Supplement of Atmos. Meas. Tech., 18, 5375–5391, 2025
https://doi.org/10.5194/amt-18-5375-2025-supplement

Atmospheric Measurement Techniques
Open Access
EGU

Supplement of

Performance evaluation of multi-source methane emission quantification models using fixed-point continuous monitoring systems

David Ball et al.

Correspondence to: Ali Lashgari (ali.lashgari@projectcanary.com)

Contents

S1 Extended Introduction to Measurement Informed Inventories and Methane Emissions Measurement and Quantification Techniques

Recent studies employing various emissions measurement techniques indicate that existing emission inventories underestimate actual methane emissions from the oil and gas sector (Johnson et al., 2017; Alvarez et al., 2018; Robertson et al., 2020; Maasakkers et al., 2021; Plant et al., 2022; Shen et al., 2022; Lu et al., 2023; Tibrewal et al., 2024; Omara et al., 2024; Williams et al., 2025a; Lu et al., 2022; Worden et al., 2022), with significant regional variations in the discrepancies between bottom-up inventories and measurement-based estimates (Ravikumar et al., 2025; Vallejo et al., 2024). A recent study found methane emission loss rates ranging from approximately 0.75% to 9.63% of natural gas production across different regions of the US (Sherwin et al., 2022). Despite variations in emission rate distributions among different oil and gas basins, recent research indicates that the majority of methane emissions are originated from facilities that emit at rates lower than 100 kg/hr (Williams et al., 2025b).

Traditional approaches for detecting methane emissions often rely on human senses (auditory, visual, and olfactory (AVO) inspections) or portable sensors used in close proximity to potential sources. These include AVO inspections, EPA Method 21 (a sensitive but labor-intensive EPA-approved method), and optical gas imaging (OGI), which uses infrared cameras to visualize methane leaks (Xia et al., 2024).

Advances in communication technologies, the Internet of Things (IoT), and reduced sensor costs have facilitated the development of next-generation emission measurement (NGEM) technologies. Unlike traditional methods, NGEM technologies operate at a distance from the source, attempting to remotely detect, localize, and quantify emission events. These technologies operate from space (satellites), air (aircraft and UAVs), or the ground (fixed, mobile, and handheld sensors). Effective and efficient methane emission detection and quantification often requires a combination of measurement technologies, commonly referred to as a multiscale measurement approach (Wang et al., 2022; Daniels et al., 2024).

Satellite and aerial remote sensing techniques can detect emissions from specific sources, with aerial methods capable of detecting emissions as low as 1-3 kg/hr, while satellites have minimum detection limits of to approximately 200 kg/hr or higher (Sherwin et al., 2023). Simulation studies based on the Fugitive Emissions Abatement Simulation Toolkit (FEAST), suggest that a minimum detection limit of 0.1–1 kg/hour is sufficient to capture almost all emissions from the oil and gas sector (Ravikumar et al., 2018). Based on the results of this study, quantifying emissions below this threshold would not substantially improve mitigation efforts.

Satellite measurements offer remote detection of large methane releases, improving transparency and independent monitoring, particularly in remote or under-studied areas. Although, their high detection thresholds and infrequent overpasses limit their utility. Aerial platforms (piloted aircraft and UAVs) in general offer lower detection limits compared to satellites, but may introduce spatial extrapolation errors when regional inventory development is the objective. Because a considerable portion of methane emissions from oil and gas operations are intermittent, and snapshot methods only provide an instantaneous measurement (Santos et al., 0), they are unable to provide robust estimates of the total mass emitted from a detected plume due to the inherent uncertainty in the required temporal extrapolation. In addition to higher detection limits, common limitations of all snapshot measurement techniques are the lack of site-specific meteorological data and information on emission event timing (e.g., duration and frequency of intermittent releases), which pose a significant challenge to the temporal extrapolation of results.

AERMOD (a Gaussian plume model) and CALPUFF (a Lagrangian puff model) are two air dispersion models commonly used for regulatory purposes. AERMOD is the EPA's recommended model for near-field applications (up to 50 km) (Houweling et al., 1999). This model relies on the assumption of steady-state conditions, which is often unrealistic in the presence of variable wind conditions (Jia et al., 2023). Some of the previous studies employed longer time averaging of wind data to satisfy this steady-state requirement. However, this approach results in masking important short-term wind variations and associated concentration enhancements that are crucial for accurate dispersion modeling at scales relevant to oil and gas facilities.

CALPUFF has not been designated as an EPA-preferred model for near-field applications. However, it may be considered as an alternative dispersion modeling method on a case-by-case basis for near-field applications involving complex winds (Wayland, 2008). CALPUFF is often preferred for long-range transport and complex terrain.

Both AERMOD and CALPUFF have limitations when applied to methane dispersion in near-source applications, such as upstream oil and gas. These limitations stem from the models' underlying assumptions and simplifications, which may not be appropriate in tens of meters-long source-to-sensor distances and lack accurate capturing of the complex atmospheric processes governing methane dispersion in upstream facilities. In a recent attempt, (Jia et al., 2025) presented a computationally efficient and scalable implementation of the Gaussian puff model for atmospheric dispersion. This model incorporates dynamic spatiotemporal thresholding to achieve a shorter runtime, making it suitable for real-time applications and large-scale deployments.

S2 Dispersion Models

The following subsections provide brief overviews of the theory underpinning the dispersion models, followed by more specific implementation details. Note that there are myriad small choices (e.g., stability class calculations, dispersion parametrization)

that must be made in the data processing and algorithmic workflow when it comes to running these dispersion models. It is outside the scope of this study to enumerate and present results from every combination of valid choices. Instead, we will provide clear justifications for the specific choices made in this study and demonstrate the efficacy of the models under these specific implementations. It should be noted that the impact of most of these higher-order decisions on the results is minimal, as they are often different approaches of approximating the same underlying phenomena. For example, there are several commonly-used functional forms and associated coefficients to describe how the dispersion of a gas plume scales with distance. While these empirical formulae may look very different (e.g., some utilize power laws while others employ logarithms), they are generally inferred by fitting these functional forms to the same underlying data, and result in similar general characteristics despite the sometimes dramatically different functional forms.

S2.1 Gaussian Plume

The commonly used Gaussian Plume model (GPM) provides a closed-form solution to the steady-state advection-diffusion equation for a single point source emitting at rate Q from height H. As a steady-state model, GPM assumes constant wind speed, wind direction, and source rate over time. Pollution concentration, C at a given location (x, y, z) is calculated as:

$$C(x,y,z) = \frac{Q}{2\pi\bar{u}\sigma_y\sigma_z}\exp\left[-\frac{y^2}{2\sigma_y^2} - \frac{(z-H)^2}{2\sigma_z^2} - \frac{(z+H)^2}{2\sigma_z^2}\right] \tag{S1}$$

In this coordinate system, x is defined as the downwind distance (aligned with the average wind velocity, the magnitude of which is denoted by $\bar{u}$), y represents the crosswind distance (perpendicular to the wind direction), and z is the receptor height above the ground. The dispersion coefficients σ_y and σ_z represent the horizontal and vertical spread of the plume, respectively. The rate at which plume spread increases with downwind distance is influenced by the stability class, a metric that represents the degree of turbulence in the atmosphere. The approximation of dispersion coefficients can be achieved using various established methods (Carruthers et al., 2009). The specific implementation details used for this study are presented later in this Section. The concentration levels modeled by the GPM represent temporal averages that capture the statistics of the turbulent dispersion by small variations in the wind field. As such, an appropriate averaging time must be chosen that is sufficiently long for small-scale turbulence effects to average out, but not so long that coherent changes in the wind direction occur.

It is important to note how different turbulent wavenumbers (k) affect a plume at different characteristic scales (L). For reference, k can be thought of simply as the inverse spatial scale of a turbulent eddy, $k = 2\pi/L_{eddy}$, where L_{eddy} represents the characteristic length scale associated with a particular turbulent eddy. Turbulence scales where $k > 1/L$ (i.e., when turbulent fluctuations are smaller than the length scale of the plume) cause plume dispersion. In other words, small-scale eddies disperse the plume randomly in a way that is captured by σ_y and σ_z. If a turbulent eddy has wavenumber comparable to (or less) than $1/L$, however, this results in meandering motions (or coherent changes in wind direction) for extended periods of time that do

not result in statistical dispersion. In this case, the Gaussian Plume will fail to capture the effects of these variations if they are averaged over.

Many applications of the Gaussian Plume are tuned to model dispersion on long length scales (kilometers or greater), and hence require long (> 30 minute) averaging periods. However, when GPM is applied to fenceline monitors placed around facilities with scales of around 100 meters, much shorter averaging periods must be used. This ensures that turbulent eddies on scales comparable to or larger than the source-sensor distance are captured temporally as coherent changes in wind direction, rather than being averaged over. As such, shorter averaging times (1-5 minutes) are more appropriate for near-source dispersion modeling applications.

When employing Gaussian Plume models, any violation of the assumptions that go into the derivation of the GPM results in unreliable concentration predictions from this dispersion model. These effects can be especially pernicious at low wind speeds. The underlying assumption of steady-state wind is often violated during low-wind conditions because the ratio between the reference velocity scale of turbulence u_{ref}, which is typically taken as the root-mean-square of the horizontal wind fluctuations u'_{rms}, scales with the mean wind speed used in Equation S1, i.e. $\frac{u_{ref}}{\bar{u}} \sim 1$. Practically, this implies that the underlying assumption for a Gaussian concentration distribution is no longer valid. In addition, frequent and significant changes in wind direction are more common at low wind speeds, making the application of GPM less optimal, as this model requires that the mean wind direction should align with the x (downwind) axis. Also note that since $\bar{u}$ is in the denominator of Equation S1, in low wind speed conditions, concentration predictions will become nonphysically high. Different corrections exist to the GPM under low wind conditions. The approach by (Carruthers et al., 1994) reconstructs concentration under low wind conditions by using a weighted average between a Gaussian-shaped plume state and a random-walk state in which the release spreads in all horizontal radial directions.

Many of the standard methods for computing the dispersion coefficients σ_y and σ_z rely first on an approximation of the Pasquill atmospheric stability class (ASC) (Mohan and Siddiqui, 1998; Venkatram, 1996). Several approaches exist for ASC approximation, including using near-surface vertical temperature gradient (Randerson, 1984) or inferring them from estimates of solar irradiation (based on cloud cover and time of day), and combining this information with wind speed to infer the stability class via a look-up table (Pasquill, 1961). The current study uses the statistics of the fluctuating horizontal wind field to estimate the ASC at minute t by computing the circular standard deviation of horizontal wind direction (σ_θ) over the trailing 15 minutes of wind data and mapping these values to a stability class via Tables 6-8a and 6-8b found in USEPA (2000). This method is chosen over the other approaches because it utilizes a direct measurement of the degree of turbulence in the atmosphere via σ_θ rather than exclusively indirect indicators (solar irradiation, wind speed, cloud cover).

The majority of the methods for calculating dispersion parameters σ_y and σ_z as a function of downwind distance and stability class provide comparable dispersion profiles that are approximated via different empirically-estimated functional forms and associated coefficients (Randerson, 1984; R. P. Hosker, 1974). This study employs the "Martin Method" (first proposed in Martin (1976)) which utilizes power law functions to describe both the horizontal and vertical spread of the plume via:

$$\sigma_y = ax^b$$
$$\sigma_z = cx^d + f \tag{S2}$$

where the parameters a, b, c, d and f depend on the ASC and can be found in a look-up table in Martin (1976). Using these functional forms for the dispersion parameters, Equation S1 is applied to every minute of measurement data for every source-sensor combination for the time period associated with a given controlled release experiment. This results in generating the previously-described sensitivity matrix $\mathbf{S}$, the elements of which represent the predicted concentration via Equation S1 (for a unit rate). In this matrix, columns correspond to individual sources and rows correspond to the location and height of a specific measurement device at a specific time. Due to the failure of the GPM to accurately capture pollutant transport under low wind speeds, concentrations from time periods with wind speeds of less than 0.5 meters per second are excised from the sensitivity matrix.

Since GPM utilizes a closed-form algebraic solution (no time-marching is needed), it is computationally fast. Several hours of measurement data can be processed via GPM and corresponding concentration predictions can be made for every source for every sensor location in milliseconds when the calculations are vectorized properly.

S2.2 Gaussian Puff

The Gaussian Puff model is a Lagrangian approach to approximating the solution to the advection-diffusion equation that makes fewer assumptions than the Gaussian Plume. More specifically, this method can capture the relevant physical effects embedded in spatially varying wind fields (i.e., it does not assume homogeneous wind fields), can handle time-varying emission rates (does not assume steady-state emission rate), and also more properly account for low wind speeds and unstable conditions when the wind vector rapidly changes (does not assume steady-state wind fields). Because of these features, Gaussian Puff is more broadly applicable than GPM and has fewer potential failure modes (e.g., low wind speeds and improper averaging times). The Gaussian Puff method approximates advection and diffusion by emitting a series of "puffs" that are advected with the wind field. The position of each puff, $\boldsymbol{x}$, as a function of time is simply:

$$\boldsymbol{x}(t) = \boldsymbol{x}_0 + \int_0^t \boldsymbol{u}(\boldsymbol{x}, t')dt' \tag{S3}$$

where $\boldsymbol{x}_0$ represents the puff's initial location (given by the emission point from which the puff originated). Note that the wind vector ($\boldsymbol{u}$) used to advance the position of the puff can be a function of both position and time. When high-frequency spatial information about the wind field is available (e.g., via multiple anemometers), these effects can be included in the evolution of the puff positions to obtain a more accurate puff trajectory. In this method, puffs are emitted at some fixed interval Δt from each point source, with each puff containing total mass $M = Q_i \Delta t$, where Q_i represents the emission rate of source i from which the puff was emitted. This mass is distributed across a three-dimensional Gaussian profile such that the concentration at a given point can be expressed via:

$$C(x,y,z,t) = \frac{Q\Delta t}{(2\pi)^{3/2}\sigma_x\sigma_y\sigma_z} \exp\left[-\frac{y^2}{2\sigma_y^2} - \frac{(z-H)^2}{2\sigma_z^2} - \frac{(z+H)^2}{2\sigma_z^2} - \frac{x^2}{2\sigma_x^2}\right]. \tag{S4}$$

Note that each coordinate (x,y,z) and associated dispersion coefficients (σ's) are functions of time (the explicit time dependence of these variables in Equation S4 is omitted for the sake of readability). In the implementation of the Gaussian Puff algorithm used in this study, higher frequency temporal wind measurements are used to advance the positions of the puffs via Equation S3. More specifically, 1Hz sampled wind measurement are interpolated to the position of each individual puff using a distance-weighted average with inverse square weights to approximate the wind field at a specific location at a specific time. In principle, this should result in more physically accurate puff trajectories that account for both high-frequency spatial and temporal variations. In other words, $u(\boldsymbol{x}, t)$ is computed via:

$$\boldsymbol{u}(\boldsymbol{x}, t) = \frac{\sum_i \boldsymbol{u}(\boldsymbol{x}_i, t)/d_i^2}{\sum_i 1/d_i^2} \tag{S5}$$

where $\boldsymbol{u}(\boldsymbol{x_i}, t)$ represents an anemometer measurement at time t from location $\boldsymbol{x}_i$ and d_i is the Euclidean distance between the puff location $\boldsymbol{x}$ and the location associated with a given anemometer measurement, $\boldsymbol{x}_i$. Note that in this study, each Canary X device was equipped with an anemometer, so the distance-weighted average of 10 anemometers is used to advance the position of the puffs. Wind measurements from 10 devices are highly redundant at such a small and simple facility; the horizontal and wind directions measured by all the devices on the facility are highly correlated and differ by only a small amount. At more complex facilities where more spatial variation of the wind field may be expected (e.g., due to significant topographical variations or obstructive complexity), then the use of multiple anemometers may be more advantageous. One puff is emitted for every second of the simulation (i.e., $\Delta t = 1$ second) and every puff's position is advanced every second.

In contrast to GPM, the Gaussian puff framework requires time advancement of each puff's spatial positioning and associated modulation of the concentration profile of each as the simulation advances. This is further coupled with processing procedures needed for the addition of new puffs and their proper accounting. Consequently, this makes the Gaussian Puff simulation for an equal number of measurement hours much slower than the GPM.

S3 Computational Fluid Dynamics

The choice of employing CFD simulations is motivated and constrained by two primary reasons: *(i)* a desire to perform three-dimensional (3D), unsteady modeling of multi-scale emissions from oil & gas facilities where the underlying wind field is directly modeled/resolved onto a computational grid while accounting for the complex effects induced by the presence of obstructions, that are otherwise not included in the previously discussed dispersion models; and *(ii)* to have a numerical framework that offers a graduated step up from the Gaussian models presented earlier while operating in effectively the same vein to permit a direct comparison of model performance.

The flow solver makes use of the finite-difference framework on a staggered Cartesian mesh and a fractional step approach, alternatively known as the predictor-corrector method, to solve the low-pass filtered incompressible Navier-Stokes equation:

$$\frac{\partial \widetilde{u_i}}{\partial x_i} = 0,$$
$$\frac{\partial \widetilde{u_i}}{\partial t} + \frac{\partial \widetilde{u_i}\widetilde{u_j}}{\partial x_j} = -\frac{1}{\rho_0}\frac{\partial \widetilde{p}}{\partial x_i} - \frac{\partial \tau_{ij}}{\partial x_j} + b\delta_{i3}. \tag{S6}$$

Solution of the momentum equation requires inverting the following Poisson equation prior to the corrector step to enforce the divergence-free condition on the 3D velocity field:

$$\frac{\partial^2(\delta\widetilde{p})}{\partial x_i^2} = \frac{\rho_0}{\Delta t}\frac{\partial \widetilde{u^*}_i}{\partial x_i}. \tag{S7}$$

Here, $x_i = (x, y, z)$ represents the horizontal (streamwise: x and spanwise: y) and the vertical z directions of the spatial coordinate system used in the CFD simulations. The asterisk on the right of Equation S7 indicates the intermediate velocity field obtained after the predictor step and does not satisfy the continuity equation. Time advancement is accomplished via a Newton-Raphson based iterative method and spatial derivatives in the momentum equation are discretized using central differences. The overall numerical approach is globally second-order accurate in both space and time. Further details on the numerical approach can be found in (Pierce, 2001). Physical obstructions are mapped onto the Cartesian grid via an immersed boundary method

(IBM) (Kim et al., 2001). The tool allows the ingestion of a digital elevation model to reproduce the undulating surface topography of complex industrial sites for use in CFD simulations. The tool is parallelized using domain decomposition coupled with the Message Passing Interface (MPI), and was employed successfully in the past to perform direct numerical simulations of both equilibrium and non-equilibrium turbulent channels flows and boundary layers over explicitly resolved rough surfaces on massively-parallel computing architectures (Ismail, 2023).

In Equations S6 and S7, $\widetilde{\ }$ represents the filtered fields whereas $u_i = (u, v, w)$, p and ρ are the instantaneous wind velocity, pressure and density, respectively. The subgrid stress τ_{ij} is modeled using the Smagorinsky approach: $\tau_{ij} \equiv \widetilde{u_i u_j} - \widetilde{u_i}\widetilde{u_j} = -2\nu_{SGS}\widetilde{S_{ij}}$. The linear constitutive relation between τ_{ij} and the filtered strain rate,

$$S_{ij} \equiv \frac{1}{2}\left(\frac{\partial \widetilde{u_i}}{\partial x_j} + \frac{\partial \widetilde{u_j}}{\partial x_i}\right) \tag{S8}$$

embedded in the Smagorinsky approach is modulated by the eddy viscosity ν_{SGS}. The eddy viscosity is related to the local characteristic length and time scales using the following mixing-length type formulation: $\nu_{SGS} = (C_S\Delta)^2|\widetilde{S}|$, where $|\widetilde{S}| = \sqrt{2\widetilde{S_{ij}}\widetilde{S_{ij}}}$ and Δ is local characteristic length scale taken as $\Delta \equiv (\mathrm{vol})^{1/3}$ (vol = local grid cell volume). We determine the Smagorinsky constant C_S using the dynamic procedure proposed by (Lilly, 1992). A wall model is included that promotes the modulation of the boundary layer at high Reynolds numbers due to surface roughness. The wall model, following (Mukha et al., 2019), imposes a no-slip condition for the horizontal velocity components and forces the computed wall shear stress by requiring the eddy viscosity be determined by,

$$\nu_{SGS} = \frac{u_*^2}{[(\widetilde{u^P}/h_p)^2 + (\widetilde{v^P}/h_p)^2]^{1/2}} \tag{S9}$$

The superscript p in equation S9 refers to the cell at a wall-normal distance h_p from the surface. The filtered horizontal velocity field at this height is used to estimate the local wall shear stress u_*^2 in the above equation using the 'law of the wall' for fully rough surfaces: $u_*^{\mathrm{est}} = \kappa\sqrt{\widetilde{u}^2 + \widetilde{v}^2}/\ln(h_p/z_0)$. The METEC facility is approximated as belonging to the open-terrain category and thus the roughness length z_0 is chosen as 0.03 m. The von-Karman constant in the law of the wall is 0.4.

Contaminants like methane are treated as passive tracers and their transport is modeled by the following filtered advection-diffusion equations:

$$\frac{\partial \widetilde{c}}{\partial t} + \frac{\partial \widetilde{c}\widetilde{u_j}}{\partial x_j} = \frac{\partial \pi_j}{\partial x_j} + S_c. \tag{S10}$$

Spatial derivatives in equation S10 are discretized by the Quadratic Upstream Interpolation for Convective Kinematics (QUICK) scheme. In the filtered scalar transport equation (Equation S10), $S_c(t)$ represents the point sources inside the fluid domain, which are approximated as 3D Gaussian functions following the approach by (Ražnjević et al., 2022). While this equation is written for a single tracer, the tool permits the inclusion of a user-specified number of scalar fields all of which, in principle, can have point sources of varying strength placed at arbitrary positions. The closure term due to the subgrid scalar flux π_j is parametrized using the following gradient-diffusion procedure: $\pi_j \equiv \widetilde{cu_j} - \widetilde{c}\widetilde{u_j} = -\nu_{SGS}/Sc_T(\partial\widetilde{c}/\partial x_j)$. The turbulent Schmidt number Sc_T is fixed at 0.7 (Qian and Venkatram, 2011).

$$\frac{\partial b}{\partial t} + \frac{\partial b\widetilde{u_j}}{\partial x_j} = \alpha_b \frac{\partial^2 b}{\partial x_j^2} \tag{S11}$$

The effect of thermal buoyancy is modeled via the Boussinesq Approximation and the equation for conservation of energy is instead replaced by the buoyancy variable: $b \equiv (\theta_v'/\theta_{v0})g$ (Van Heerwaarden et al., 2017). Here, θ_v', θ_{v0} and g are the fluctuating potential temperature, the reference virtual reference potential temperature and the acceleration due to gravity. The diffusivity coefficient for the buoyancy field α_b is treated like Sc_T.

No-slip and impermeability conditions are applied on the bottom surface for the horizontal and vertical components of the wind field, respectively. A zero-flux condition is used for the scalar fields at the lower surface. Periodic and zero-flux boundary conditions are imposed on the lateral boundary surfaces for velocity and scalar fields, respectively.

As our focus is on near-surface pollutant dispersion, we do not construct the entire boundary layer. Instead, the ceiling of the domain is capped at 200 m, and a no-flux condition is imposed at this height. The horizontal mesh is uniformly spaced with a resolution of $1.5\mathrm{mx}1.5\mathrm{m}$. The mesh is stretched vertically to accommodate the ground-imposed anisotropy: the grid spacing at the ground is 0.2 m, and is gradually relaxed to 2 m towards the top boundary, with the maximum expansion ratio between successive grid points in the z direction remaining below 1.10.

We further leverage the periodicity and uniform grid spacing in the horizontal direction to perform a discrete Fourier transform of the Poisson equation (Equation S7) into a set of linearly independent 1D modified Helmholtz equations that are then solved using the tridiagonal algorithm. Figure S1a shows the 3D digital elevation model (DEM) of the METEC facility after it is mapped onto the Cartesian grid. The positions of the virtual sensors are highlighted using purple spheres. It should be noted that the METEC site exhibits an effectively flat surface, that justifies its earlier classification as open terrain: the only significant obstructions are the three tanks at the center of the facility that have a height of $\sim 4m$.

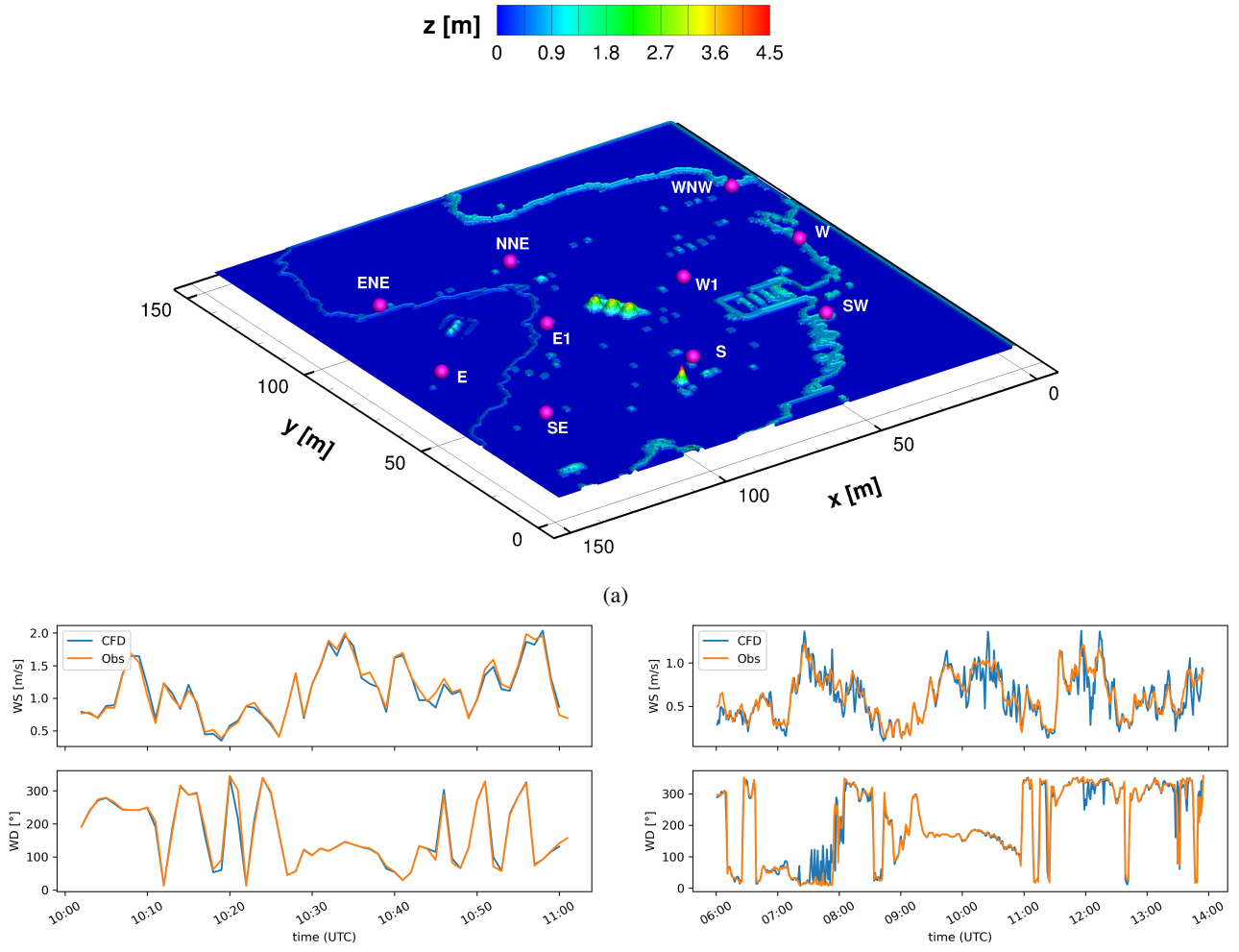


Figure S1. *(a)* Visualizing the 3D surface terrain after mapping the digital elevation model of the facility onto the CFD mesh. The color map indicates the local elevation z in meters and the magenta spheres identify the position of the point sensors. *(b,c)* Comparison of the minute-averaged computed wind speed and direction at the virtual NE sensor height from the CFD simulations with onsite observations for two experiments: one from *(b)* Feb. 08, 2024 and another from *(c)* April. 11, 2024.

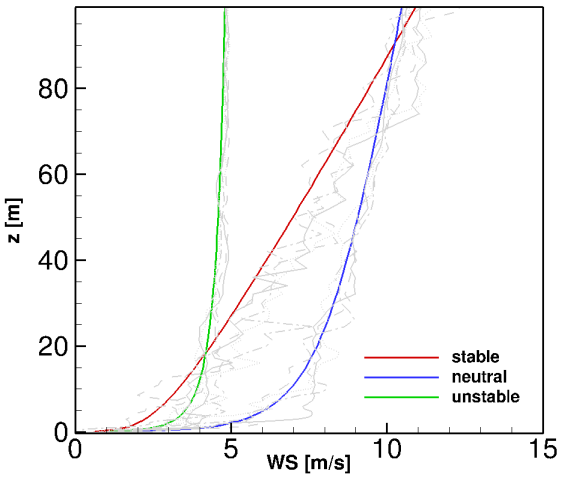
Preliminary simulations were performed to dissect the signal detected at virtual sensor stations both with the effect of the DEM included and without it, and found negligible value in modeling the tanks as obstacles for the present scenario. The downwash effects presumably induced by the obstacles typically extend less than 5 times the mean obstacle height (LEONARDI et al., 2003). The closest sensors to the tanks are however at least over 20m from the centroid of the tanks. Regardless, the simulations

presented later in this investigation include the effect of the 3D DEM.

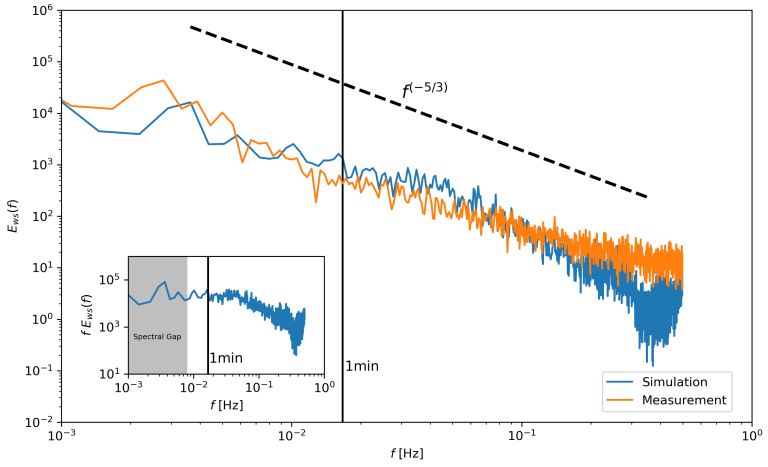
A novelty of these CFD simulations rests in the use of facility-specific wind measurements collected by an on-site anemometer to reconstruct the horizontally-averaged boundary layer profile. This in turn is employed to interpolate the horizontal forcing terms needed at each time step on the right-hand side of the u and v components of the momentum equation and the buoyancy transport equation. This approach bears some resemblance to the method used by (Wiersema et al., 2020) for their forced microscale simulations. The on-site measurements by the anemometer are combined with the local estimate of the Obhukov lengthscale L provided by the High-Resolution Rapid Refresh (HRRR) atmospheric forecasting model (Dowell et al., 2022). Vertically varying mean velocity and thermal buoyancy profiles mapped to the parametric equations of the Monin-Obhukov similarity theory (MOST) are constructed using L and onsite measurements referenced earlier at the sensor height as inputs. The MOST matching profiles that are used to infer the forcing term are periodically updated during the course of the simulation using time-evolving onsite information. This allows the simulated boundary layer to respond to changing on-site conditions while still retaining the resolved component acquired from solving the filtered momentum and buoyancy transport equations.

The time traces of the instantaneous and minute-averaged horizontal wind speed ($WS = \sqrt{U^2 + V^2}$) and wind direction ($WD = 270^\circ - tan^{-1}(V/U)$) on the virtual north-east (NE) sensor from simulations of two different cases are compared with onsite measurements in Figures S1b and S1c. It is evident that the simulations are able to faithfully reproduce the varying wind field as measured by onsite monitors. The wall-normal variation of the mean wind speed for three selected test cases is presented in Figure S2a. The three cases are chosen such that each of them falls within the stable (red), neutral (blue) and unstable (green) regimes based on extracted L from the HRRR model. The gray profiles represent several instantaneous transects extracted from the simulations. The qualitative trend of the vertical variation in mean wind speed reproduced by the simulations under different atmospheric stability regimes compares favorably with the results reported by (Bre, 2018).

To further demonstrate that the CFD simulations are indeed reproducing eddies representative of the expected turbulence cascade, Figure S2b compares the mean one-dimensional power spectrum E_{WS} of the horizontal wind speed for the neutral test case presented in Figure S2a. Since our simulations are performed using a constant time step of $\Delta t = 0.1s$, the WS at the virtual sensor locations is averaged up to the resolution of $1Hz$ before computing the power spectrum to allow a one-to-one comparison with the power spectrum computed via onsite measurements. Figure S2b identifies a five-order reduction in energy across the range of frequencies captured in this simulation case. The presence of the turbulence cascade with a distinct inertial subrange of scales was identified using Kolmogorov's -5/3 power law (Pope, 2000). The inset of Figure S2b presents the premultiplied power spectrum fE_{WS} for the same simulation. There is a decent qualitative correspondence with the fE_{WS} reported by (Schalkwijk et al., 2014) from their year-long LES of an ABL. Specifically, the range of frequencies that can be classified as being part of the so-called spectral gap extends below the 1-minute level, which is typically above the general upper limit of frequencies typically assumed for the validity spectral gap. Additionally, there is a precipitous drop in spectral



(a)



(b)

Figure S2. *(a)* Variation of the mean wind speed in the vertical z direction for three representative cases from the CSU METEC study dataset: stable (red), neutral (blue) and unstable (green). Gray lines indicate instantaneous traces for each of the three cases. *(b)* Power spectrum profiles of the horizontal wind speed from the simulation for the neutral case in *(a)* compared against spectral profiles of the measured on-field data at 2m height. The inset shows the premultiplied power spectrum for this simulation case.

power for the range of scales above 0.05Hz, which is in agreement with the simulation by (Schalkwijk et al., 2014). These
315 features in the CFD simulations give confidence that the more salient pieces of flow physics relevant to the turbulent dispersion

of gas are being captured numerically.

S4 Inversion Frameworks

S4.1 Least-Squares Optimization

The commonly used method of least-squares optimization (LSQ) seeks to minimize the sum of the squared differences between observed and predicted values (residuals). In this application, using the sensitivity matrix $\mathbf{S}$ and the measured concentrations $\boldsymbol{b}$, this optimization problem is formulated as:

$$\min_{\boldsymbol{Q}} \sum_{i}^{N} (\mathbf{S}\boldsymbol{Q} - \boldsymbol{b})^2 : Q \in \mathbb{R}+ \tag{S12}$$

where there are N measurements and Q cannot be negative. In this context, the system is consistently over-determined, with the sensitivity matrix $\mathbf{S}$ having more linearly independent rows (measurements) than columns (sources). The physical interpretation of this system being overdetermined is that there is sufficient sensor coverage, wind variability, and duration of experiment to provide enough information to enable robust emission rate estimation for every source.

However, under unfavorable atmospheric conditions (such as a sustained wind direction with low variability due to horizontal turbulence during an emission event), linear independence is not guaranteed. In such cases, even with a large number of measurement data rows, the lack of linear independence may lead to an under-determined system. Furthermore, there may be cases when the wind never carries emissions from a specific source, i, towards a sensor. In this case, the ith column of $\mathbf{S}$ will entirely consist of zeros. As a result, any arbitrary rate for source i would equally satisfy the measurement data.

The data collected for this study exhibits sufficient wind variability, high enough sensor density, and long enough event duration to ensure that the system is always overdetermined. For example, a 30-minute experiment (the shortest duration in this study) that is monitored by 10 sensors generates 300 sensor minutes of data ($\mathbf{S}$ has 300 rows). In an extremely unlikely scenario with minimal wind variability, 95% of the rows of $\mathbf{S}$ may not be linearly independent, and the rank $\mathbf{S}$ is still 15, substantially larger than the five sources. Future work will explore the required information density and distribution to make robust emission rate estimates for CMS deployments that more closely resemble real-world conditions, encompassing sensor density, placement, and source count.

In practice, adding a degree of regularization that encourages sparsity in the inferred rate vector can significantly improve the detection and localization statistics. It will also prevent unrealistically high emission rates by requiring stronger evidence (in the

goodness of fit) before assigning a high rate. For this purpose, a standard Lasso regression is implemented which incorporates an L1 norm penalty on the rate vector in the optimization problem:

$$\min_{\boldsymbol{Q}} \sum_{i=1}^{N} (\mathbf{S}\boldsymbol{Q} - \boldsymbol{b})^2 + \sum_{j=1}^{M} \alpha |Q_j| : Q \in \mathbb{R} + . \tag{S13}$$

This approach introduces a hyperparameter α that penalizes nonzero rate entries with a magnitude proportional to the rate. Consequently, the solver prioritizes sparser solutions, assigning high-rate values only when they lead to a significant improvement in data fitting, thereby avoiding unphysically high rate estimates.

S4.2 Markov-Chain Monte-Carlo

While a computationally complex and expensive approach to a linear optimization problem may seem unnecessary, it offers distinct benefits for specific applications, warranting its inclusion. In the context of the current study, this approach offers three key advantages. Unlike simpler optimizers like LSQ which is constrained to only minimizing the sum of squared residuals, the use of various error evaluation functions is permitted in this approach. Second, this method offers greater control over the use of prior information in the context of a rate inference. This enables more explicit promotion of sparsity, rate regularization, and even the incorporation of operational or independently gathered data to further inform estimates. For instance, SCADA data combined with emissions factors can generate time-dependent priors, favoring higher rates when a known equipment piece is in operating mode. Finally, a full posterior distribution approximation is crucial for continuous estimators. These estimators compute rates as a function of time on streaming data. They recursively take in the last known rate estimate and combine it with new incoming data to make an updated rate estimate. The posterior distribution of a given timestep can be used as the prior information in the subsequent timestep. This approach enhances emission rate estimation accuracy during periods of limited information. Prior information can be used to propagate source-specific rates through these periods until the proper signal is delivered from the given source to the system (the CMS observes a given source). A full description of a continuous state estimator employing a recursive Bayesian framework is beyond of the scope of this paper and will be addressed in future work.

In the context of a Bayesian parameter inference problem, Markov-Chain Monte-Carlo technique (MCMC) constructs a Markov chain by drawing samples from a probability distribution and performing acceptance/rejection sampling. This chain samples the posterior distribution of the model parameters with a frequency that is proportional to the posterior probability. The objective is to estimate the posterior probability density function $P(D|Q)$ of rate vector Q, given some input data D. This can be achieved by Bayes' theorem:

$$P(Q|D) = \frac{P(D|Q)P(Q)}{P(D)} \tag{S14}$$

where $P(D|Q)$ is the likelihood, representing how well the data fits a given set of model parameters, $P(Q)$ is the prior, and $P(D)$ is the evidence. Note that $P(D)$ is inconsequential in the context of the existing problem, because this calculation only considers relative probabilities between different Q when performing acceptance sampling, and $P(D)$ does not depend on Q, so it cancels out. As evident in Equation S14, the likelihood function and the prior must be defined in order to run this calculation. An appropriate prior to promote sparsity is the spike and slab prior, with a higher probability (the spike) at zero rate. The distribution of nonzero rates can be specified as $\beta(Q_i)$:

$$P(\boldsymbol{Q}) = \prod_i S(Q_i) \tag{S15}$$

$$S(Q_i) = \begin{cases} H & \text{if } Q_i = 0 \\ \beta(Q_i) & \text{if } Q_i > 0 \end{cases} \tag{S16}$$

Here, H represents the height of the spike. The relative magnitude of H compared to the characteristic value of $\beta(Q_i)$, defines the strength of the sparsity bias in the rate inference. The function $\beta(Q_i)$ can be defined using any other prior information available. For the present study, a constant value of $\beta(Q_i) = 1$ is employed, and H is set to 5. In other words, the algorithm favors zero rates over nonzero rates by a factor of 5, without imposing any additional prior on the nonzero distribution. A likelihood function is then defined using a standard approach of employing the chi-squared statistic via:

$$P(D|Q) = \exp\left[-\chi^2/2\right] \tag{S17}$$

where

$$\chi^2(\mathbf{S}, \boldsymbol{Q}, \boldsymbol{b}) = \sum_i^N \frac{(\mathbf{S}\boldsymbol{Q} - \boldsymbol{b})^2}{\sigma_i^2}. \tag{S18}$$

Here, $\mathbf{S}$, $\boldsymbol{Q}$, and b represent the sensitivity matrix, rate vector, and measurement vector, respectively and σ_i denotes the uncertainty associated with each measurement-prediction combination. In principle, σ_i can be a function of the particular dispersion model's error characteristics under certain conditions (i.e., a function of wind speed, stability class, etc.)

A range of sampling schemes with various complexity levels be can employed for approximating $P(Q|D)$. Examples include Metropolis-Hastings, Gibbs, and Hamiltonian (see Dunson and Johndrow (2019) for a recent review and history of MCMC sampling schemes). In this study, we implement the basic Metropolis-Hasting algorithm. A concise description is outlined here for clarity.

First, a random rate vector is chosen as an initial starting point. Next, a random step is taken in the rate vector. The defined likelihood and prior functions are applied to both the old $\boldsymbol{Q}$ and the new $\boldsymbol{Q}$ to compute the ratio of new to old posterior probability. Then, a random number between 0 and 1 is generated. If this value is less than the ratio of the new to the old posterior probabilities, the new sample is accepted and added to the chain. Otherwise, the sample is rejected and the old rate estimate is retained as the most recent sample. This iterative process continues until pre-defined stopping criteria are met, which may include a maximum number of iterations, convergence of the distribution, or other criteria.

The result of this calculation is a "chain" of rate vectors that can then be analyzed further for the given purpose. Means or medians across marginalized distributions can be determined to infer the most common or best-fit rate values for each source. In addition, the standard deviations or inner-quartile ranges of the marginalized distributions can be calculated as a measure of uncertainty per source, and the covariance between different source rate estimates can be computed to gain a deeper understanding of the underlying distribution. For example, closely located sources may demonstrate a negative covariance. This indicates that the algorithm recognizes the need for rate allocation to one of the sources, but struggles to differentiate them. In this example, an equivalent data fit can be achieved by distributing the rate arbitrarily between these two sources.

Further processing of the rate estimates based on the covariances between sources can be applied to promote sparsity or to highlight potential instances of source confusion or mislocalization. It can effectively act as a quality control measure to flag rate inferences that may be error-prone. In this work, the best-fit rate from a Markov chain of rate vectors is inferred by computing the median of each element of $\boldsymbol{Q}$. If the median is 0 (more than 50% of the samples for a given source are at 0 rate), that source rate is set to 0. If the median value is nonzero, then the zero-rate samples are excised and the median of the "slab" portion of the distribution is recomputed as the best-fit rate.

S5 Evaluation of Forward Model Accuracy

In general, it is expected that the more accurate forward model for concentration predictions should result in the best quantification estimates. This Appendix tests whether this is indeed the case. To directly compare the accuracy of the three dispersion modeling techniques (Plume, Puff, CFD), each forward model is applied to actual release rates and locations. Predicted and

measured concentrations are then compared using several error metrics. A random set of 60 experiments are chosen for this analysis. This limited subset is used due to the high computational cost of running CFD against all 347 experiments.

S5.1 Metrics

In order to compare the predicted concentration estimates against the measurement data, we apply the following performance metrics suggested by Chang and Hanna (2004) for this exact purpose. More specifically, the normalized mean squared error (NMSE), fractional bias (FB), geometric mean bias (MG), geometric variance (VG), normalized standard deviation (NSD), and the fraction of predictions within a factor of x (Fac$x$). Mathematically, these are expressed as

$$NMSE = \frac{\overline{(X_o - X_p)^2}}{\overline{X_o}\,\overline{X_p}}, \tag{S19}$$

$$FB = \frac{2(\overline{X_o} - \overline{X_p})}{(\overline{X_o} + \overline{X_p})}, \tag{S20}$$

$$MG = \exp(\overline{\ln X_o} - \overline{\ln X_p}), \tag{S21}$$

$$VG = \exp(\overline{(\ln X_o - \ln X_p)^2}), \tag{S22}$$

$$NSD = \frac{\sigma_p}{\sigma_o}. \tag{S23}$$

In these equations, X_o represents the observed concentration, X_p refers to the predicted concentrations (computed via $\mathbf{S}\boldsymbol{Q'}$, where $\mathbf{S}$ is the sensitivity matrix associated with a given model and $\boldsymbol{Q'}$ is the actual rate vector), and σ is the standard deviation of either the observed of predicted values. Fac$x$ is simply the fraction of predicted values that are within a factor of x of the measurements. In order to compensate for the fact that a few outliers can significantly impact the NMSE and FB, MG and VG are introduced using logarithmically-scaled data to offer error statistics that are less sensitive to a small number of outliers. For the purposes of computing MG and VG, measurement/prediction concentration pairs where at least one of the values is less than 1 ppm are removed from the calculation. This prevents contamination of mean MG and VG by predictions/measurements by sensors upwind of the source or under conditions when the predictions/measurements are near the instrument sensitivity. In other words, the MG and VG statistics are computed only on prediction/measurement pairs for which both "agree" that the sensor in question is receiving an appreciable amount of signal from a source. This helps balance the statistics from being skewed by an overabundance of $X_o = 0, X_p = 0$ cases, which are the majority of sensor readings.

S5.2 Results

The results of applying all three forward models to the same subset of 60 experiments and computing the metrics defined in S5.1 are shown in Figure S3.

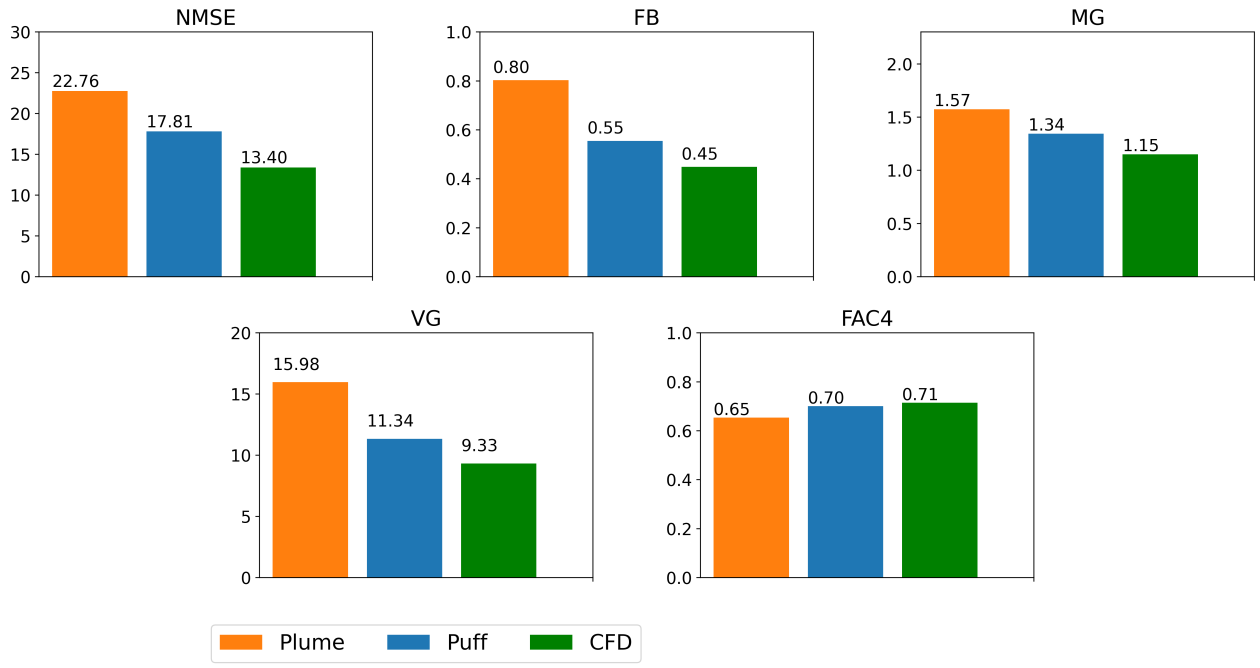


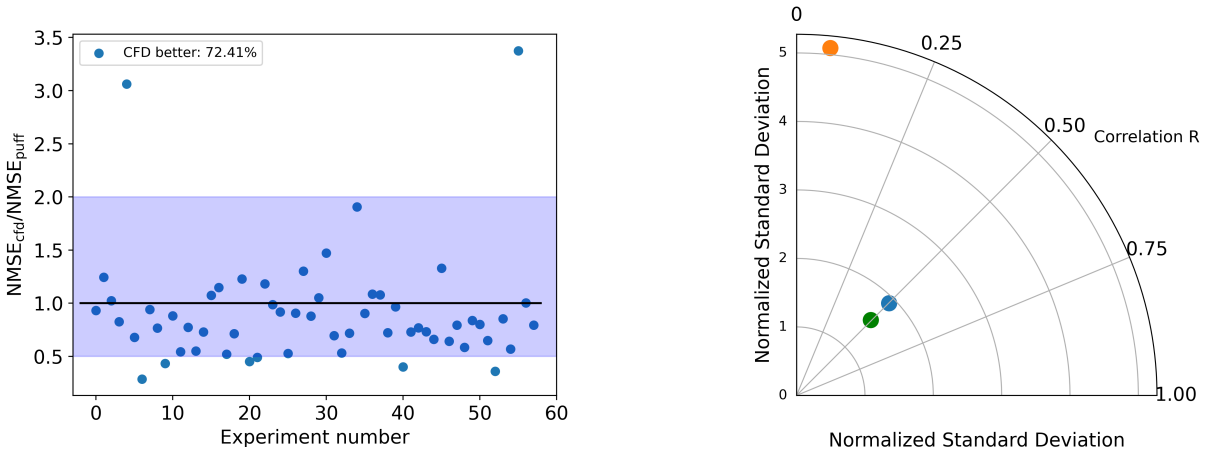
Figure S3. Comparison of the mean measures of the model skill for the Gaussian plume (orange), Gaussian puff (blue) and CFD (green) techniques for the selected 60 experiments.

It is evident considering the results in Figure S3 that increasingly complex dispersion modeling (Plume, Puff, CFD, in that order) results in uniformly and monotonically improving error metrics. In other words, it is the case for each of these 5 metrics that the Puff outperforms Plume, but is outperformed by CFD. More specifically, the NMSE for the CFD model shows an improvement of 45% and 25% over the plume and puff models, respectively. To further illustrate this, Figure S4a shows the ratio of $NMSE_{cfd}/NMSE_{puff}$ for the 60 experiments, and demonstrates that the CFD outperforms the Puff model in terms of NMSE for 72% of cases and a lower MG in 63% and 86%, of the experiments when compared to the puff and plume models, respectively. A large ($>> 1$) NMSE indicates that the associated error distribution is more likely to be lognormal than normal (Chang and Hanna, 2004)).The higher values of geometric variance (VG) for the plume model indicate that the lognormal distribution of the error has comparatively much larger scatter than both puff and CFD, which show much more comparable variance in error, as is evident in Figure S4a, showing that 95% of the experiments have $\mathrm{NMSE_{cfd}/NMSE_{puff}}$ within a factor of 2.

Both fractional bias (FB) and mean geometric bias (MG), are indicators of systematic error. Considering these metrics across all dispersion models, we see that the models generally underestimate the measured concentrations, although the magnitude of the bias decreases with increasing complexity of the forward model. The systematic underestimation of concentrations across all three approaches indicates a combination of two potential underlying issues: first, that the simulated plumes are overly diffusive, resulting in smaller than actual concentrations. Second, it is also possible that there are more complex wind dynamics that are not being properly captured by any of the dispersion models which could result in concentration enhancements on sensors that the dispersion models think are being "missed" by the plume. Either one of these effects could plausibly cause a systematic underestimation of concentrations. With this said, it is evident that the CFD is more properly capturing salient features in the plume physics that result in substantially better concentration predictions, however there is still significant room for improvement.

In order to further understand the predictive capabilities of the CFD dispersion model under different atmospheric stability classes, we stratified the experiments by stability class and computed the error metrics in each ASC. In general, we found that CFD consistently outperforms the puff model for experiments falling in the unstable and neutral ASC regimes. The relative performance of the CFD is the poorest for the stable regime: CFD only outperformed the Puff in 54% of these cases. It is worth noting that while the dispersion profiles in the plume and puff models are determined by the ASC alone, the structures produced by CFD simulations are far more complex and depend on interactions between particular details of the modeling approach including the local grid resolution, the dynamic balance between thermal and momentum forcing, and potentially many other factors. As such, it is difficult to pin down the exact physical reasons why the CFD modeling results in such an improvement.

Of all the metrics shown in Figure S3, FAC4 shows the lowest progressive improvement from plume to puff to CFD; in fact, the FAC4 value for the CFD (0.71) is nearly identical to that for puff (0.70). This is due to the well recognized challenge of combining observed and simulated plumes both in time and space. As Weil et al. (1992) note, variations in wind direction of 20-40 degrees may result in complete failure of overlap of simulated and observed plume despite manifesting similar magnitudes and shapes. Nevertheless, the current FAC4 results show a favorable comparison with the FAC2 and FAC5 numbers reported by Wiersema et al. (2020) from their microscale and multiscale CFD simulations of the Join Urban 2003 experiment, which modeled the single-point release of a tracer in an urban environment. The relative comparison across the following three performance metrics: NSD, correlation coefficient R and normalized root mean square error (NRMSE), which are related by the law of cosines, can be performed using the single diagram method of Taylor (2001). An ideal case for these metrics would be located on the horizontal axis at the origin of the polar plot shown in Figure S4b. For the current experiments, it is clear that both puff and CFD vastly outperform the plume model, the CFD slightly edging out puff with near identical correlation coefficient R but an appreciably lower NSD.



(a) (b)

Figure S4. *(a)* $NMSE_{cfd}/NMSE_{puff}$ from the experimental dataset for CFD. The shaded region indicates the range of cases within a factor of 2 of the scenario with NMSE=1. *(b)* Visualizing the model performance metrics via the single nomogram method proposed by Taylor (2001). Each dot corresponds to the mean value across all experiments used in the CFD dataset. Colors in *(b)* follow the labeling identified in figure S3.

S6 Relative Error Distribution Associated With Quantification Methods

Figure S5 shows box-and-whisker plots of the relative error distribution associated with each quantification method. In this Figure, the orange lines show the median of the relative error distribution while the blue-shaded region shows the two inner quartile regions (IQR) of the distribution, encompassing the middle 50% of the data. The whiskers indicate the first and third quartiles, extending to 1.5 times the IQR, and the circles denote data points outside this range, representing outliers. Several key features are apparent in this figure. First, the Plume-based quantification methods (first and third columns) show a tendency toward lower median relative errors suggesting a more pronounced low bias compared to Puff-based models. This observation is consistent with the $\overline{E}$ and ΔC statistics represented in the main manuscript. Second, the IQR narrows when transitioning from the LSQ inversion methods to MCMC, indicating less scatter in the relative error distributions for the MCMC inversion method. This trend is also reflected in the $F2$ statistic from the main manuscript. Finally, the combination of Puff MCMC has a smaller upper range of percent errors, with the maximum relative error being around 300%, compared to all of the other methods that have outlying relative error points above 400%.

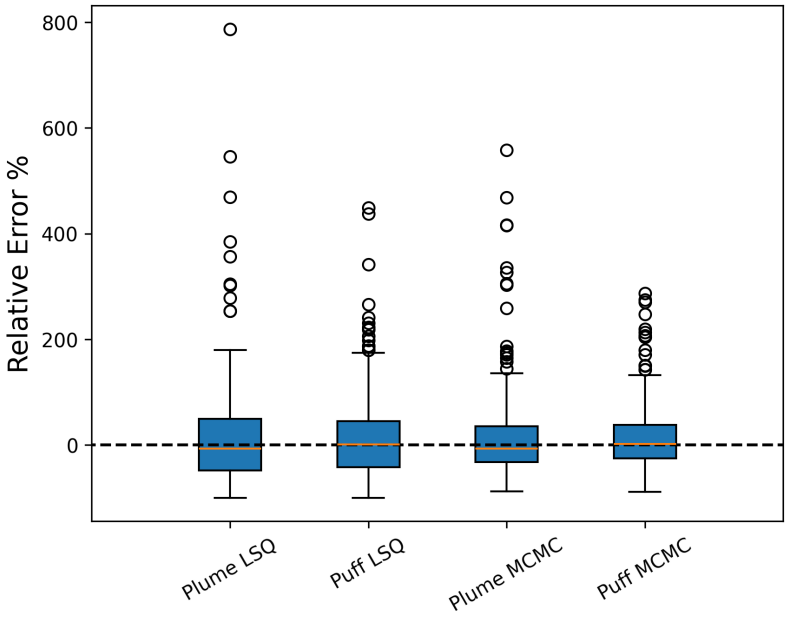


Figure S5. Box and whisker plots of relative error distributions for 4 different quantification methods.

References

Monin-Obukhov similarity theory and its application to wind flow modelling over complex terrain, J WIND ENG IND AEROD, 182, 308–321, https://doi.org/10.1016/j.jweia.2018.09.026, 2018.

Alvarez, R. A., Zavala-Araiza, D., Lyon, D. R., Allen, D. T., Barkley, Z. R., Brandt, A. R., Davis, K. J., Herndon, S. C., Jacob, D. J., Karion, A., Kort, E. A., Lamb, B. K., Lauvaux, T., Maasakkers, J. D., Marchese, A. J., Omara, M., Pacala, S. W., Peischl, J., Robinson, A. L., Shepson, P. B., Sweeney, C., Townsend-Small, A., Wofsy, S. C., and Hamburg, S. P.: Assessment of methane emissions from the U.S. oil and gas supply chain, Science, 361, 186–188, https://doi.org/10.1126/science.aar7204, 2018.

Carruthers, D., Holroyd, R., Hunt, J., Weng, W., Robins, A., Apsley, D., Thompson, D., and Smith, F.: UK-ADMS: A new approach to modelling dispersion in the earth's atmospheric boundary layer, J WIND ENG IND AEROD, 52, 139–153, https://doi.org/https://doi.org/10.1016/0167-6105(94)90044-2, 1994.

Carruthers, D., Weng, W., Hunt, J., Holroyd, R., McHugh, C., and Dyster, S.: Plume/puff spread and mean concentration module specifications, 3, 2009.

Chang, J. C. and Hanna, S. R.: Air quality model performance evaluation, METEOROL ATMOS PHYS, 87, 167–196, https://doi.org/10.1007/s00703-003-0070-7, 2004.

Daniels, W. S., Jia, M., and Hammerling, D. M.: Detection, localization, and quantification of single-source methane emissions on oil and gas production sites using point-in-space continuous monitoring systems, Elementa, 12, https://doi.org/10.1525/elementa.2023.00110, 2024.

Dowell, D. C., Alexander, C. R., James, E. P., Weygandt, S. S., Benjamin, S. G., Manikin, G. S., Blake, B. T., Brown, J. M., Olson, J. B., Hu, M., Smirnova, T. G., Ladwig, T., Kenyon, J. S., Ahmadov, R., Turner, D. D., Duda, J. D., and Alcott, T. I.: The High-Resolution Rapid

Refresh (HRRR): An Hourly Updating Convection-Allowing Forecast Model. Part I: Motivation and System Description, WEATHER FORECAST, 37, 1371 – 1395, https://doi.org/10.1175/WAF-D-21-0151.1, 2022.

Dunson, D. B. and Johndrow, J. E.: The Hastings algorithm at fifty, BIOMETRIKA, 107, 1–23, https://doi.org/10.1093/biomet/asz066, 2019.

Houweling, S., Kaminski, T., Dentener, F., Lelieveld, J., and Heimann, M.: Inverse modeling of methane sources and sinks using the adjoint of a global transport model, J GEOPHYS RES-ATMOS, 104, 26 137–26 160, https://doi.org/10.1029/1999JD900428, 1999.

Ismail, U.: Direct Numerical Simulation of a Turbulent Boundary Layer Encountering a Smooth-to-Rough Step Change, ENERGIES, 16, https://doi.org/10.3390/en16041709, 2023.

Jia, M., Daniels, W., and Hammerling, D.: Comparison of the Gaussian plume and puff atmospheric dispersion models for methane modeling on oil and gas sites, ChemRxiv, https://doi.org/10.26434/chemrxiv-2023-hc95q-v2, 2023.

Jia, M., Fish, R., Daniels, W. S., Sprinkle, B., and Hammerling, D.: A fast and lightweight implementation of the Gaussian puff model for near-field atmospheric transport of trace gasses, SCI REP-UK, 15, 18 710, https://doi.org/10.1038/s41598-025-99491-x, 2025.

Johnson, M. R., Tyner, D. R., Conley, S., Schwietzke, S., and Zavala-Araiza, D.: Comparisons of Airborne Measurements and Inventory Estimates of Methane Emissions in the Alberta Upstream Oil and Gas Sector, ACS ES&T, 51, 13 008–13 017, https://doi.org/10.1021/acs.est.7b03525, pMID: 29039181, 2017.

Kim, J., Kim, D., and Choi, H.: An Immersed-Boundary Finite-Volume Method for Simulations of Flow in Complex Geometries, J COMPUT PHYS, 171, 132–150, https://doi.org/10.1006/jcph.2001.6778, 2001.

LEONARDI, S., ORLANDI, P., SMALLEY, R. J., DJENIDI, L., and ANTONIA, R. A.: Direct numerical simulations of turbulent channel flow with transverse square bars on one wall, J FLUID MECH, 491, 229–238, https://doi.org/10.1017/S0022112003005500, 2003.

Lilly, D. K.: A proposed modification of the Germano subgrid-scale closure method, PHYS FLUIDS A-FLUID, 4, 633–635, https://doi.org/10.1063/1.858280, 1992.

Lu, X., Jacob, D. J., Wang, H., Maasakkers, J. D., Zhang, Y., Scarpelli, T. R., Shen, L., Qu, Z., Sulprizio, M. P., Nesser, H., Bloom, A. A., Ma, S., Worden, J. R., Fan, S., Parker, R. J., Boesch, H., Gautam, R., Gordon, D., Moran, M. D., Reuland, F., Villasana, C. A. O., and Andrews, A.: Methane emissions in the United States, Canada, and Mexico: evaluation of national methane emission inventories and 2010–2017 sectoral trends by inverse analysis of in situ (GLOBALVIEWplus CH4 ObsPack) and satellite (GOSAT) atmospheric observations, ATMOS CHEM PHYS, 22, 395–418, https://doi.org/10.5194/acp-22-395-2022, 2022.

Lu, X., Jacob, D. J., Zhang, Y., Shen, L., Sulprizio, M. P., Maasakkers, J. D., Varon, D. J., Qu, Z., Chen, Z., Hmiel, B., Parker, R. J., Boesch, H., Wang, H., He, C., and Fan, S.: Observation-derived 2010-2019 trends in methane emissions and intensities from US oil and gas fields tied to activity metrics, P NATL ACAD SCI USA, 120, e2217900 120, https://doi.org/10.1073/pnas.2217900120, 2023.

Maasakkers, J. D., Jacob, D. J., Sulprizio, M. P., Scarpelli, T. R., Nesser, H., Sheng, J., Zhang, Y., Lu, X., Bloom, A. A., Bowman, K. W., Worden, J. R., and Parker, R. J.: 2010–2015 North American methane emissions, sectoral contributions, and trends: a high-resolution inversion of GOSAT observations of atmospheric methane, ATMOS CHEM PHYS, 21, 4339–4356, https://doi.org/10.5194/acp-21-4339-2021, 2021.

Martin, D. O.: Comment On"The Change of Concentration Standard Deviations with Distance", JAPCA J AIR WASTE MA, 26, 145–147, https://doi.org/10.1080/00022470.1976.10470238, 1976.

Mohan, M. and Siddiqui, T.: Analysis of various schemes for the estimation of atmospheric stability classification, ATMOS ENVIRON, 32, 3775–3781, https://doi.org/https://doi.org/10.1016/S1352-2310(98)00109-5, 1998.

Mukha, T., Rezaeiravesh, S., and Liefvendahl, M.: A library for wall-modelled large-eddy simulation based on OpenFOAM technology, COMPUT PHYS COMMUN, 239, 204–224, https://doi.org/10.1016/j.cpc.2019.01.016, 2019.

Omara, M., Himmelberger, A., MacKay, K., Williams, J. P., Benmergui, J., Sargent, M., Wofsy, S. C., and Gautam, R.: Constructing a measurement-based spatially explicit inventory of US oil and gas methane emissions, EARTH SYST SCI DATA, 2024, 1–25, https://doi.org/10.5194/essd-2024-72, 2024.

Pasquill, F.: The estimation of the dispersion of windborne material, METEOROL MAG, 90, 20–49, 1961.

Pierce, C. D.: Progress-variable approach for large-eddy simulation of turbulent combustion, Ph.D. thesis, Stanford University, California, 2001.

Plant, G., Kort, E. A., Brandt, A. R., Chen, Y., Fordice, G., Negron, A. M. G., Schwietzke, S., Smith, M., and Zavala-Araiza, D.: Inefficient and unlit natural gas flares both emit large quantities of methane, Science, 377, 1566–1571, https://doi.org/10.1126/science.abq0385, 2022.

Pope, S. B.: Turbulent Flows, Cambridge University Press, 2000.

Qian, W. and Venkatram, A.: Performance of Steady-State Dispersion Models Under Low Wind-Speed Conditions, BOUND LAYER METEOROL, 138, 475–491, https://doi.org/10.1007/s10546-010-9565-1, 2011.

R. P. Hosker, J.: A Comparison of Estimation Procedures for Over-Water Plume Dispersion, in: Symposium on Atmospheric Diffusion and Air Pollution held September 9-13, 1974 in Santa Barbara, California., vol. 60, pp. 669–669, ISSN 0096-3941, https://doi.org/10.1029/eo060i038p00669-03, 1974.

Randerson, D.: Atmospheric science and power production, Tech. rep., USDOE Technical Information Center, Oak Ridge, TN, https://doi.org/10.2172/6503687, 1984.

Ravikumar, A. P., Wang, J., McGuire, M., Bell, C. S., Zimmerle, D., and Brandt, A. R.: "Good versus good enough?" Empirical tests of methane leak detection sensitivity of a commercial infrared camera, ACS ES&T, 52, 2368–2374, https://doi.org/10.1021/acs.est.7b04945, 2018.

Ravikumar, A. P., Li, H., Yang, S. L., and Smith, M. L.: Developing Measurement-Informed Methane Emissions Inventory Estimates at Midstream Compressor Stations, ACS ES&T Air, 2, 358–367, https://doi.org/10.1021/acsestair.4c00237, 2025.

Ražnjević, A., van Heerwaarden, C., and Krol, M.: Evaluation of two common source estimation measurement strategies using large-eddy simulation of plume dispersion under neutral atmospheric conditions, ATMOS MEAS TECH, 15, 3611–3628, https://doi.org/10.5194/amt-15-3611-2022, 2022.

Robertson, A. M., Edie, R., Field, R. A., Lyon, D., McVay, R., Omara, M., Zavala-Araiza, D., and Murphy, S. M.: New Mexico Permian Basin Measured Well Pad Methane Emissions Are a Factor of 5–9 Times Higher Than U.S. EPA Estimates, ACS ES&T, 54, 13 926–13 934, https://doi.org/10.1021/acs.est.0c02927, pMID: 33058723, 2020.

Santos, A., Mollel, W., Duggan, G. P., Hodshire, A., Vora, P., and Zimmerle, D.: Using Measurement-Informed Inventory to Assess Emissions in the Denver-Julesburg Basin, ACS ES&T Air, 0, null, https://doi.org/10.1021/acsestair.5c00089, 0.

Schalkwijk, J., Jonker, H., and Siebesma, P.: Turbulent Scales in the Boundary Layer : A Year-Long Large-Eddy Simulation YOGA : A Year-Long Large-Eddy Simulation, 2014.

Shen, L., Gautam, R., Omara, M., Zavala-Araiza, D., Maasakkers, J. D., Scarpelli, T. R., Lorente, A., Lyon, D., Sheng, J., Varon, D. J., Nesser, H., Qu, Z., Lu, X., Sulprizio, M. P., Hamburg, S. P., and Jacob, D. J.: Satellite quantification of oil and natural gas methane emissions in the US and Canada including contributions from individual basins, ATMOS CHEM PHYS, 22, 11 203–11 215, https://doi.org/10.5194/acp-22-11203-2022, 2022.

Sherwin, E., Rutherford, J., Zhang, Z., Chen, Y., Wetherley, E., Yakovlev, P., Berman, E., Jones, B., Thorpe, A., Ayasse, A., Duren, R., Brandt, A., and Cusworth, D.: Quantifying oil and natural gas system emissions using one million aerial site measurements, RES SQ, https://doi.org/10.21203/rs.3.rs-2406848/v1, 2022.

Sherwin, E. D., Rutherford, J. S., Chen, Y., Aminfard, 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.

Taylor, K. E.: Summarizing multiple aspects of model performance in a single diagram, J GEOPHYS RES-ATMOS, 106, 7183–7192, https://doi.org/https://doi.org/10.1029/2000JD900719, 2001.

Tibrewal, K., Ciais, P., Saunois, M., Martinez, A., Lin, X., Thanwerdas, J., Deng, Z., Chevallier, F., Giron, C., Albergel, C., Tanaka, K., Patra, P., Tsuruta, A., Zheng, B., Belikov, D., Niwa, Y., Janardanan, R., Maksyutov, S., Segers, A., Tzompa-Sosa, Z. A., Bousquet, P., and Sciare, J.: Assessment of methane emissions from oil, gas and coal sectors across inventories and atmospheric inversions, COMMUN EARTH ENVIRON, 5, 26, https://doi.org/10.1038/s43247-023-01190-w, 2024.

USEPA: Meteorological Monitoring Guidance for Regulatory Modeling Applications, Tech. rep., ISBN 1428901949, http://www.epa.gov/scram001/guidance/met/mmgrma.pdf, 2000.

Vallejo, V., Nguyen, Q., and Ravikumar, A. P.: Geospatial variation in carbon accounting of hydrogen production and implications for the US Inflation Reduction Act, NAT ENERGY, 9, 1571–1582, https://doi.org/10.1038/s41560-024-01653-0, 2024.

Van Heerwaarden, C. C., Van Stratum, B. J., Heus, T., Gibbs, J. A., Fedorovich, E., and Mellado, J. P.: MicroHH 1.0: A computational fluid dynamics code for direct numerical simulation and large-eddy simulation of atmospheric boundary layer flows, GEOSCI MODEL DEV, 10, 3145–3165, https://doi.org/10.5194/gmd-10-3145-2017, 2017.

Venkatram, A.: An examination of the Pasquill-Gifford-Turner dispersion scheme, ATMOS ENVIRON, 30, 1283–1290, https://doi.org/10.1016/1352-2310(95)00367-3, 1996.

Wang, J. L., Daniels, W. S., Hammerling, D. M., Harrison, M., Burmaster, K., George, F. C., and Ravikumar, A. P.: Multiscale Methane Measurements at Oil and Gas Facilities Reveal Necessary Frameworks for Improved Emissions Accounting, ACS ES&T, 56, 14 743–14 752, https://doi.org/10.1021/acs.est.2c06211, pMID: 36201663, 2022.

Wayland, R.: Clarification of Regulatory Status of CALPUFF for Near-Field Applications, 24, 2020, 2008.

Weil, J. C., Sykes, R. I., and Venkatram, A.: Evaluating Air-Quality Models: Review and Outlook, J APPL METEOROL CLIM, 31, 1121 – 1145, https://doi.org/10.1175/1520-0450(1992)031<1121:EAQMRA>2.0.CO;2, 1992.

Wiersema, D. J., Lundquist, K. A., and Chow, F. K.: Mesoscale to microscale simulations over complex terrain with the immersed boundary method in the weather research and forecasting model, MON WEATHER REV, 148, 577–595, https://doi.org/10.1175/MWR-D-19-0071.1, 2020.

Williams, J. P., Omara, M., Himmelberger, A., Zavala-Araiza, D., MacKay, K., Benmergui, J., Sargent, M., Wofsy, S. C., Hamburg, S. P., and Gautam, R.: Small emission sources in aggregate disproportionately account for a large majority of total methane emissions from the US oil and gas sector, ATMOS CHEM PHYS, 25, 1513–1532, https://doi.org/10.5194/acp-25-1513-2025, 2025a.

Williams, J. P., Omara, M., Himmelberger, A., Zavala-Araiza, D., MacKay, K., Benmergui, J., Sargent, M., Wofsy, S. C., Hamburg, S. P., and Gautam, R.: Small emission sources in aggregate disproportionately account for a large majority of total methane emissions from the US oil and gas sector, ATMOS CHEM PHYS, 25, 1513–1532, https://doi.org/10.5194/acp-25-1513-2025, 2025b.

Worden, J. R., Cusworth, D. H., Qu, Z., Yin, Y., Zhang, Y., Bloom, A. A., Ma, S., Byrne, B. K., Scarpelli, T., Maasakkers, J. D., Crisp, D., Duren, R., and Jacob, D. J.: The 2019 methane budget and uncertainties at 1 degree resolution and each country through

Bayesian integration Of GOSAT total column methane data and a priori inventory estimates, ATMOS CHEM PHYS, 22, 6811–6841, https://doi.org/10.5194/acp-22-6811-2022, 2022.

Xia, H., Strayer, A., and Ravikumar, A. P.: The role of emission size distribution on the efficacy of new technologies to reduce methane emissions from the oil and gas sector, ACS ES&T, 58, 1088–1096, https://doi.org/10.1021/acs.est.3c05245, 2024.