



High-resolution greenhouse gas flux inversions using a machine learning surrogate model for atmospheric transport

Nikhil Dadheeck^{1,★}, Tai-Long He^{1,a,★}, and Alexander J. Turner¹

¹Department of Atmospheric and Climate Science, University of Washington, Seattle, WA, USA

^anow at: School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138, USA

★These authors contributed equally to this work.

Correspondence: Alexander J. Turner (turneraj@uw.edu)

Received: 20 September 2024 – Discussion started: 26 September 2024

Revised: 7 March 2025 – Accepted: 9 March 2025 – Published: 21 May 2025

Abstract. Quantifying greenhouse gas (GHG) emissions is critically important for projecting future climate and assessing the impact of environmental policy. Estimating GHG emissions using atmospheric observations is typically done using source–receptor relationships (i.e., “footprints”). Constructing these footprints can be computationally expensive and is rapidly becoming a computational bottleneck for studying GHG fluxes at high spatio-temporal resolution using dense observations. Here, we demonstrate a computationally efficient GHG flux inversion framework using a machine learning emulator for atmospheric transport (FootNet) as a surrogate for the full-physics model. The footprints generated by FootNet are at approximately 1 km resolution. We update the architecture of the deep-learning model to improve the performance in a GHG flux inversion. We find that the posterior fluxes estimated with FootNet footprints are in good agreement with the posterior fluxes estimated with STILT footprints. We observe that the more simplistic representation of transport in the machine learning model helps to mitigate transport errors. This flux inversion using a machine learning surrogate model requires only meteorological data, GHG measurements, and prior fluxes. Constructing footprints using FootNet is 650 times faster than the full-physics atmospheric transport model on similar hardware. This speedup allows for the computation of footprints “on the fly” during the GHG flux inversion (i.e., computed as needed, rather than archiving for future use) and makes near-real-time emission monitoring computationally possible. This work alleviates a major computational bottleneck with inferring GHG fluxes with next-generation dense observing systems.

1 Introduction

Carbon dioxide (CO₂) and methane are the two most powerful greenhouse gases (GHGs). Together, they account for more than 85 % of the total GHG radiative forcing since preindustrial times (IPCC, 2023). As such, it is important to quantify the GHG sources and sinks in order to project future climate. Near-real-time quantification of GHG emissions is key to identifying the intermittent super-emitters, which often dominate the emission budget. However, the large computational and storage costs associated with full-physics at-

mospheric transport models in the current inversion framework limit our ability to perform near-real-time emissions monitoring from urban to global scales (Roten et al., 2021; Varon et al., 2023; Cartwright et al., 2023; Fillola et al., 2023; Nayagam et al., 2023; Steiner et al., 2024; Janardanan et al., 2024). Here, we use FootNet (He et al., 2025), a computationally efficient deep-learning model, to emulate a full-physics atmospheric transport model and conduct GHG flux inversions. This work shows the feasibility of using a machine learning (ML) emulator for near-real-time computation of source–receptor relationships and to infer hourly GHG

emission fluxes at the kilometer scale from atmospheric observations.

Previous work has shown the importance of point sources for methane emissions (e.g., Brandt et al., 2014; Zavala-Araiza et al., 2015; Frankenberg et al., 2016; Duren et al., 2019; Lauvaux et al., 2022; Chen et al., 2022; Cusworth et al., 2022; Sherwin et al., 2023; He et al., 2024) and urban and localized sources for CO₂ (e.g., Hutyra et al., 2014; Janardanan et al., 2016; Turner et al., 2020; Wu et al., 2020; Kiel et al., 2021). These sources represent a small geographical area and yet dominate the GHG emission budget. The emission fluxes from these “super emitters” are often observed to have a heavy-tailed distribution (Brandt et al., 2014; Zavala-Araiza et al., 2015; Frankenberg et al., 2016; Duren et al., 2019; Chen et al., 2022; Cusworth et al., 2022). In other words, a small number of point sources are responsible for a large fraction of the total emission budget, despite representing a small fraction of the land area. As such, studying these point sources requires densely spaced measurements due to their localized nature. Fortunately, there has been a proliferation of next-generation, dense observing systems for GHGs over the past decade, including both spaceborne instruments, e.g., OCO-2 (O’Dell et al., 2012; Crisp et al., 2012; Eldering et al., 2012; Hammerling et al., 2012), OCO-3 (Eldering et al., 2019; Taylor et al., 2020), and TROPOMI (Veefkind et al., 2012), and low-cost urban monitoring networks (e.g., BEACO₂N: Shusterman et al., 2016).

Measurements from dense observing systems can be used to quantify surface fluxes. The inverse modeling of GHG emissions often requires one to know the upwind region of influence on the atmospheric measurements. This region of influence is also known as the “source–receptor relationship” or the measurement “footprint” (see Rodgers, 2000). The *i*th observation (y_i) can be related to the *m* surface fluxes using the associated footprint:

$$y_i = \mathbf{h}_i \mathbf{x} + b_i, \quad (1)$$

where \mathbf{h}_i is the $1 \times m$ footprint for the *i*th observation, \mathbf{x} is an $m \times 1$ vector of surface fluxes, and b_i is the background concentration for the observation. Here, \mathbf{h}_i describes the sensitivity of observation y_i to *m* surface fluxes with units similar to ppm($\mu\text{mol m}^{-2} \text{s}^{-1}$)⁻¹. Similarly, the vector of the *n* observations ($\mathbf{y}; n \times 1$) can be related to the surface fluxes as

$$\mathbf{y} = \mathbf{H} \mathbf{x} + \mathbf{b}, \quad (2)$$

where \mathbf{b} is $n \times 1$ vector of background concentrations, and \mathbf{H} is an $n \times m$ Jacobian matrix representing the atmospheric transport, such that the *i*th row of \mathbf{H} describes the sensitivity of the *i*th observation (y_i) to the *m* surface fluxes.

Researchers construct this source–receptor relationship (\mathbf{H}) by running a Lagrangian model *n* times or an Eulerian model *m* times. The choice of construction of \mathbf{H} will depend on the size of both *n* and *m*. There are approximately 800 observations and 15 million state vector elements in

each inversion run for this study. Therefore, constructing \mathbf{H} will require approximately 800 simulations in the Lagrangian framework or around 15 million simulations in the Eulerian framework. The Eulerian models are not grid agnostic, and the corresponding computational expense increases as the spatial resolution increases (Steiner et al., 2024). For example, the Integrated Methane Inversion (IMI) is an Eulerian-based framework focusing on the regional scale but is limited to 25 km at present (Varon et al., 2022). Variational methods such as 4D-Var can be used with large state and observation spaces. However, it requires computing an adjoint, which is a computationally expensive process. Additionally, this process iteratively minimizes the cost function with many forward runs and, as such, can not be parallelized. The computation cost of 4D-Var is independent of the number of observations but can still be very large. It also requires storing many checkpoint files, which can become very large for high spatial resolution and can have high storage costs. Gaussian plume models are known for their simplicity and are often used for point source modeling (Bovensmann et al., 2010; Nassar et al., 2017; Wang et al., 2020). However, these models typically assume favorable conditions such as constant winds and flat topography, which may not always be the case.

Here, we focus on constructing the source–receptor relationship using a full-physics Lagrangian particle dispersion model (LPDM). As such, this requires constructing *n* footprints, i.e., one footprint for each measurement.

The LPDM can be used to construct a footprint by advecting an ensemble of particles backwards in time from the measurement sites using archived meteorology (e.g., Lin et al., 2003). These Lagrangian trajectories are agnostic to the choice of grid and can be easily mapped to a high spatial resolution. Additionally, for this study, *n* is small as compared to *m*, and, therefore, it is computationally efficient to use the Lagrangian method of constructing \mathbf{H} .

The computation of footprints from a full-physics LPDM is an embarrassingly parallel problem and, as such, can easily be parallelized. However, this does not overcome the sheer volume of data that would need to be generated for a high-resolution GHG flux inversion. The construction of footprints can quickly become computationally intractable as the number of measurements increases, as is the case with dense observing systems.

In addition to the computational cost of constructing footprints described above, there is a storage cost associated with these footprints that increases with the number of measurements. This can also become burdensome as the number of observations increases. As an example, Turner et al. (2018) examined point sources in the Barnett Shale region in Texas and found that generating 1 week of footprints for hourly measurements made by a geostationary satellite instrument required more than 15 million simulations to construct the footprints and over 4 TB to store them. This region represents less than 1 % of the United States. As such, many previous studies investigating point sources have focused on a

subset of cases with favorable atmospheric conditions, allowing them to utilize a simplified representation of atmospheric transport. For example, recent work estimated CO₂ emissions from individual power plants using a Gaussian plume (e.g., Nassar et al., 2017, 2021; Guo et al., 2023), which is only valid for steady winds.

Recent work have explored the potential of machine learning and other analytical methods to address computational bottlenecks when working with LPDMs. For example, Roten et al. (2021) developed an interpolation method using non-linear weighted averaging to compute footprints. Cartwright et al. (2023) developed a convolution-based variational autoencoder, which predicts footprints using a spatio-temporal Gaussian emulator. Brecht et al. (2023) used neural networks for super-resolution to improve trajectory calculations with the FLEXible PARTicle dispersion model (FLEXPART). Fillola et al. (2023) developed an emulator using gradient-boosted regression trees to predict the influence for each grid, one at a time. These studies provide a proof-of-concept for emulating full-physics LPDM. However, they still require running the full-physics LPDM simulations a significant number of times to conduct flux inversions. Fillola et al. (2023) developed a stand-alone emulator that can predict the near field of the footprints at approximately 35 km × 23 km resolution. However, they had to use LPDM simulations for the far field while using the emulator in an inversion framework. Hence, these proof-of-concept studies are not entirely independent of LPDM simulations in an inversion framework. Additionally, their spatial resolutions are very coarse, which is not ideal for high-resolution emission inventories.

Here we use FootNet, a computationally efficient deep-learning model that emulates the atmospheric transport (He et al., 2025), to conduct GHG flux inversion at the kilometer scale. FootNet is a deep-learning model based on a U-Net architecture and trained on outputs from the Stochastic Time-Inverted Lagrangian Transport model (STILT; Lin et al., 2003). FootNet computes footprints at 1 km × 1 km resolution and is independent of the parent LPDM after the training process. As such, FootNet can be readily used in GHG flux inversions and, once trained, does not require STILT model outputs to predict the footprint. Here, we evaluate the performance of FootNet in GHG flux inversion using the San Francisco Bay Area in northern California as a case study.

2 Impacts of COVID-19 regulations on urban CO₂ fluxes in the San Francisco Bay Area as a case study

This study focuses on the San Francisco Bay Area in northern California. We use hourly atmospheric CO₂ measurements from the Berkeley Environmental Air Quality and CO₂ Network (BEACO₂N; Shusterman et al., 2016; Turner et al., 2016; Shusterman et al., 2018; Turner et al., 2020; Asimow et al., 2024). BEACO₂N is a dense urban moni-

toring network with monitoring sites spaced approximately 2 km apart. This study uses hourly CO₂ measurements for the period of 2 February 2020 to 2 May 2020. Measurements are from approximately 35 BEACO₂N sites. These observations were used in a recent study from Turner et al. (2020), who evaluated the impact of COVID-19 restrictions on urban CO₂ fluxes. This high-resolution GHG flux inversion from Turner et al. (2020) will serve as a reference study to evaluate the performance of flux inversions using an ML model as a surrogate for the full atmospheric transport model.

As in Turner et al. (2020) and He et al. (2025), we use meteorological fields from the High-Resolution Rapid Refresh (HRRR) model at 3 km × 3 km spatial resolution to drive the STILT LPDM. Footprints from the STILT LPDM are then used as training data for FootNet (see He et al., 2025, for details). The training data are sampled from 2018 to 2019, as such, the timeline of this case study is not involved in the training data. FootNet predicts footprints at 1 km × 1 km resolution over a 400 km × 400 km spatial region centered on the San Francisco Bay Area. The setup for the CO₂ flux inversion is the same as in Turner et al. (2020). As such, we can compare the performance of the GHG flux inversion from FootNet with previously published work. Differences from the setup in Turner et al. (2020) will be emphasized in the text that follows.

3 Inferring CO₂ fluxes at high spatio-temporal resolution from atmospheric observations

Our goal in this work is to infer the fluxes of CO₂ using atmospheric observations. Bayesian inference is commonly used when estimating CO₂ fluxes using atmospheric observations. Building on the framework described in Sect. 1, we use Bayesian inference to relate the probability density function (PDF) of the posterior ($P(x|y)$) to the observation likelihood PDF ($P(y|x)$) and prior PDF ($P(x)$) as follows:

$$P(x|y) \propto P(y|x)P(x). \quad (3)$$

Assuming a normal distribution for both $P(y|x)$ and $P(x)$ yields a closed-form solution for the posterior distribution:

$$P(x|y) \propto \exp \left[-\frac{1}{2}(y - \mathbf{Hx})^T \mathbf{R}^{-1} (y - \mathbf{Hx}) - \frac{1}{2}(\mathbf{x} - \mathbf{x}_a)^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}_a) \right], \quad (4)$$

where \mathbf{x}_a is the prior, \mathbf{R} is the observational error covariance matrix, and \mathbf{B} is the prior error covariance matrix. The maximum a posteriori probability can be obtained by finding the minimum of the negative log-likelihood term within the ex-

ponential. The resultant cost function is

$$\begin{aligned} \mathcal{J}(x) = & \frac{1}{2}(y - \mathbf{H}x)^T \mathbf{R}^{-1} (y - \mathbf{H}x) \\ & + \frac{1}{2}(x - x_a)^T \mathbf{B}^{-1} (x - x_a). \end{aligned} \quad (5)$$

Minimizing the cost function provides a closed-form estimate for the posterior fluxes (\hat{x}):

$$\hat{x} = x_a + (\mathbf{H}\mathbf{B})^T (\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1} (y - \mathbf{H}x_a), \quad (6)$$

where \hat{x} is the posterior fluxes. It is important to note here that \mathbf{B} is an $m \times m$ matrix that is often computationally intractable for large state vectors. In this case study, m is larger than 15 million, and \mathbf{B} is computationally intractable. We use a Kroenecker product to decompose \mathbf{B} into temporal and spatial submatrices. Yadav and Michalak (2013) proposed a computationally efficient algorithm for serial computation of \mathbf{HB} and \mathbf{HBH}^T using this Kroenecker product without explicitly forming \mathbf{B} . We further reduce the computation time of \mathbf{HB} computation by updating their algorithm to implement parallel computation (see Appendix A).

4 Relating observations to surface fluxes using footprints

This study uses footprints generated by both STILT and FootNet to relate surface fluxes to observations. Both of these models are driven by the same parent HRRR meteorology. STILT is a Lagrangian model built on top of the Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT) model and can produce either time-resolved (e.g., hourly) footprints or time-integrated footprints (i.e., the temporal dimension has been summed, yielding a 2-D matrix). STILT footprints generated for this study are hourly, going 72 h backwards in time or until the trajectories leave the spatial domain (whichever is shorter). This is the same setup that was used in Turner et al. (2020). The FootNet model computes time-integrated footprints because it was deemed computationally infeasible to emulate the time-resolved footprints. Therefore, we need to devise an approach to allocate the FootNet footprints backwards in time. Here, we use exponentially decaying weights to allocate the footprint across all 72 back hours. These weights are normalized such that the summation of the weights adds up to 1 and conserves the total magnitude of the footprint. This section compares the temporal allocation using exponential decay on time-integrated STILT footprints with the hourly time-resolved STILT footprints in a GHG flux inversion to investigate the additional error induced before using it with FootNet footprints in Sect. 5.

Figure 1 shows the difference between hourly time-resolved footprints and the time-integrated footprints with exponentially decayed weights. The time-resolved footprints have a smaller sensitivity “blob” that moves as we go backward in time. This indicates that, for any given time step, only

the emission sources in the small shaded region influence our observation. This highly time-resolved method assumes that the numerical schemes used for the transport and advection are highly accurate. In contrast, the time-integrated footprint with exponentially decayed weights assumes a time-invariant spatial structure with decreasing magnitude at previous time steps. This plume decays with time such that time steps close to the time of observation have higher weights as compared to the plume 72 h before the time of observation, which has negligible influence. This temporal allocation of the footprint will likely induce additional error.

We assess this error using a pair of flux inversions: (1) conduct a GHG flux inversion with time-resolved STILT footprints, and (2) conduct a second GHG flux inversion using STILT footprints where the footprints have been time-integrated and temporally allocated as described above. This pair of GHG flux inversions will isolate the impact of this temporal allocation. Please note that all other inversion parameters are the same between these tests.

Figure 2 shows the results of these two GHG flux inversions in the Bay Area. It is important to emphasize that both of the inversions use the same set of footprints derived from the STILT model. The only difference is that one set of STILT footprints has been time-integrated and then reallocated temporally using exponential decay. The scatter plots compare observed CO₂ concentrations with the simulated CO₂ concentrations using prior and posterior fluxes on a validation dataset that was not used in the flux inversion. We use the same seed such that same validation concentrations are sampled in both cases. Overall, the posterior emission fluxes are in agreement. Interestingly, the second case, where the footprints have been temporally reallocated, does a better job at simulating independent observations with both the prior fluxes and the posterior fluxes. This can be seen in both the correlation and the mean squared error. Upon close examination of the scatter plots, it can be seen that the time-resolved scatter plots have a cluster of simulated concentrations around 410 ppm (approximately background signal), even though actual concentrations are higher than that. This pattern is visible in the CO₂ concentrations simulated using both prior and posterior fluxes. On the other hand, this pattern is not visible in the CO₂ concentrations simulated using the posterior fluxes of the exponential decay case.

The poor performance of the time-resolved footprints is likely driven by transport errors in the STILT simulations. The time-resolved footprints are a more realistic representation of the source–receptor relationship, but not necessarily more accurate. The simulated transport could have large biases that will propagate into time-resolved footprints. For example, errors in the wind speed could lead to parcels advecting too fast and attributing fluxes to the wrong spatial location. In these flux inversions, the ocean and the San Francisco Bay are assumed to be neither sources nor sinks. It seems that the sensitivity blobs in the time-resolved footprints are quickly advected over the ocean, and hence the

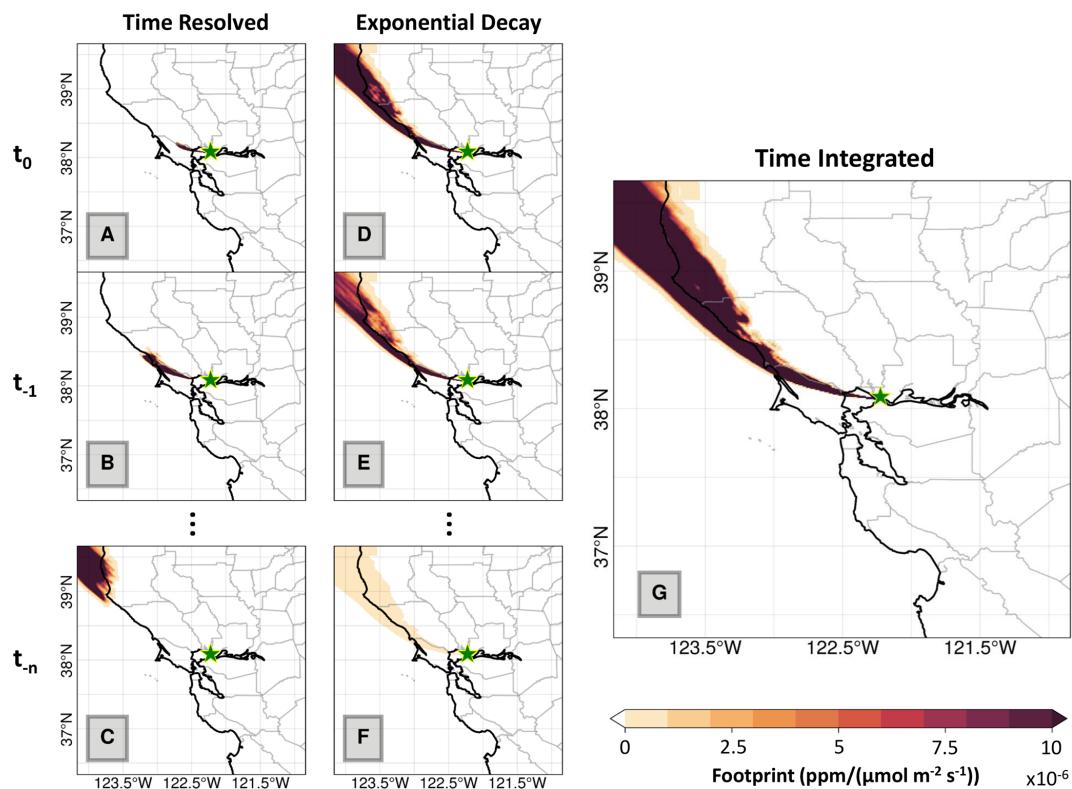


Figure 1. Two methods of representing temporal patterns in footprints for a single receptor. The green stars indicate the receptor location for the footprint. Panels (a) – (c) show the time-resolved footprints computed directly from STILT. Panels (d) – (f) show the time-integrated STILT footprints with the pattern allocated temporally using an exponential decay. Both representations have the same time-integrated spatial pattern (panel g).

simulated concentrations are closer to the background signal, even though the observed concentrations are higher. As such, the time-integrated footprints may be mitigating these transport errors and allowing GHG fluxes to be attributed correctly. Additionally, this pair of GHG flux inversions indicates that the exponential-decay-based temporal allocation of footprints should not induce significant errors in the GHG flux inversion. All of the GHG flux inversion experiments that follow (i.e., Sect. 5) will use FootNet footprints and this temporal allocation of the footprints.

5 GHG flux inversions with a machine learning surrogate model

Using the temporal allocation strategy described above, we can now assess the performance of an ML-based surrogate transport model within a GHG flux inversion. Figure 3 shows conceptually how this ML model for atmospheric transport can be used within a GHG flux inversion. Figure 3a shows the conventional GHG inversion framework using an LPDM. This framework begins with obtaining meteorological data for the spatio-temporal region of interest. These meteorological data are used to drive the full-physics LPDM and, in turn, construct the footprints for all of the observations. Construc-

tion of these footprints is computationally expensive, and the footprints are typically archived prior to conducting the GHG flux inversion. Archiving these footprints can have large data storage requirements for dense observations at high spatio-temporal resolution. Following the construction and archival of the footprints, researchers can then estimate GHG fluxes via Bayesian inference, as described in Sect. 3.

Figure 3b shows the process to estimate GHG fluxes using an ML model as a surrogate for the full-physics atmospheric transport model. The initial and final steps of the process are identical to the process described above. The difference is in the construction and archival of the footprints. We detail two approaches: (1) construct and archive the footprints, or (2) construct the footprints “on the fly” during the GHG flux inversion. The former approach is similar to the process using the full-physics model, but the construction of the footprints uses the ML-based surrogate model (e.g., FootNet). The latter approach of computing the footprints on the fly during the inversion is only feasible if the computation of the footprints can be done in near-real time; otherwise, additional computational expense would make the inversion prohibitively slow. The computation of a time-integrated footprint by the ML-based surrogate model described by He et al. (2025) takes less than a second and, as such, may be sufficiently fast to

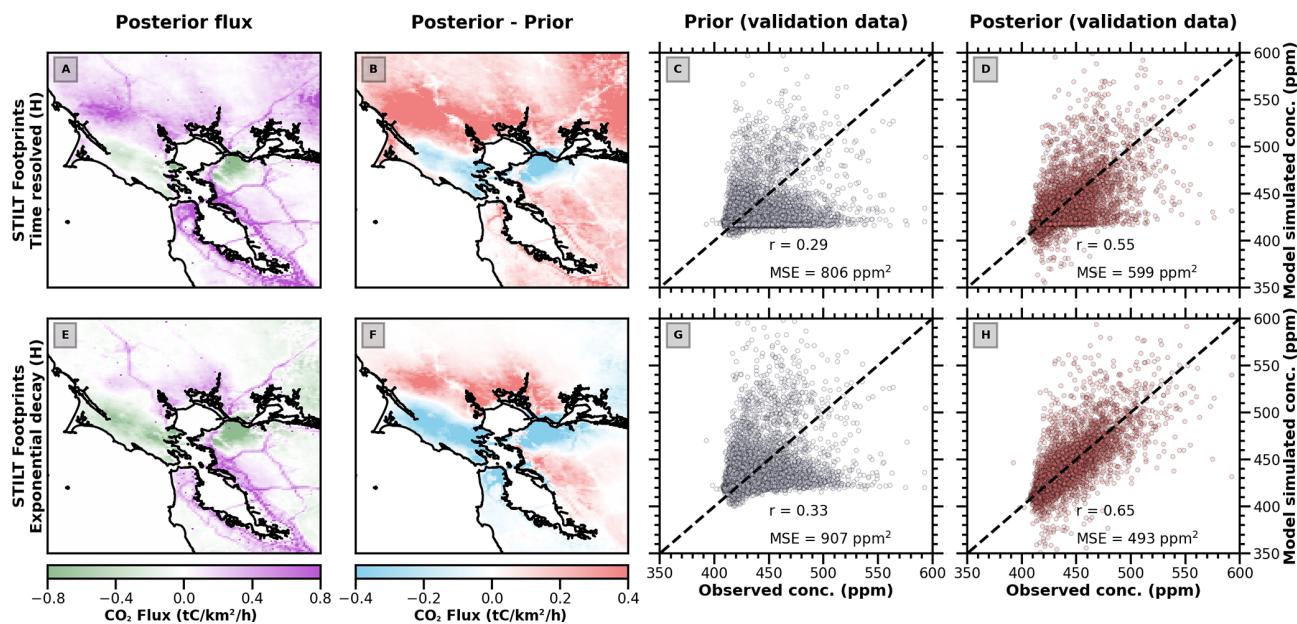


Figure 2. Comparison between CO₂ fluxes derived using time-resolved and time-integrated footprints with exponential decay for the STILT model. Panels (a)–(d) show flux inversion results using time-resolved footprints from STILT. Panels (e)–(h) show flux inversion results using time-integrated footprints with an exponential decay. Panels (a) and (e) show the posterior fluxes averaged over the study period from 2 February 2020 to 2 May 2020. Panels (b) and (f) show the difference between the posterior and prior fluxes. Panels (c) and (g) show a comparison of the model simulated concentrations using the prior fluxes against independent observations withheld from the flux inversion. Panels (d) and (h) show the same comparison as is panels (c) and (g), but using the posterior fluxes.

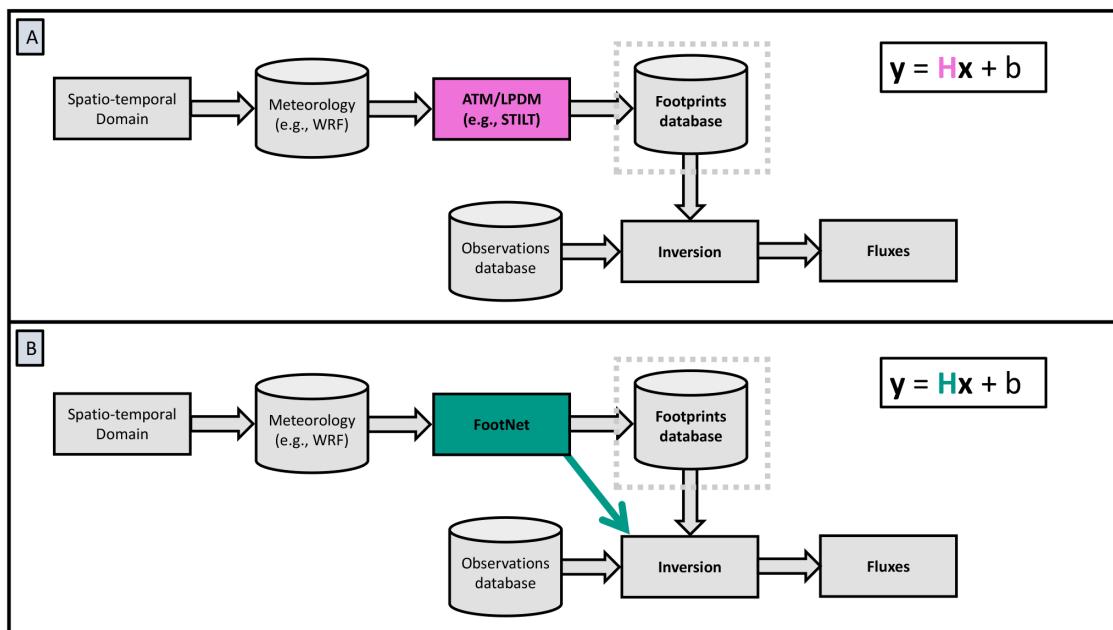


Figure 3. Conventional and FootNet-based GHG inversion frameworks. Panel (a) shows the conventional GHG flux inversion framework, which uses a full-physics LPDM simulation to compute footprints. Panel (b) shows the proposed GHG flux inversion framework using FootNet model to compute footprints. The arrow in panel (b) indicates how FootNet may be used to compute footprints on the fly as they are needed within the GHG flux inversion, bypassing the data storage step.

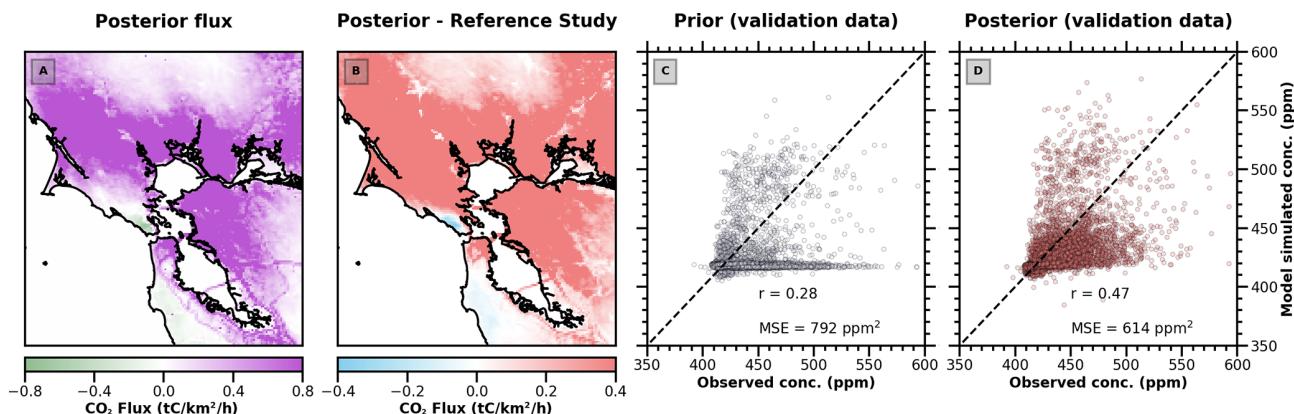


Figure 4. Urban CO₂ flux inversion results in the San Francisco Bay Area using FootNet v1 (He et al., 2025). Panel (a) shows the posterior CO₂ fluxes averaged over the study period. Panel (b) shows the difference between the posterior fluxes inferred with FootNet v1 and fluxes inferred with a full-physics model (STILT). Panel (c) shows a comparison of CO₂ concentrations simulated with FootNet v1 and the prior fluxes against independent observations. Panel (d) shows the same as in panel (c) but for posterior fluxes inferred from using FootNet v1. The reference study is Turner et al. (2020).

facilitate computation of footprints on the fly as they are needed in the GHG flux inversion. In the following sections, we evaluate the performance and the computational expense of GHG flux inversion frameworks using FootNet. These results will be compared to a GHG flux inversion using a full-physics atmospheric transport model (STILT).

5.1 GHG flux inversion with FootNet v1

Figure 4 shows the results of a flux inversion using FootNet v1, the model described in He et al. (2025), with the flux inversion setup from Turner et al. (2020), who evaluated the impact of COVID-19 regulations on urban CO₂ fluxes in the San Francisco Bay Area. Briefly, we conduct hourly flux inversions at 1 km spatial resolution for overlapping 96 h windows. Each 96 h window includes a 36 h buffer on the 24 h period of interest. The prior error covariance matrix is decomposed using a Kronecker product (e.g., Yadav and Michalak, 2013). Upwind concentrations are taken from NOAA observations in the Pacific and from AmeriFlux measurements in the Sacramento Delta.

From Fig. 4c and d, we observe that the posterior fluxes inferred using FootNet v1 perform substantially better than the prior fluxes when compared against independent validation data. Figure 4a shows changes in physically meaningful locations, such as freeways in the San Francisco Bay Area, as well as regions dominated by the biosphere. However, from Fig. 4b, we can see that these posterior fluxes do not appear to be in agreement with the fluxes inferred from Turner et al. (2020) using the full-physics model. FootNet v1 finds substantially higher fluxes throughout the domain than the inversion using the full-physics model. This suggests that FootNet v1 simulates realistic spatial patterns well but may be generating weaker footprints than STILT and may not be appropriately scaled. Additionally, the large differences in the

far field seen in Fig. 4b indicate that there could be an imbalance between the near field and far field of the footprints from FootNet v1.

Figure 5 shows the cumulative influence of footprints for the full-physics model (STILT) and FootNet v1. These cumulative influence plots give an idea of the spatial regions that the observations are sensitive to. As alluded to above, there are strong similarities in the spatial patterns, but FootNet v1 does indeed find a larger contribution from distant regions than STILT. This larger region of influence would result in FootNet allocating larger fluxes to distant regions than a flux inversion using STILT. This comparison of the regions of influence suggests that the FootNet model should be updated to improve the performance within a GHG flux inversion. We hypothesize two methods for improving the performance of FootNet: (1) changing parameters in the deep-learning architecture for FootNet and (2) adding input features to FootNet.

5.2 Updating the FootNet model to improve performance in flux inversions

As a first test, we train an additional variant of the FootNet model with an alternate formulation of the cost function (i.e., the cost function used to construct FootNet, not the cost function for a GHG flux inversion). This FootNet variant still relies on the underlying U-Net architecture. The FootNet v1 model used a mean-squared error (MSE; i.e., \mathcal{L}_2 -norm) cost function in the gradient descent optimization. We hypothesize that an \mathcal{L}_1 -norm cost function may improve the balance between the near field and the far field. This is because the MSE cost function is sensitive to outliers, in contrast to an \mathcal{L}_1 -norm that is less sensitive to outliers (see Bishop, 2006).

Figure 6a and b show the cumulative influence for FootNet v1 using both an MSE and \mathcal{L}_1 norm cost function. Both formulations of the cost function yield large spatial regions

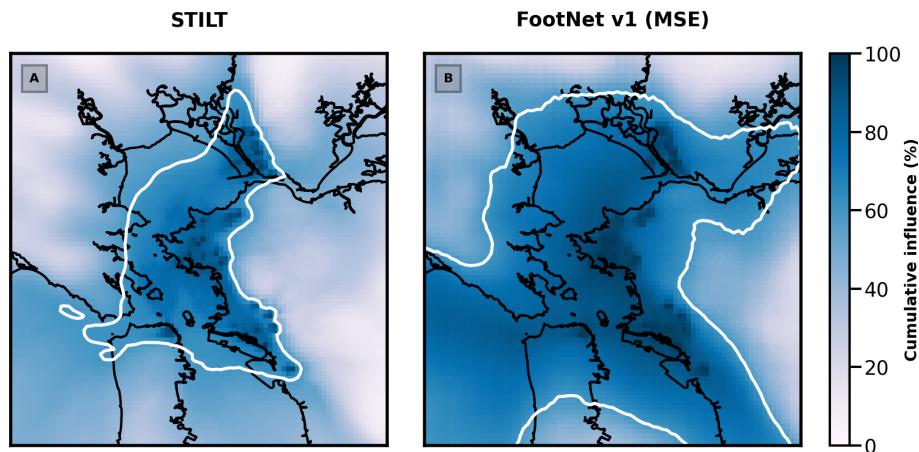


Figure 5. Cumulative region of influence for the BEACO₂N network in the San Francisco Bay Area. Panel (**a**) shows the cumulative influence computed using STILT, a full-physics model. Panel (**b**) shows the cumulative influence computed using FootNet v1 with an MSE cost function. The white contour represents the region encapsulating the top 40 % of the total influence on the BEACO₂N network (i.e., the region to which the observations are most sensitive).

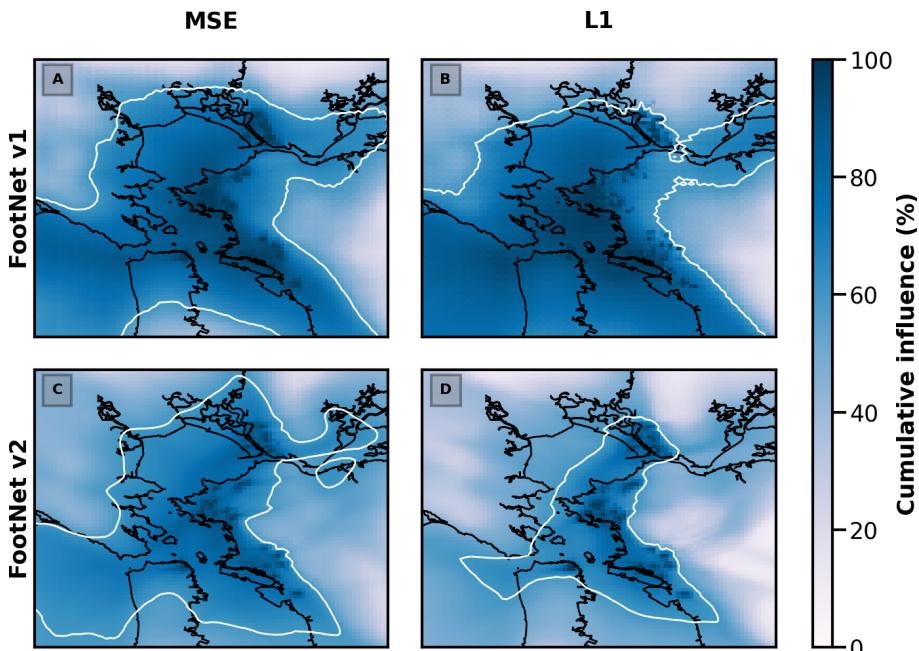


Figure 6. Same as Fig. 5 but for different variants of FootNet. Panels (**a**) and (**b**) show the cumulative influence for FootNet v1. Panels (**c**) and (**d**) show the cumulative influence for FootNet v2. Panels (**a**) and (**c**) use MSE cost functions. Panels (**b**) and (**d**) use \mathcal{L}_1 -norm cost functions.

of influence. Based on this, we conclude that changing the cost function alone is insufficient to rectify the imbalance in the near-field and far-field footprints.

Two other FootNet model parameters we evaluate are the choice of activation function and the formulation of a log-transformation of the training data. The construction of FootNet v1 uses the rectified linear unit (ReLU) activation functions to introduce nonlinearity in the deep-learning model architecture (He et al., 2025). We assess the performance of

parametric rectified linear unit (PReLU) activation functions. These PReLU activation functions have parameters that can be tuned during the training process, giving FootNet additional degrees of freedom. FootNet v1 also uses a logarithmic transformation of the training data to help identify large-scale spatial patterns. We modify the logarithmic transformation to add a small number and ensure positivity, as follows:

$$\Gamma(x^*) = \log(x + \epsilon) - \log(\epsilon), \quad (7)$$

where x is a real number and ϵ is a small value ($\epsilon = 10^{-3}$). Both of these updated parameter choices improved the performance of FootNet but did not fully rectify the near-field and far-field imbalance.

Finally, we trained two more FootNet models using additional input features shown in Table 1. We define these models with additional input features as “FootNet v2”. FootNet v1 only uses meteorological data at the time of measurement and 6 h prior. Here, we add meteorological data 12 h back in time. These meteorological data include 10 m zonal and meridional winds, surface pressure, and planetary boundary layer (PBL) height. Winds are important for the model to learn advection, and surface pressure also acts as a proxy for the region’s topography. We also use PBL height, as it is an important input for computing footprints from the trajectories of the particles in a full-physics LPDM (e.g., STILT; Lin et al., 2003). The choice of 12 h is because many trajectories have not yet left the domain within 6 h, and, as such, meteorological information from 12 h before the observation time may be important in constructing the footprint and assessing fluxes. We also include distance from the receptor (both linear distance and an exponentially decaying distance). Distance from the receptor may help the model learn the optimal decay rate for the spatial pattern of the footprint (i.e., help the imbalance of the near field and far field). Finally, we include a spatial mask inferred from a network of Gaussian plumes. The spatial mask is based on winds at the time of measurement, 6, 12, 18, and 24 h back in time and may help identify the important regions influencing our measurement.

Figure 6 shows the cumulative influence plots for FootNet v2 using both MSE and \mathcal{L}_1 -norm cost functions. The 60th percentile contours show a stronger resemblance to that of the full-physics STILT model (see Fig. 5a). The balance between the near field and far field is more in line with the cumulative influence inferred by the STILT model. As hypothesized, the model with the MSE cost function has a larger region of influence than the model using the \mathcal{L}_1 -norm cost function. This allows the model to optimize the spatial decay structure of the footprints as it moves radially outward from the receptor location. Notably, the MSE-based cost function indicates a larger sensitivity over the ocean as compared to the STILT footprint and the \mathcal{L}_1 -norm. This is likely due to FootNet simulating “smoother” footprints than STILT (STILT exhibits sharp gradients at the edge of the footprint). The FootNet v2 model using an \mathcal{L}_1 -norm cost exhibits a smaller region of influence than the STILT model.

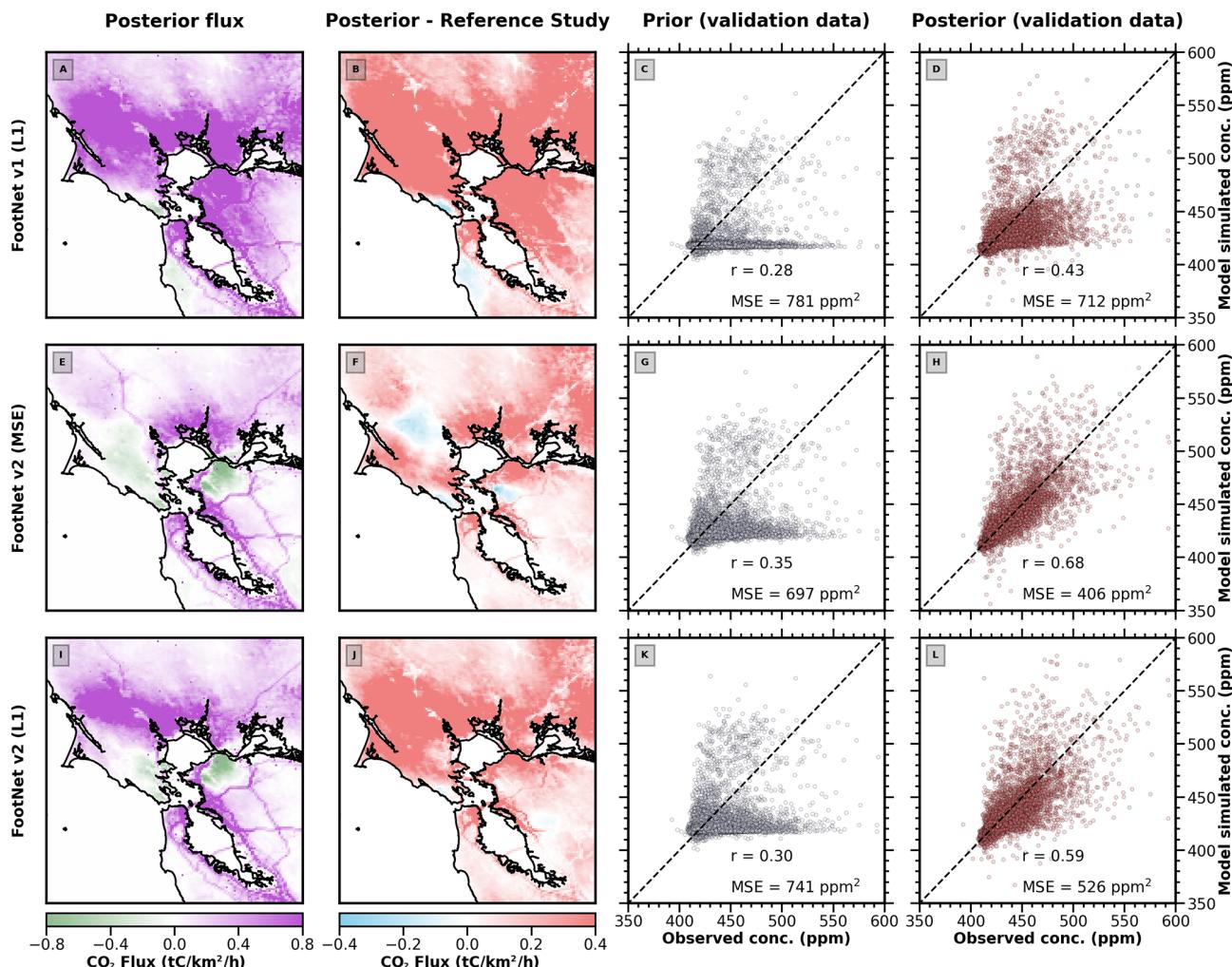
We can now assess the performance of the four models shown in Fig. 6 in the context of realistic GHG flux inversions. In all cases, the models will be compared against the results of GHG flux inversion using the full-physics STILT model (Turner et al., 2020), and they will be evaluated against validation data from CO₂ observations withheld from the flux inversion.

Figure 7 shows the comparison of these models against Turner et al. (2020) and the validation data (note, one of the cases is shown above in Fig. 4). As shown in Fig. 6, the near-field and far-field imbalance persists in FootNet v1 using an \mathcal{L}_1 -norm cost function. Both variants of FootNet v2 show marked improvement in the near-field and far-field balance. Additionally, both variants of FootNet v2 perform substantially better against independent validation data (see right column). Between the two variants of FootNet v2, we find that the mean squared error (MSE)-based cost function performs best. This conclusion is based on the performance against independent validation data and comparison to the posterior fluxes from Turner et al. (2020). The comparison against independent validation shows a strong correlation ($r = 0.68$) and no systematic biases in the residuals. The mean squared error against independent validation data is the lowest of the FootNet models tested (MSE = 405 ppm²). FootNet v2 with an MSE cost function also shows the smallest deviations from the posterior fluxes inferred from Turner et al. (2020), who used a computationally expensive full-physics model to relate the fluxes to observations. As such, we select FootNet v2 with an MSE cost function as the final model.

Figure 8 shows a direct comparison of FootNet v2 using an MSE cost function against the posterior fluxes inferred from Turner et al. (2020) using a full-physics model (STILT). The comparison shows the average fluxes for 6 weeks before COVID-19 shelter-in-place orders, 6 weeks during shelter-in-place orders, the difference, and a comparison against independent validation data. Figure 8a and e show a strong similarity in their CO₂ fluxes before the COVID-19 shelters were in place. We note some disagreements in the far field, such as the northern and eastern parts of the domain. Figure 8b and f correspond to the period during the shelter-in-place measures that decreased anthropogenic emissions. Both sets of posterior fluxes are in agreement during this period, with clearly visible reductions in emissions from freeways and anthropogenic sources. The only notable difference is near Tomales Bay and Point Reyes, in the western portion of the domain. Figure 8c and g show the difference in the CO₂ fluxes between these two periods. Both inversions largely agree in the Bay Area, where the observations have the largest influence. Some disagreements can be seen to the east of Tomales Bay and in the Sacramento Delta. Finally, the right column of Fig. 8 shows the comparison against independent validation data for both flux inversions. Interestingly, we observe the FootNet v2 posterior fluxes to perform *better* than the posterior fluxes inferred using STILT. This is seen in both the correlation coefficient ($r = 0.68$ for FootNet and $r = 0.65$ for STILT) and mean squared error (MSE = 405 ppm² for FootNet and MSE = 493 ppm² for STILT). This begs the question: “why would a machine learning surrogate model perform better than the full-physics model in a GHG flux inversion?”

Table 1. Input features for FootNet.

Features	FootNet v1	FootNet v2
U10M (m s^{-1})	✓	✓
V10M (m s^{-1})	✓	✓
PBL height (m)	✓	✓
Surface pressure (hPa)	✓	✓
Gaussian plume at the time of measurement	✓	✓
Gaussian plume 6 h before measurement	✓	
Meteorology (U10, V10, PBL, Sfc Pres.) at the time of measurement	✓	✓
Meteorology (U10, V10, PBL, Sfc Pres.) 6 h before measurement	✓	✓
Meteorology (U10, V10, PBL, Sfc Pres.) 12 h before measurement		✓
Distance from the receptor (m)		✓
Combined Gaussian plumes network (mask)		✓

**Figure 7.** Same as Fig. 5 but for three other variants of FootNet: FootNet v1 trained with an \mathcal{L}_1 -norm cost function (a–d), FootNet v2 trained with an MSE cost function (e–h), and FootNet v2 trained with an \mathcal{L}_1 -norm cost function (i–l). The reference study is Turner et al. (2020).

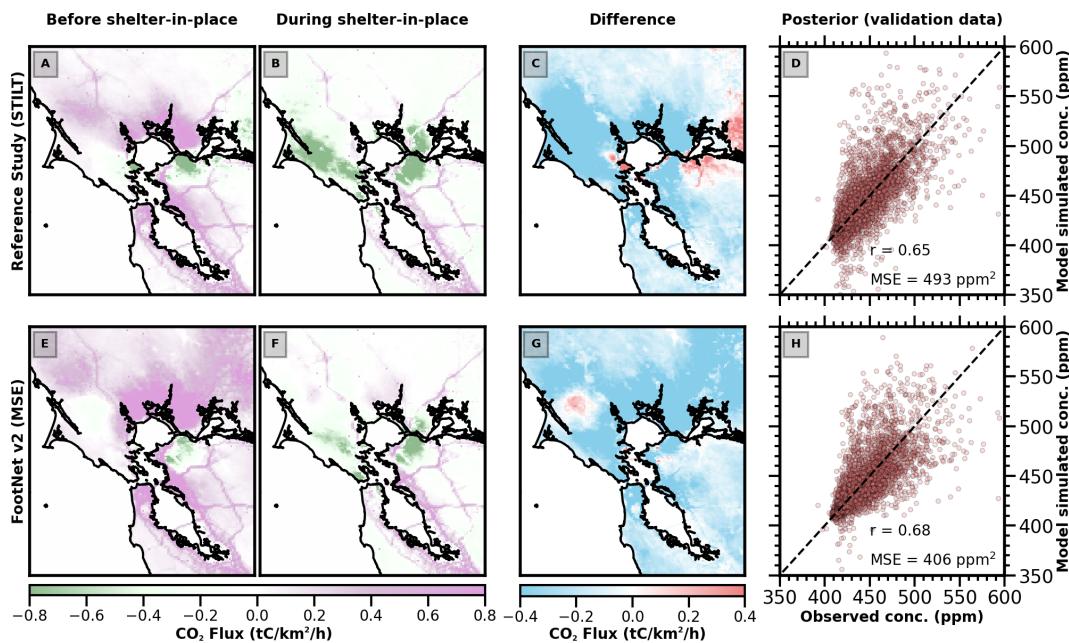


Figure 8. Urban CO₂ fluxes in the San Francisco Bay area inferred using atmospheric observations. Panels (a)–(d) show the posterior CO₂ fluxes inferred using footprints from STILT, a full-physics transport model. Panels (e)–(h) show the posterior CO₂ fluxes inferred using FootNet v2 with an MSE cost function. Panels (a) and (e) show the CO₂ fluxes averaged over 6 weeks before the COVID-19 shelter-in-place orders. Panels (b) and (f) show the CO₂ fluxes averaged over 6 weeks during the COVID-19 shelter-in-place order. Panels (c) and (g) show the difference between these time periods. Panels (d) and (h) compare predicted CO₂ concentrations using the posterior fluxes with independent CO₂ observations withheld from the flux inversion. Note that both inversions use time-integrated footprints that have been temporally allocated based on an exponentially decaying weight.

FootNet was designed to emulate the STILT model (a full-physics atmospheric transport model); as such, it is surprising to see the ML-based surrogate model outperform the full-physics model it was trained on when used in a GHG flux inversion. The explanation for this paradox is that, while STILT is a more realistic representation of the transport it is not necessarily more accurate. ML models often give predictions that tend toward the mean. In the context of atmospheric transport and footprints, this results in a FootNet simulating a smoother and more diffuse spatial pattern than STILT (He et al., 2025). When used in a GHG flux inversion, the sharp gradients simulated with STILT mean that small errors in wind speed or direction could lead to fluxes being allocated to the incorrect spatial location. In the context of the GHG flux inversion, this diffuse spatial pattern simulated by FootNet can potentially mitigate transport errors in the flux inversion. An important takeaway from this work is that using an ML-based surrogate model in a flux inversion can *potentially* outperform the computationally expensive full-physics model. However, this result is unlikely to be universally true, as there will be cases where smoother spatial patterns induce errors in the flux inversion (e.g., when the true trajectories are localized). The performance of the ML-based surrogate model will almost certainly vary on a region-by-region basis, and additional tests are needed to assess the extent of this finding.

6 Computational cost of the GHG flux inversions

Table 2 shows the computational and storage cost analysis for both STILT and FootNet v2 for the 3-month study period. The construction of footprints for all of the measurements using the full-physics-based STILT model is computationally expensive. It takes roughly an hour (upper bound) to construct a single footprint using STILT. As such, it can take more than 8 years to construct all the footprints required for this study if one were to construct them sequentially. Parallel computation of the footprints on a 32-core machine can reduce the time to 3 months. Using multiple nodes can further reduce the time to a few days. However, this reduction in time comes with an infrastructure cost. Given the computational expense in generating these footprints, researchers typically store these time-resolved footprints, which can take approximately 470 GB of space for this study period.

The computation of footprints using FootNet v2 is fast. It takes 3.5 h to compute the same footprints using a single NVIDIA A2 GPU card on 1 core and 24 h using a 32-core machine; this is a 650× and 85× speedup on similar hardware, respectively. This time includes reading input data, constructing footprints, and writing to disk. He et al. (2025) mention that it takes 0.08 s to compute a single footprint after loading input data. This speedy computation from FootNet allows for near-real-time construction of footprints

Table 2. Construction and storage cost to compute footprints with STILT and FootNet^a.

STILT			FootNet		
Serial	Parallel	Parallel (multi-node)	CPU	GPU (archival ^b)	GPU (on the fly ^c)
1 core ^d	32 cores ^d	320 cores ^d	32 cores ^d	1 GPU ^e	1 GPU ^e
Construction ^f	8.4 years	3.15 months	10 d	1.1 d	3.5 h
Storage	467 GB	467 GB	467 GB	45 GB	45 GB
					up to 3 min ^g
					0 GB ^h

^a Computation times may change slightly based on factors such as hardware usage and resource availability. ^b Flux inversion using archived footprints.

^c Footprints computed as they are needed during the flux inversion. ^d Intel® Xeon® Gold 6226R CPU @ 2.90 GHz. ^e NVIDIA A2 GPU, 16 GB GDDR6.

^f Constructing 73 703 footprints. ^g ~ 800 observations for a 1-day inversion with a 96 h window. ^h No archival of footprints in the on-the-fly framework.

within the GHG flux inversion. In other words, one can compute footprints on the fly as they are needed, rather than storing the footprints on the disk. Additionally, it may allow researchers to conduct sensitivity runs on meteorological parameters (e.g., the PBL height) during the flux inversion by including meteorological parameters as a hyperparameter in the inversion. We note that the storage cost is lower for FootNet because STILT is saving time-resolved footprints.

The serial computation of the **HB** matrix, based on the computationally efficient algorithm proposed by Yadav and Michalak (2013), can also become computationally expensive. In this study, we implemented a parallel computation of **HB** matrix based on the algorithm detailed in Yadav and Michalak (2013). This parallel implementation is approximately 27 times faster than the serial implementation. Appendix A discusses the parallel implementation of **HB** matrix computation.

7 Conclusions

Near-real-time quantification of greenhouse gas (GHG) fluxes is important for monitoring, reporting, and verification of GHG fluxes to ensure climate goals are met. Here we demonstrate how machine learning (ML) models can be used as a surrogate for the full-physics atmospheric transport models in a GHG flux inversion, alleviating a computational bottleneck. This work updates the deep-learning architecture of FootNet v1 (He et al., 2025) to improve the performance in a GHG flux inversion. This updated deep-learning model for atmospheric transport (FootNet v2) outperforms the full-physics model in an inversion estimating urban CO₂ fluxes at high spatio-temporal resolution in the San Francisco Bay Area. Further tests are required to investigate the generalizability of this finding.

A potential barrier to using FootNet within a GHG flux inversion is that FootNet computes a 2-D spatial pattern of time-integrated footprints, whereas the full-physics model (STILT) generates time-resolved footprints. To overcome this, we temporally allocate the footprints using an exponentially decaying weight such that there is a time-invariant spatial structure with decreasing magnitude at previous time

steps. Further, we compare predicted concentrations after conducting flux inversion using both the time-resolved and time-integrated footprints with temporal allocation. We observe that time-integrated footprints perform better, as they can mitigate transport errors in the time-resolved representation. The time-resolved footprints are a more realistic representation of the source–receptor relationship, but not necessarily more accurate. Additional tests are needed to understand the transport errors in the time-resolved footprint and the broader applicability of the exponential-decay footprints. This overcomes a potential barrier to using the ML-based surrogate model in a flux inversion.

A preliminary flux inversion using FootNet v1 suggested there was a bias in the balance between the near-field and far-field footprints. We constructed additional variants of the FootNet model and evaluated them in an urban CO₂ flux inversion. Performance was evaluated against independent observations that were withheld from the flux inversion. Ultimately, we developed a new model (FootNet v2) that includes additional input features. We find that FootNet v2 outperforms the full-physics model in the flux inversion in this particular study. This is likely because the FootNet v2 footprints have a smoother spatial structure than the full-physics model. This smoother spatial structure can help mitigate transport errors. Additionally, this machine learning surrogate model allows for a 650× speedup in the construction of the footprints as compared to the full-physics model. This speedup allows for on-the-fly computation of footprints during the inversion, as opposed to archiving footprints prior to the GHG flux inversion.

Previous work has shown that the distribution of GHG sources may be skewed, with a “heavy tail” of super emitters. This suggests that the assumption of a Gaussian distribution for the prior PDFs may not be accurate. Stochastic methods such as Markov chain Monte Carlo can allow one to specify non-Gaussian prior PDFs, as well as jointly solve for meteorology (e.g., uncertainties in PBL height). However, it is currently infeasible to implement with traditional models, as it requires evaluation of the forward model many times, which is computationally intractable. FootNet can compute the footprints in near real time, making it feasible to use

these methods to estimate posterior emissions. This can be one potential application of machine learning surrogates of atmospheric transport in improving the flux estimates.

Overall, FootNet alleviates a computational bottleneck when working with dense GHG observing systems, such as those from urban monitoring networks and next-generation satellite measurements (e.g., MethaneSat and Carbon-I). The computational efficiency of FootNet allows for near-real-time emission monitoring of GHGs, along with other non-reactive trace gases. This work demonstrated the utility of FootNet in quantifying urban CO₂ fluxes in a case study, and future work is needed to extend this framework to a larger region, such as the contiguous US, or total column measurements. Nevertheless, ML-based surrogate models such as FootNet represent a promising direction for efficiently interpreting the growing volume of observational data from next-generation observing systems.

Appendix A: Parallel implementation of HB matrix multiplication

Here, we use a shared-memory parallelization technique to compute \mathbf{HB} from $\mathbf{H}_{n \times pr}$, the temporal prior error covariance matrix ($\mathbf{D}_{p \times q}$), and the spatial prior error covariance matrix ($\mathbf{E}_{r \times t}$). This method is similar to the algorithm described in Yadav and Michalak (2013). The primary difference from Yadav and Michalak (2013) is that we form \mathbf{HB} in shared memory and use a multi-threading approach to iterate over the q columns of \mathbf{D} simultaneously, such that a thread performs the following operation on the k th column of \mathbf{D} :

1. Multiply each $(n \times r)$ block of \mathbf{H} by the elements of the k th column of \mathbf{D} and add these blocks.
2. Multiply the resulting $n \times r$ matrix by $\mathbf{E}_{r \times t}$ to obtain the k th $n \times t$ column block of \mathbf{HB} matrix.
3. Update the k th $n \times t$ column block of \mathbf{HB} matrix.
4. End the thread operation.

This method is limited by the number of threads and the memory available for the matrix multiplication. The larger the number of threads, the faster the multiplication can be, provided that there is enough memory available for each thread.

Code availability. The code for this study is available at <https://github.com/nd349/Bayesian> (last access: 15 May 2025) and <https://doi.org/10.5281/zenodo.13750963> (Dadheeche, 2024).

Data availability. CO₂ data are available at <http://beacon.berkeley.edu/Sites.aspx> (Cohen Research, 2025). The FootNet training data (<https://doi.org/10.5281/zenodo.12803736>, He, 2024) used in this study are the same as in He et al. (2025). The posterior

fluxes are uploaded to <https://doi.org/10.5281/zenodo.13750963> (Dadheeche, 2024).

Author contributions. ND, TLH, and AJT designed the research study. ND and TLH trained the models, conducted the flux inversions, and analyzed the results. ND, TLH, and AJT wrote the manuscript.

Competing interests. The contact author has declared that none of the authors has any competing interests.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors.

Acknowledgements. This work is supported by a NASA FIESST grant (grant no. 80NSSC22K1557) to Nikhil Dadheeche and a NASA Early Career Faculty grant (grant no. 80NSSC21K1808) to Alexander J. Turner and Tai-Long He. We acknowledge funding from the Environmental Defense Fund, whose work is supported by gifts from Signe Ostby, Scott Cook, and the Valhalla Foundation. This project is supported in part by Schmidt Sciences through the VESRI program. Nikhil Dadheeche also acknowledges funding from the Integral Environmental Big Data Research Fund of the University of Washington. We thank the three anonymous reviewers for their thoughtful feedback and constructive comments. Their suggestions helped improve the clarity and quality of this paper.

Financial support. This research has been supported by the National Aeronautics and Space Administration (grant nos. 80NSSC22K1557 and 80NSSC21K1808). This work received funding from the Environmental Defense Fund, whose work is supported by gifts from Signe Ostby, Scott Cook, and the Valhalla Foundation. This project is supported in part by Schmidt Sciences through the VESRI program. Nikhil Dadheeche was also supported by the Integral Environmental Big Data Research Fund of the University of Washington.

Review statement. This paper was edited by Abhishek Chatterjee and reviewed by three anonymous referees.

References

- Asimow, N. G., Turner, A. J., and Cohen, R. C.: Sustained Reductions of Bay Area CO₂ Emissions 2018–2022, Environ. Sci. Technol., 58, 6586–6594, <https://doi.org/10.1021/acs.est.3c09642>, 2024.

- Bishop, C. M.: Pattern Recognition and Machine Learning, Springer Science+Business Media, LLC, ISBN 978-0-387-31073-2, 2006.
- Bovensmann, H., Buchwitz, M., Burrows, J. P., Reuter, M., Krings, T., Gerilowski, K., Schneising, O., Heymann, J., Tretner, A., and Erzinger, J.: A remote sensing technique for global monitoring of power plant CO₂ emissions from space and related applications, *Atmos. Meas. Tech.*, 3, 781–811, <https://doi.org/10.5194/amt-3-781-2010>, 2010.
- Brandt, A. R., Heath, G. A., Kort, E. A., O’Sullivan, F., Pétron, G., Jordaan, S. M., Tans, P., Wilcox, J., Gopstein, A. M., Arent, D., Wofsy, S., Brown, N. J., Bradley, R., Stucky, G. D., Eardley, D., and Harriss, R.: Methane Leaks from North American Natural Gas Systems, *Science*, 343, 733–735, <https://doi.org/10.1126/science.1247045>, 2014.
- Brech, R., Bakels, L., Bihlo, A., and Stohl, A.: Improving trajectory calculations by FLEXPART 10.4+ using single-image super-resolution, *Geosci. Model Dev.*, 16, 2181–2192, <https://doi.org/10.5194/gmd-16-2181-2023>, 2023.
- Cartwright, L., Zammit-Mangion, A., and Deutscher, N. M.: Emulation of greenhouse-gas sensitivities using variational autoencoders, *Environmetrics*, 34, e2754, <https://doi.org/10.1002/env.2754>, 2023.
- Chen, Y., Sherwin, E. D., Berman, E. S., Jones, B. B., Gordon, M. P., Wetherley, E. B., Kort, E. A., and Brandt, A. R.: Quantifying Regional Methane Emissions in the New Mexico Permian Basin with a Comprehensive Aerial Survey, *Environ. Sci. Technol.*, 56, 4317–4323, <https://doi.org/10.1021/acs.est.1c06458>, 2022.
- Cohen Research: Berkeley Environmental Air-quality & CO₂ Network (BEACO2N), University of California Berkeley [data set], <http://beacon.berkeley.edu>, last access: 15 May 2025.
- Crisp, D., Fisher, B. M., O’Dell, C., Frankenberg, C., Basilio, R., Bösch, H., Brown, L. R., Castano, R., Connor, B., Deutscher, N. M., Eldering, A., Griffith, D., Gunson, M., Kuze, A., Mandrake, L., McDuffie, J., Messerschmidt, J., Miller, C. E., Morino, I., Natraj, V., Notholt, J., O’Brien, D. M., Oyafuso, F., Polonsky, I., Robinson, J., Salawitch, R., Sherlock, V., Smyth, M., Suto, H., Taylor, T. E., Thompson, D. R., Wennberg, P. O., Wunch, D., and Yung, Y. L.: The ACOS CO₂ retrieval algorithm – Part II: Global X_{CO₂} data characterization, *Atmos. Meas. Tech.*, 5, 687–707, <https://doi.org/10.5194/amt-5-687-2012>, 2012.
- Cusworth, D. H., Thorpe, A. K., Ayasse, A. K., Stepp, D., Heckler, J., Asner, G. P., Miller, C. E., Yadav, V., Chapman, J. W., Eastwood, M. L., Green, R. O., Hmiel, B., Lyon, D. R., and Duren, R. M.: Strong methane point sources contribute a disproportionate fraction of total emissions across multiple basins in the United States, *P. Natl. Acad. Sci. USA*, 119, e2202338 119, <https://doi.org/10.1073/pnas.2202338119>, 2022.
- Dadheeche, N.: High-resolution greenhouse gas flux inversions using a machine learning surrogate model, Zenodo [data set and code], <https://doi.org/10.5281/zenodo.13750963>, 2024.
- Duren, R. M., Thorpe, A. K., Foster, K. T., Rafiq, T., Hopkins, F. M., Yadav, V., Bue, B. D., Thompson, D. R., Conley, S., Colombi, N. K., Frankenberg, C., McCubbin, I. B., Eastwood, M. L., Falk, M., Herner, J. D., Croes, B. E., Green, R. O., and Miller, C. E.: California’s methane super-emitters, *Nature*, 575, 180–184, <https://doi.org/10.1038/s41586-019-1720-3>, 2019.
- Eldering, A., Boland, S., Solish, B., Crisp, D., Kahn, P., and Gunson, M.: High precision atmospheric CO₂ measurements from space: The design and implementation of OCO-2, in: 2012 IEEE aerospace conference, Big Sky, MT, USA, 3–10 March 2012, IEEE, <https://doi.org/10.1109/AERO.2012.6187176>, pp. 1–10, 2012.
- Eldering, A., Taylor, T. E., O’Dell, C. W., and Pavlick, R.: The OCO-3 mission: measurement objectives and expected performance based on 1 year of simulated data, *Atmos. Meas. Tech.*, 12, 2341–2370, <https://doi.org/10.5194/amt-12-2341-2019>, 2019.
- Fillola, E., Santos-Rodriguez, R., Manning, A., O’Doherty, S., and Rigby, M.: A machine learning emulator for Lagrangian particle dispersion model footprints: a case study using NAME, *Geosci. Model Dev.*, 16, 1997–2009, <https://doi.org/10.5194/gmd-16-1997-2023>, 2023.
- Frankenberg, C., Thorpe, A. K., Thompson, D. R., Hulley, G., Kort, E. A., Vance, N., Borchardt, J., Krings, T., Gerilowski, K., Sweeney, C., Conley, S., Bue, B. D., Aubrey, A. D., Hook, S., and Green, R. O.: Airborne methane remote measurements reveal heavy-tail flux distribution in Four Corners region, *P. Natl. Acad. Sci. USA*, 113, 9734–9739, <https://doi.org/10.1073/pnas.1605617113>, 2016.
- Guo, W., Shi, Y., Liu, Y., and Su, M.: CO₂ emissions retrieval from coal-fired power plants based on OCO-2/3 satellite observations and a Gaussian plume model, *J. Clean. Prod.*, 397, 136 525, <https://doi.org/10.1016/j.jclepro.2023.136525>, 2023.
- Hammerling, D. M., Michalak, A. M., and Kawa, S. R.: Mapping of CO₂ at high spatiotemporal resolution using satellite observations: Global distributions from OCO-2, *J. Geophys. Res.-Atmos.*, 117, D06306, <https://doi.org/10.1029/2011JD017015>, 2012.
- He, T.-L.: STILT footprints data set 2, Zenodo [data set], <https://doi.org/10.5281/zenodo.12803736>, 2024.
- He, T.-L., Boyd, R. J., Varon, D. J., and Turner, A. J.: Increased methane emissions from oil and gas following the Soviet Union’s collapse, *P. Natl. Acad. Sci. USA*, 121, e2314600121, <https://doi.org/10.1073/pnas.2314600121>, 2024.
- He, T.-L., Dadheeche, N., Thompson, T. M., and Turner, A. J.: FootNet v1.0: development of a machine learning emulator of atmospheric transport, *Geosci. Model Dev.*, 18, 1661–1671, <https://doi.org/10.5194/gmd-18-1661-2025>, 2025.
- Hutyra, L. R., Duren, R., Gurney, K. R., Grimm, N., Kort, E. A., Larson, E., and Shrestha, G.: Urbanization and the carbon cycle: Current capabilities and research outlook from the natural sciences perspective, *Earths Future*, 2, 473–495, <https://doi.org/10.1002/2014EF000255>, 2014.
- IPCC: Climate change 2023: synthesis report. Contribution of working groups I, II and III to the sixth assessment report of the intergovernmental panel on climate change, The Australian National University, <https://doi.org/10.59327/IPCC/AR6-9789291691647.001>, 2023.
- Janardanan, R., Maksyutov, S., Oda, T., Saito, M., Kaiser, J. W., Ganshin, A., Stohl, A., Matsunaga, T., Yoshida, Y., and Yokota, T.: Comparing GOSAT observations of localized CO₂ enhancements by large emitters with inventory-based estimates, *Geophys. Res. Lett.*, 43, 3486–3493, 2016.
- Janardanan, R., Maksyutov, S., Wang, F., Nayagam, L., Sahu, S. K., Mangaraj, P., Saunois, M., Lan, X., and Matsunaga, T.: Country-level methane emissions and their sectoral trends during 2009–2020 estimated by high-resolution inversion of GOSAT

- and surface observations, Environ. Res. Lett., 19, 034007, <https://doi.org/10.1088/1748-9326/ad2436>, 2024.
- Kiel, M., Eldering, A., Roten, D. D., Lin, J. C., Feng, S., Lei, R., Lauvaux, T., Oda, T., Roehl, C. M., Blavier, J.-F., and Iraci, L. T.: Urban-focused satellite CO₂ observations from the Orbiting Carbon Observatory-3: A first look at the Los Angeles megacity, Remote Sens. Environ., 258, 112314, <https://doi.org/10.1016/j.rse.2021.112314>, 2021.
- Lauvaux, T., Giron, C., Mazzolini, M., d'Aspremont, A., Duren, R., Cusworth, D., Shindell, D., and Ciais, P.: Global assessment of oil and gas methane ultra-emitters, Science, 375, 557–561, <https://doi.org/10.1126/science.abj4351>, 2022.
- Lin, J. C., Gerbig, C., Wofsy, S. C., Andrews, A. E., Daube, B. C., Davis, K. J., and Grainger, C. A.: A near-field tool for simulating the upstream influence of atmospheric observations: The Stochastic Time-Inverted Lagrangian Transport (STILT) model, J. Geophys. Res.-Atmos., 108, 4493, <https://doi.org/10.1029/2002JD003161>, 2003.
- Nassar, R., Hill, T. G., McLinden, C. A., Wunch, D., Jones, D. B. A., and Crisp, D.: Quantifying CO₂ Emissions From Individual Power Plants From Space, Geophys. Res. Lett., 44, 10045–10053, <https://doi.org/10.1002/2017GL074702>, 2017.
- Nassar, R., Mastrogiacomo, J.-P., Bateman-Hemphill, W., McCracken, C., MacDonald, C. G., Hill, T., O'Dell, C. W., Kiel, M., and Crisp, D.: Advances in quantifying power plant CO₂ emissions with OCO-2, Remote Sens. Environ., 264, 112579, <https://doi.org/10.1016/j.rse.2021.112579>, 2021.
- Nayagam, L., Maksyutov, S., Oda, T., Janardanan, R., Trisolino, P., Zeng, J., Kaiser, J. W., and Matsunaga, T.: A top-down estimation of subnational CO₂ budget using a global high-resolution inverse model with data from regional surface networks, Environ. Res. Lett., 19, 014031, <https://doi.org/10.1088/1748-9326/ad0f74>, 2023.
- O'Dell, C. W., Connor, B., Bösch, H., O'Brien, D., Frankenberg, C., Castano, R., Christi, M., Eldering, D., Fisher, B., Gunson, M., McDuffie, J., Miller, C. E., Natraj, V., Oyafuso, F., Polonsky, I., Smyth, M., Taylor, T., Toon, G. C., Wennberg, P. O., and Wunch, D.: The ACOS CO₂ retrieval algorithm – Part 1: Description and validation against synthetic observations, Atmos. Meas. Tech., 5, 99–121, <https://doi.org/10.5194/amt-5-99-2012>, 2012.
- Rodgers, C. D.: Inverse methods for atmospheric sounding: theory and practice, vol. 2, World Scientific, ISBN 981-02-2740-X, 2000.
- Roten, D., Wu, D., Fasoli, B., Oda, T., and Lin, J. C.: An Interpolation Method to Reduce the Computational Time in the Stochastic Lagrangian Particle Dispersion Modeling of Spatially Dense XCO₂ Retrievals, Earth Space Sci., 8, e2020EA001343, <https://doi.org/10.1029/2020EA001343>, 2021.
- Sherwin, E. D., Rutherford, J. S., Chen, Y., Aminfarid, S., Kort, E. A., Jackson, R. B., and Brandt, A. R.: Single-blind validation of space-based point-source detection and quantification of onshore methane emissions, Sci. Rep.-UK, 13, 3836, <https://doi.org/10.1038/s41598-023-30761-2>, 2023.
- Shusterman, A. A., Teige, V. E., Turner, A. J., Newman, C., Kim, J., and Cohen, R. C.: The BErkeley Atmospheric CO₂ Observation Network: initial evaluation, Atmos. Chem. Phys., 16, 13449–13463, <https://doi.org/10.5194/acp-16-13449-2016>, 2016.
- Shusterman, A. A., Kim, J., Lieschke, K. J., Newman, C., Wooldridge, P. J., and Cohen, R. C.: Observing local CO₂ sources using low-cost, near-surface urban monitors, Atmos. Chem. Phys., 18, 13773–13785, <https://doi.org/10.5194/acp-18-13773-2018>, 2018.
- Steiner, M., Peters, W., Luijx, I., Henne, S., Chen, H., Hammer, S., and Brunner, D.: European CH₄ inversions with ICON-ART coupled to the CarbonTracker Data Assimilation Shell, Atmos. Chem. Phys., 24, 2759–2782, <https://doi.org/10.5194/acp-24-2759-2024>, 2024.
- Taylor, T. E., Eldering, A., Merrelli, A., Kiel, M., Somkuti, P., Cheng, C., Rosenberg, R., Fisher, B., Crisp, D., Basilio, R., Bennett, M., Cervantes, D., Chang, A., Dang, L., Frankenberg, C., Haemmerle, V. R., Keller, G. R., Kurosu, T., Laughner, J. L., Lee, R., Marchetti, Y., Nelson, R. R., O'Dell, C. W., Osterman, G., Pavlick, R., Roehl, C., Schneider, R., Spiers, G., To, C., Wells, C., Wennberg, P. O., Yelamanchili, A., and Yu, S.: OCO-3 early mission operations and initial (vEarly) XCO₂ and SIF retrievals, Remote Sens. Environ., 251, 112032, <https://doi.org/10.1016/j.rse.2020.112032>, 2020.
- Turner, A. J., Shusterman, A. A., McDonald, B. C., Teige, V., Harley, R. A., and Cohen, R. C.: Network design for quantifying urban CO₂ emissions: assessing trade-offs between precision and network density, Atmos. Chem. Phys., 16, 13465–13475, <https://doi.org/10.5194/acp-16-13465-2016>, 2016.
- Turner, A. J., Jacob, D. J., Benmergui, J., Brandman, J., White, L., and Randles, C. A.: Assessing the capability of different satellite observing configurations to resolve the distribution of methane emissions at kilometer scales, Atmos. Chem. Phys., 18, 8265–8278, <https://doi.org/10.5194/acp-18-8265-2018>, 2018.
- Turner, A. J., Kim, J., Fitzmaurice, H., Newman, C., Worthington, K., Chan, K., Wooldridge, P. J., Köehler, P., Frankenberg, C., and Cohen, R. C.: Observed Impacts of COVID-19 on Urban CO Emissions, Geophys. Res. Lett., 47, e2020GL090037, <https://doi.org/10.1029/2020GL090037>, 2020.
- Varon, D. J., Jacob, D. J., Sulprizio, M., Estrada, L. A., Downs, W. B., Shen, L., Hancock, S. E., Nesser, H., Qu, Z., Penn, E., Chen, Z., Lu, X., Lorente, A., Tewari, A., and Randles, C. A.: Integrated Methane Inversion (IMI 1.0): a user-friendly, cloud-based facility for inferring high-resolution methane emissions from TROPOMI satellite observations, Geosci. Model Dev., 15, 5787–5805, <https://doi.org/10.5194/gmd-15-5787-2022>, 2022.
- Varon, D. J., Jacob, D. J., Hmiel, B., Gautam, R., Lyon, D. R., Omara, M., Sulprizio, M., Shen, L., Pendergrass, D., Nesser, H., Qu, Z., Barkley, Z. R., Miles, N. L., Richardson, S. J., Davis, K. J., Pandey, S., Lu, X., Lorente, A., Borsdorff, T., Maasakers, J. D., and Aben, I.: Continuous weekly monitoring of methane emissions from the Permian Basin by inversion of TROPOMI satellite observations, Atmos. Chem. Phys., 23, 7503–7520, <https://doi.org/10.5194/acp-23-7503-2023>, 2023.
- Veefkind, J., Aben, I., McMullan, K., Förster, H., de Vries, J., Otter, G., Claas, J., Eskes, H., de Haan, J., Kleipool, Q., van Weele, M., Hasekamp, O., Hoogeveen, R., Landgraf, J., Snel, R., Tol, P., Ingmann, P., Voors, R., Kruizinga, B., Vink, R., Visser, H., and Levelt, P.: TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications, Remote Sens. Environ., 120, 70–83, <https://doi.org/10.1016/j.rse.2011.09.027>, 2012.

- Wang, Y., Broquet, G., Bréon, F.-M., Lespinas, F., Buchwitz, M., Reuter, M., Meijer, Y., Loescher, A., Janssens-Maenhout, G., Zheng, B., and Ciais, P.: PMIF v1.0: assessing the potential of satellite observations to constrain CO₂ emissions from large cities and point sources over the globe using synthetic data, *Geosci. Model Dev.*, 13, 5813–5831, <https://doi.org/10.5194/gmd-13-5813-2020>, 2020.
- Wu, D., Lin, J. C., Oda, T., and Kort, E. A.: Space-based quantification of per capita CO₂ emissions from cities, *Environ. Res. Lett.*, 15, 035004, <https://doi.org/10.1088/1748-9326/ab68eb>, 2020.
- Yadav, V. and Michalak, A. M.: Improving computational efficiency in large linear inverse problems: an example from carbon dioxide flux estimation, *Geosci. Model Dev.*, 6, 583–590, <https://doi.org/10.5194/gmd-6-583-2013>, 2013.
- Zavalá-Araiza, D., Lyon, D. R., Alvarez, R. A., Davis, K. J., Harriss, R., Herndon, S. C., Karion, A., Kort, E. A., Lamb, B. K., Lan, X., Marchese, A. J., Pacala, S. W., Robinson, A. L., Shepson, P. B., Sweeney, C., Talbot, R., Townsend-Small, A., Yacovitch, T. I., Zimmerle, D. J., and Hamburg, S. P.: Reconciling divergent estimates of oil and gas methane emissions, *P. Natl. Acad. Sci. USA*, 112, 15597–15602, <https://doi.org/10.1073/pnas.1522126112>, 2015.