{A1}$$

Where $\mathbf{J}$ is a row stochastic (i.e. row-sums are all unity) $k \times p$ weight matrix ($k < p$). Variance of $\tilde{\mathbf{s}}$ can be written as $\mathbf{J\Sigma J^t}$ where var($\tilde{\mathbf{s}}$) = $\mathbf{J}$var($\hat{\mathbf{s}}$)$\mathbf{J^t}$=$\mathbf{J\Sigma J^t}$. The general structure of $\mathbf{J}$ is as follows:

$$J = \begin{bmatrix} \cdot & j_{12} & j_{13} & \cdots & \cdots & \cdots \\ j_{21} & \cdots & j_{2r+1} & j_{2r+2} & \cdots & \cdots \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ \cdots & \cdots & \cdots & j_{km} & \cdots & \cdots \end{bmatrix} = \begin{bmatrix} \mathbf{j_1}' \\ \mathbf{j_2}' \\ \vdots \\ \mathbf{j_k}' \end{bmatrix} \tag{A2}$$

Essentially, $\mathbf{J}$ can have any number of non-zero entries in a row that may or may not be consecutive. This is because although on a map, adjacent grids are averaged, they may not be adjacent upon vectorization. Moreover, geometry of the map may not

https://doi.org/10.5194/gmd-2022-89
Preprint. Discussion started: 23 May 2022

Geoscientific Model Development Discussions

be exactly square or rectangular. This means, depending on the upscaling factor and geometry, for any particular grid, there may or may not any neighboring grid for averaging. However, the rows are linearly independent as nearby grids are considered once for averaging. Zeros are represented as dots in the example structure of $\mathbf{J}$ (A2). So, the properties of $\mathbf{J}$ are as follows:

1. $\mathbf{J1} = \mathbf{1}$ or $\sum_{j=i}^{p} j_{ij} = 1 \;\; \forall i = 1, \ldots, k$
2. $\mathbf{j_i' j_l} = \mathbf{0}$ for $i \neq l$

We can rearrange the columns of $\mathbf{J}$ and the rows of $\mathbf{\Sigma}$ accordingly without loss of any structure such that non-zero elements are consecutive for each row of $\mathbf{J}$. Matrix $\mathbf{J\Sigma J'}$ under column permutation can be written as:

$$\mathbf{J\Sigma J^t} = \mathbf{J^P \Sigma^P J^{Pt}} = \begin{bmatrix} \mathbf{J^P{}_1{}^t} & 0 & \ldots & 0 \\ 0 & \mathbf{J^P{}_2{}^t} & \ldots & 0 \\ \vdots & \vdots & \ddots & . \\ 0 & 0 & \ldots & \mathbf{J^P{}_k{}^t} \end{bmatrix}^{k \times p} \begin{bmatrix} \mathbf{\Sigma^P_{11}} & \mathbf{\Sigma^P_{12}} & \ldots & \mathbf{\Sigma^P_{1k}} \\ \mathbf{\Sigma^P_{21}} & \mathbf{\Sigma^P_{22}} & \ldots & . \\ \vdots & \vdots & \ddots & . \\ \mathbf{\Sigma^P_{k1}} & . & \ldots & \mathbf{\Sigma^P_{kk}} \end{bmatrix}^{p \times p} \begin{bmatrix} \mathbf{J^P{}_1{}^t} & 0 & \ldots & 0 \\ 0 & \mathbf{J^P{}_2{}^t} & \ldots & 0 \\ \vdots & \vdots & \ddots & . \\ 0 & 0 & \ldots & \mathbf{J^P{}_k{}^t} \end{bmatrix}^{p \times k} \tag{A3}$$

$$= \begin{bmatrix} \mathbf{J^P{}'_1 \Sigma^P_{11} J^P_1} & . & \ldots & \mathbf{J^P{}'_1 \Sigma^P_{1k} J^P_k} \\ . & \mathbf{J^P{}'_2 \Sigma^P_{22} J^P_2} & \ldots & . \\ \vdots & \vdots & \ddots & . \\ \mathbf{J^P{}'_k \Sigma^P_{k1} J^P_1} & . & \ldots & \mathbf{J^P{}'_k \Sigma^P_{kk} J^P_1} \end{bmatrix}^{k \times k} \tag{A4}$$

where $\mathbf{J^P}$ and Σ^p are the column and row permuted $\mathbf{J}$ and Σ respectively. Note that, any $\mathbf{J^P_i{}'}$ is a row-vector of dimension $(1, d_i)$, and $\Sigma^p{}_{ii}$ is a square matrix of dimension (d_i, d_i) where $\sum_{i=1}^{k} d_i = p$. Thus, diagonal entry $\mathbf{J^P_i{}^t \Sigma^P_{ii} J^P_i}$ is a scalar quantity. For any i^{th} diagonal entry, the corresponding scalar quantity can be written as $\sum_{jl} a^p_{ij} a^p_{il} \sigma^p_{jl}$ where superscript p refers to the corresponding matrix after permutation. By symmetry of $\mathbf{\Sigma^P}$, this reduces to

$$\mathbf{J^P_i{}^t \Sigma_{ii} J^P_i} = \sum_{l} {a^p_{il}}^2 {\sigma^p_{ll}}^2 + 2 \sum_{j>l} a^p_{ij} a^p_{ij} \sigma^p_{jl} \tag{A5}$$

By Cauchy Squartz inequality on σ^p_{jl}, this can be written as

$$\sum_{l} {a^p_{il}}^2 {\sigma^p_{ll}}^2 - 2 \sum_{j>l} a^p_{ij} a^p_{ij} \sigma^p_{jj} \sigma^p_{ll} \leq \sum_{l} {a^p_{il}}^2 {\sigma^p_{ll}}^2 + 2 \sum_{j>l} a^p_{ij} a^p_{ij} \sigma^p_{jl} \leq \sum_{l} {a^p_{il}}^2 {\sigma^p_{ll}}^2 + 2 \sum_{j>l} a^p_{ij} a^p_{ij} \sigma^p_{jj} \sigma^p_{ll} \tag{A6}$$

$$\left(a^p_{i1} \sqrt{\sigma^p_{i1}} - \sum_{l \geq 2} a^p_{il} \sqrt{\sigma^p_{il}} \right)^2 \leq \sum_{l} {a^p_{il}}^2 {\sigma^p_{ll}}^2 + 2 \sum_{j>l} a^p_{ij} a^p_{ij} \sigma^p_{jj} \sigma^p_{ll} \leq \left(\sum_{il} a^p_{il} \sqrt{\sigma^p_{ll}} \right)^2 \tag{A7}$$

$$\min_l \sigma^p_{ll} \left(a^p_{i1} - \sum_{l \geq 2} a^p_{il} \right)^2 \leq \sum_{l} {a^p_{il}}^2 {\sigma^p_{ll}}^2 + 2 \sum_{j>l} a^p_{ij} a^p_{ij} \sigma^p_{jj} \sigma^p_{ll} \leq \max_l \sigma^p_{ll} \left(\sum_{il} a^p_{il} \right)^2 \tag{A8}$$

https://doi.org/10.5194/gmd-2022-89
Preprint. Discussion started: 23 May 2022

Geoscientific Model Development Discussions

This implies (by property 1 of the weight matrix $\mathbf{J}$) that the ith diagonal element is bounded by:

$$\min_l \sigma^p_{ll} \left(a^p_{i1} - \sum_{l \geq 2} a^p_{il} \right)^2 \leq \mathbf{J_i' \Sigma_{ii} J_i} \leq \max_l \sigma^{\mathbf{p}}_{\mathbf{ll}} \leq \sum_{\mathbf{r=1}}^{\mathbf{d_i}} \sigma^{\mathbf{p}}{}_{\mathbf{rr}} \tag{A9}$$

where $\sum_{r=1}^{d_i} \sigma^p{}_{rr}$ is the sum of the marginal variance of the ith block of un-averaged $\hat{\mathbf{s}}$. Thus, sum of the marginal variance of $\tilde{\mathbf{s}}$ which is the sum of the ith diagonal $\mathbf{J^p{}_i{}^t \Sigma_{ii} J^p_i}$ is also smaller or equals to the sum total of marginal variance of $\hat{\mathbf{s}}$. Clearly, we see that under upscaling or averaging, diagonal of the variance matrix shrinks in magnitude from the un-averaged one. As a consequence, it implies that marginal variance of the posterior mean decreases.

Appendix B: Proof of the reduction in model resolution when upscaling is performed

Upscaled footprint operator $\tilde{\mathbf{H}}$ can be written as:

$$\tilde{\mathbf{H}} = \mathbf{HB} \quad \text{where } \mathbf{B} \text{ is the upscaling matrix} \tag{B1}$$

Dimension $\mathbf{B}$ is the dimension of transpose of $\mathbf{J}$. Form of $\mathbf{B}$ is similar to form of $\mathbf{J}$ explained in A2. Non-zero entries of $\mathbf{B}$ are in the same place as $\mathbf{J}'$ with magnitude being replaced by unity. This is evident from the fact that footprint operator is summed instead of average over the same grids for upscaling. Properties of $\mathbf{B}$ are as follows:

1. $\mathbf{B1 = 1}$
2. $\mathbf{JB = diag(N)^{k \times k}}$ where $\mathbf{N}$ is the vector of number of neighboring grids for any particular grid i.e. $\mathbf{N} = (N_1, \ldots, N_k)$
3. $\mathbf{BJ} = \begin{bmatrix} \mathbf{C_1} & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \mathbf{C_2} & \ldots & \mathbf{0} \\ \vdots & \vdots & \ddots & . \\ \mathbf{0} & . & \ldots & \mathbf{C_k} \end{bmatrix}^{\mathbf{p \times p}}$ is a block diagonal matrix Any block $\mathbf{C_i}$ of $\mathbf{JA}$ can be expressed as a varying dimension (depending on the number of neighboring grids of any particular grid) matrix of form:

$$\mathbf{C_i} = \begin{bmatrix} \frac{1}{N_i} & \cdots & \frac{1}{N_i} \\ \vdots & \ddots & \vdots \\ \frac{1}{N_i} & \cdots & \frac{1}{N_i} \end{bmatrix}^{\mathbf{N_i \times N_i}} = \frac{\mathbf{1}}{\mathbf{N_i}} \mathbf{11'} \tag{B2}$$

4. $\mathbf{BJ}$ is symmetric and positive semi-definite

Proof: $Det(\mathbf{BJ} - \lambda \mathbf{I}) = \mathbf{Det(C_1} - \lambda \mathbf{I}) \ldots . \mathbf{Det(C_k} - \lambda \mathbf{I})$. So, eigen values of $\mathbf{BJ}$ are the list of eigen values of the block matrices. It can be proved that 1 and 0 are the only two distinct eigen values of $\mathbf{C_i}$ for any i. Below here is a brief

https://doi.org/10.5194/gmd-2022-89
Preprint. Discussion started: 23 May 2022


Geoscientific Model Development Discussions

argument on that:

$\left(\frac{1}{N_i}\mathbf{11'}\right)\mathbf{1}=\frac{1}{N_i}\mathbf{1N_i}=\mathbf{1}\cdot\mathbf{1}$ implies one eigen value of $\mathbf{C_i}$ is 1. Observe that, $rank\left(\frac{1}{N_i}\mathbf{11'}\right)=rank(\mathbf{1})=\mathbf{1}$. Hence, dimension of null space $dim\left(\mathcal{N}\left(\frac{1}{N_i}\mathbf{11'}\right)\right)=k-rank\left(\frac{1}{N_i}\mathbf{11'}\right)=k-1$. This implies that the other eigen value of $\mathbf{C_i}$ is 0 with multiplicity $k-1$.

So, not only $\mathbf{C_i}$ is symmetric but the eigen values $\mathbf{C_i}$ are always non negative. Consequently, all eigen values of $\mathbf{BJ}$ are of similar form or $\mathbf{BJ}$ is symmetric positive semidefinite.

Model resolution matrix for inversion can be written as $\frac{\partial\hat{\mathbf{s}}}{\partial\mathbf{z}}\mathbf{H}$ where $\mathbf{H}$ is the footprint operator. Post inversion aggregated model-resolution can be written as:

$$\frac{\partial\tilde{\mathbf{s}}}{\partial\mathbf{z}}\tilde{\mathbf{H}}=\mathbf{A}\frac{\partial\hat{\mathbf{s}}}{\partial\mathbf{z}}\mathbf{HB} \quad \text{By equation A1 and B1} \tag{B3}$$

The question is what happens to the trace of the model-resolution under the upscaled case? We prove it for the simple batch Bayesian case. Proof for the geostatistical case is similar and left for the enthusiastic readers.

Lemma 1.

$$\mathbf{Mres}=\mathbf{QH'}\psi^{-\mathbf{1}}\mathbf{H}$$

$$\mathbf{Mres_{agg}}=\mathbf{JQH'}\psi^{-\mathbf{1}}\mathbf{HB} \quad \textit{by B3, then}$$

$$trace(\mathbf{Mres_{agg}})\leq trace(\mathbf{Mres}) \tag{B4}$$

Proof: Model resolution for the aggregated case can be written as:

$$trace(\mathbf{Mres_{agg}})=\mathbf{trace}(\mathbf{JQH'}\psi^{-\mathbf{1}}\mathbf{HB})=\mathbf{trace}(\mathbf{BJQH'}\psi^{-\mathbf{1}}\mathbf{H})=\mathbf{trace}(\mathbf{WS}) \text{ where } \mathbf{W}=\mathbf{BJ},\ \mathbf{S}=\mathbf{QH'}\psi^{-\mathbf{1}}\mathbf{H} \tag{B5}$$

Where $\mathbf{S}$ and $\mathbf{W}$ are both of dimension $p\times p$. $\mathbf{S}$ is a positive semidefinite matrix since both $\mathbf{Q}$ and $\mathbf{H'}\psi^{-\mathbf{1}}\mathbf{H}$ are positive semidefinite. For $\mathbf{W}^{p\times p}$ and $\mathbf{S}^{p\times p}$ positive semidefinite, trace of their product can be bounded by the following quantities (see Kleinman and Athans, 1968 and discussion in Fang et al., 1994):

$$\lambda_{min}(\mathbf{W})\mathbf{trace}(\mathbf{S})\leq\mathbf{trace}(\mathbf{WS})\leq\lambda_{\mathbf{min}}(\mathbf{W})\mathbf{trace}(\mathbf{S}) \tag{B6}$$

By Property 4 of the weight matrix $\mathbf{B}$, we know that $\lambda_{min}(\mathbf{W})=\mathbf{0}$ and $\lambda_{max}(\mathbf{W})=\mathbf{1}$, hence the above reduces to $0\leq trace(\mathbf{WS})\leq\mathbf{1}\cdot\mathbf{trace}(\mathbf{S})$. Hence is the proof by B5.

https://doi.org/10.5194/gmd-2022-89
Preprint. Discussion started: 23 May 2022


Geoscientific Model Development Discussions

Author contributions. V.Y., and S.G. contributed equally in preparing the manuscript.

Competing interests. The authors declare no competing interest.

Acknowledgements. The authors thank Anna Karion, Kimberly Mueller, James Whetstone (National Institute of Standards and technology, NIST), and Daniel Cusworth (University of Arizona, UA) for their review and advice on the manuscript. This work was partially funded by NIST's Greenhouse Gas Measurements Program. Support to University of Notre Dame provided by NIST grant 70NANB19H132. Support for JPL was provided via an interagency agreement between NIST and NASA. A portion of this research was carried out at JPL, California Institute of Technology, under a contract with NASA (80NM0018D0004).

https://doi.org/10.5194/gmd-2022-89
Preprint. Discussion started: 23 May 2022

References

Berk, R., Brown, L., Buja, A., Zhang, K., and Zhao, L.: Valid post-selection inference, The Annals of Statistics, pp. 802–837, 2013.

Bouchard, M., Jousselme, A.-L., and Doré, P.-E.: A proof for the positive definiteness of the Jaccard index matrix, International Journal of Approximate Reasoning, 54, 615–626, 2013.

Brasseur, G. P. and Jacob, D. J.: Modeling of atmospheric chemistry, Cambridge University Press, 2017.

Cha, S.-H.: Comprehensive survey on distance/similarity measures between probability density functions, City, 1, 1, 2007.

Conley, S., Franco, G., Faloona, I., Blake, D. R., Peischl, J., and Ryerson, T.: Methane emissions from the 2015 Aliso Canyon blowout in Los Angeles, CA, Science, 351, 1317–1320, 2016.

Enting, I. G.: Inverse problems in atmospheric constituent transport, Cambridge University Press, 2002.

Fang, Y., Loparo, K. A., and Feng, X.: Inequalities for the trace of matrix product, IEEE Transactions on Automatic Control, 39, 2489–2490, 1994.

Gelman, A. and Hill, J.: Data analysis using regression and multilevel/hierarchical models, Cambridge university press, 2006.

Ghosh, S., Mueller, K., Prasad, K., and Whetstone, J.: Accounting for transport error in inversions: An urban synthetic data experiment, Earth and Space Science, 8, e2020EA001 272, 2021.

Gourdji, S., Hirsch, A., Mueller, K., Yadav, V., Andrews, A., and Michalak, A.: Regional-scale geostatistical inverse modeling of North American CO 2 fluxes: a synthetic data study, Atmospheric Chemistry and Physics, 10, 6151–6167, 2010.

Hastie, T., Tibshirani, R., and Wainwright, M.: Statistical learning with sparsity, Monographs on statistics and applied probability, 143, 143, 2015.

Johnson, J. W.: A heuristic method for estimating the relative weight of predictor variables in multiple regression, Multivariate behavioral research, 35, 1–19, 2000.

Jolliffe, I. T. and Cadima, J.: Principal component analysis: a review and recent developments, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 374, 20150 202, 2016.

Kitanidis, P. K.: On the geostatistical approach to the inverse problem, Advances in Water Resources, 19, 333–342, 1996.

Kleinman, D. and Athans, M.: The design of suboptimal linear time-varying systems, IEEE Transactions on Automatic Control, 13, 150–159, 1968.

Lauvaux, T., Miles, N. L., Deng, A., Richardson, S. J., Cambaliza, M. O., Davis, K. J., Gaudet, B., Gurney, K. R., Huang, J., O'Keefe, D., et al.: High-resolution atmospheric inversion of urban CO2 emissions during the dormant season of the Indianapolis Flux Experiment (INFLUX), Journal of Geophysical Research: Atmospheres, 121, 5213–5236, 2016.

Lin, J., Gerbig, C., Wofsy, S., Andrews, A., Daube, B., Davis, K., and Grainger, C.: A near-field tool for simulating the upstream influence of atmospheric observations: The Stochastic Time-Inverted Lagrangian Transport (STILT) model, Journal of Geophysical Research: Atmospheres, 108, 2003.

Link, W. A. and Doherty Jr, P. F.: Scaling in sensitivity analysis, Ecology, 83, 3299–3305, 2002.

MacKay, D. J., Mac Kay, D. J., et al.: Information theory, inference and learning algorithms, Cambridge university press, 2003.

MatlabLivescript: version 9.9.0 (R2020b), The MathWorks Inc., Natick, Massachusetts, https://www.mathworks.com/help/matlab/matlab_prog/what-is-a-live-script-or-function.html.

Michalak, A. M., Bruhwiler, L., and Tans, P. P.: A geostatistical approach to surface flux estimation of atmospheric trace gases, Journal of Geophysical Research: Atmospheres, 109, 2004.

https://doi.org/10.5194/gmd-2022-89
Preprint. Discussion started: 23 May 2022

Michalak, A. M., Randazzo, N. A., and Chevallier, F.: Diagnostic methods for atmospheric inversions of long-lived greenhouse gases, Atmospheric Chemistry and Physics, 17, 7405–7421, 2017.

Nielsen, F.: On the Jensen–Shannon symmetrization of distances relying on abstract means, Entropy, 21, 485, 2019.

Rabitz, H.: Systems analysis at the molecular scale, Science, 246, 221–226, 1989.

Rodgers, C. D.: Inverse methods for atmospheric sounding: theory and practice, vol. 2, World scientific, 2000.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., and Tarantola, S.: Global sensitivity analysis: the primer, John Wiley & Sons, 2008.

Tarantola, A.: Inverse problem theory and methods for model parameter estimation, SIAM, 2005.

Turányi, T.: Sensitivity analysis of complex kinetic systems. Tools and applications, Journal of mathematical chemistry, 5, 203–248, 1990.

Vafaei, N., Ribeiro, R. A., and Camarinha-Matos, L. M.: Selecting normalization techniques for the analytical hierarchy process, in: Doctoral Conference on Computing, Electrical and Industrial Systems, pp. 43–52, Springer, 2020.

Wikle, C. K. and Berliner, L. M.: A Bayesian tutorial for data assimilation, Physica D: Nonlinear Phenomena, 230, 1–16, 2007.

Yadav, V., Duren, R., Mueller, K., Verhulst, K. R., Nehrkorn, T., Kim, J., Weiss, R. F., Keeling, R., Sander, S., Fischer, M. L., et al.: Spatio-temporally resolved methane fluxes from the Los Angeles Megacity, Journal of Geophysical Research: Atmospheres, 124, 5131–5148, 2019.