Supplementary Materials A: Mathematical Theories and Derivations

1 Modeling Gas Dispersions using the Gaussian Plume Model

From the Stockie derivation [Stockie, 2011], the distribution of the plume perpendicular to the wind direction has a 2-D Gaussian form. That is, the gas concentration at location (x, y, z)[m] takes the general form:

$$c(x, y, z; \tilde{x}, \tilde{y}, \tilde{z}) = \frac{10^6}{\rho_{\text{CH4}}} \frac{s}{2\pi u |\mathbf{\Sigma}|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}\boldsymbol{\omega}^T \mathbf{\Sigma}^{-1} \boldsymbol{\omega}\right\}, \tag{1}$$

where $\rho_{\text{CH4}} \in \mathbb{R}^+$ is the density of methane [kg/m^3], the term $10^6/\rho_{\text{CH4}}$ ensures the gas concentration is in parts per million [PPM], u is the wind speed [m/s], s is the source emission rate [kg/s] for a source at location $(\tilde{x}, \tilde{y}, \tilde{z})$[m], and:

$$\begin{aligned} \boldsymbol{\omega} &= (\delta_H, \delta_V), \\ \mathbf{\Sigma} &= \begin{bmatrix} \sigma_H^2 & 0 \\ 0 & \sigma_V^2 \end{bmatrix}, \end{aligned} \tag{2}$$

with:

$$\begin{pmatrix} \delta_R \\ \delta_H \\ \delta_V \end{pmatrix} = \begin{pmatrix} \cos\theta_u & \sin\theta_u & 0 \\ -\sin\theta_u & \cos\theta_u & 0 \\ 0 & 0 & 1 \end{pmatrix} \times \left(\begin{bmatrix} x \\ y \\ z \end{bmatrix} - \begin{bmatrix} \tilde{x} \\ \tilde{y} \\ \tilde{z} \end{bmatrix} \right), \tag{3}$$

for θ_u being the wind direction [rad] and with the parametrization of σ_H and σ_V based on δ_R. Therefore, this simplifies to:

$$c(x, y, z; \tilde{x}, \tilde{y}, \tilde{z}) = \frac{10^6}{\rho_{\text{CH4}}} \frac{s}{2\pi u \sigma_H \sigma_V} \exp\left\{-\frac{1}{2}\left(\frac{\delta_H^2}{\sigma_H^2} + \frac{\delta_V^2}{\sigma_V^2}\right)\right\}. \tag{4}$$

So far we have only included horizontal and vertical offsets. However, we can extend equation (1) to account for gas reflections against surrounding boundaries. Assuming there are no horizontal boundaries, the horizontal term remains:

$$\exp\left\{-\frac{\delta_H^2}{2\sigma_H^2}\right\}. \tag{5}$$

However, the atmospheric boundary layer (ABL) with height P and the ground form vertical boundaries against which the gas reflects. Denoting the source height by H the vertical reflections are captured by:

$$\sum_{j=1}^{n_{refl}} \Bigg[\exp\Bigg\{ -\frac{1}{2}\frac{(2\lfloor (j+1)/2 \rfloor P + (-1)^j(\delta_V + H) - H)^2}{\sigma_V^2} \Bigg\} + \exp\Bigg\{ -\frac{1}{2}\frac{(2\lfloor j/2 \rfloor P + (-1)^{j-1}(\delta_V + H) + H)^2}{\sigma_V^2} \Bigg\} \Bigg]. \tag{6}$$

The first exponential term in equation (6) corresponds to the reflection against the ABL and the second exponential term to the reflection against the ground. n_{refl} denotes the number of reflections against the ABL and against the ground. Adding the vertical and horizontal offset and the vertical reflections we finally obtain the formula :

$$\begin{aligned} c(x,y,z;\tilde{x},\tilde{y},\tilde{z}) = \frac{10^6}{\rho_{\mathrm{CH4}}} \frac{s}{2\pi u \sigma_H \sigma_V} \exp\Bigg\{ -\frac{\delta_H^2}{2\sigma_H^2} \Bigg\} \times \Bigg(\exp\Bigg\{ -\frac{\delta_V^2}{2\sigma_V^2} \Bigg\} \\ + \sum_{j=1}^{n_{refl}} \Bigg[\exp\Bigg\{ -\frac{1}{2}\frac{(2\lfloor (j+1)/2 \rfloor P + (-1)^j(\delta_V + H) - H)^2}{\sigma_V^2} \Bigg\} \\ + \exp\Bigg\{ -\frac{1}{2}\frac{(2\lfloor j/2 \rfloor P + (-1)^{j-1}(\delta_V + H) + H)^2}{\sigma_V^2} \Bigg\} \Bigg] \Bigg). \end{aligned} \tag{7}$$

2 MCMC Posterior Derivations

2.1 Conditional Posterior $p(\sigma^2|\boldsymbol{d}, \boldsymbol{s}, \boldsymbol{\beta})$

The conditional posterior distribution for the variance of the measurement error, σ^2, can be written as:

$$\begin{aligned} p(\sigma^2|\boldsymbol{d}, \boldsymbol{s}, \boldsymbol{\beta}) &= \frac{p(\boldsymbol{d}|\boldsymbol{s}, \sigma^2)p(\boldsymbol{s})p(\boldsymbol{\beta})p(\sigma^2)}{\displaystyle\int p(\boldsymbol{d}|\boldsymbol{s}, \sigma^2)p(\boldsymbol{s})p(\boldsymbol{\beta})p(\sigma^2)d\sigma^2}, \\ &= \frac{p(\boldsymbol{d}|\boldsymbol{s}, \boldsymbol{\beta}, \sigma^2)p(\sigma^2)}{\displaystyle\int p(\boldsymbol{d}|\boldsymbol{s}, \boldsymbol{\beta}, \sigma^2)p(\sigma^2)d\sigma^2}. \end{aligned}$$

To make it clearer lets denote $R = \displaystyle\int p(\boldsymbol{d}|\boldsymbol{s}, \boldsymbol{\beta}, \sigma^2)p(\sigma^2)d\sigma^2$, i.e. the normalising constant.

$$\begin{aligned} p(\sigma^2|\boldsymbol{d}, \boldsymbol{s}, \boldsymbol{\beta}) &= \frac{1}{R} \times p(\boldsymbol{d}|\boldsymbol{s}, \boldsymbol{\beta}, \sigma^2)p(\sigma^2), \\ &= \frac{1}{R}\Big(2\pi\sigma^2\Big)^{-\frac{n_{obs}}{2}} \exp\Bigg\{ -\frac{1}{2\sigma^2}\sum^{n_{obs}}(\boldsymbol{d} - \boldsymbol{\beta} - A\boldsymbol{s})^2 \Bigg\} \frac{b^a}{\Gamma(a)}\Big(\frac{1}{\sigma^2}\Big)^{a+1} \exp\Bigg\{ -\frac{b}{\sigma^2} \Bigg\}. \end{aligned}$$

We can now absorb the terms that do not depend on σ^2 into R. We obtain:

$$\begin{aligned}
p(\sigma^2|\boldsymbol{d},\boldsymbol{s},\boldsymbol{\beta}) &= \frac{1}{R}\Big(\sigma^2\Big)^{-\frac{n_{obs}}{2}-a-1}\exp\left\{-\frac{1}{2\sigma^2}\sum^{n_{obs}}(\boldsymbol{d}-\boldsymbol{\beta}-A\boldsymbol{s})^2-\frac{b}{\sigma^2}\right\},\\
&= \frac{1}{R}\Big(\sigma^2\Big)^{-(\frac{n_{obs}}{2}+a)-1}\exp\left\{-\frac{1}{\sigma^2}\left(\frac{\sum^{n_{obs}}(\boldsymbol{d}-\boldsymbol{\beta}-A\boldsymbol{s})^2}{2}+b\right)\right\}.
\end{aligned}$$

This corresponds to the Inverse-Gamma distribution:

$$\sigma^2|\boldsymbol{d},\boldsymbol{s}\sim\text{Inv-Gamma}\left(\frac{n_{obs}}{2}+a\ ,\ b+\frac{\sum^{n_{obs}}(\boldsymbol{d}-\boldsymbol{\beta}-A\boldsymbol{s})^2}{2}\right).$$

2.2 Conditional Posterior $p(\boldsymbol{\beta}|\boldsymbol{d},\boldsymbol{s},\sigma^2)$

The conditional posterior distribution for the background concentration $\boldsymbol{\beta}$ can be written as:

$$\begin{aligned}
p(\boldsymbol{\beta}|\boldsymbol{d},\boldsymbol{s},\sigma^2) &= \frac{p(\boldsymbol{d}|\boldsymbol{s},\boldsymbol{\beta},\sigma^2)p(\boldsymbol{s})p(\boldsymbol{\beta})p(\sigma^2)}{\displaystyle\int p(\boldsymbol{d}|\boldsymbol{s},\boldsymbol{\beta},\sigma^2)p(\boldsymbol{s})p(\boldsymbol{\beta})p(\sigma^2)d\boldsymbol{\beta}},\\
&= \frac{p(\boldsymbol{d}|\boldsymbol{s},\boldsymbol{\beta},\sigma^2)p(\boldsymbol{\beta})}{\displaystyle\int p(\boldsymbol{d}|\boldsymbol{s},\boldsymbol{\beta},\sigma^2)p(\boldsymbol{\beta})d\boldsymbol{\beta}}.
\end{aligned}$$

To make it clearer lets denote $R=\displaystyle\int p(\boldsymbol{d}|\boldsymbol{s},\boldsymbol{\beta},\sigma^2)p(\boldsymbol{\beta})d\boldsymbol{\beta}$, i.e. the normalising constant.

$$\begin{aligned}
p(\boldsymbol{\beta}|\boldsymbol{d},\boldsymbol{s},\sigma^2) &= \frac{1}{R}\times p(\boldsymbol{d}|\boldsymbol{s},\boldsymbol{\beta},\sigma^2)p(\boldsymbol{\beta}),\\
&= \frac{1}{R}\left(2\pi\sigma^2\right)^{-\frac{n_{obs}}{2}}\exp\left\{-\frac{1}{2\sigma^2}\left(\boldsymbol{d}-\boldsymbol{\beta}-A\boldsymbol{s}\right)^2\right\}\\
&\quad\times(2\pi)^{-\frac{n_{obs}}{2}}\left|\Sigma\right|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}\left(\boldsymbol{\beta}-\boldsymbol{\mu}\right)^T\Sigma^{-1}\left(\boldsymbol{\beta}-\boldsymbol{\mu}\right)\right\},\\
&= \frac{1}{R}\exp\left\{-\frac{1}{2}\left[\boldsymbol{\beta}^T(\frac{1}{\sigma^2}\mathbb{I}+\Sigma^{-1})\boldsymbol{\beta}-2\boldsymbol{\beta}^T\left(\frac{1}{\sigma^2}(\boldsymbol{d}-A\boldsymbol{s})+\Sigma^{-1}\boldsymbol{\mu})\right)\right]\right\}.
\end{aligned}$$

This is the kernel of a multivariate Gaussian distribution, therefore:

$$\boldsymbol{\beta}|\boldsymbol{d},\boldsymbol{s},\sigma^2\sim\mathrm{N}\left(\left(\frac{1}{\sigma^2}\mathbb{I}+\Sigma^{-1}\right)^{-1}\left(\frac{1}{\sigma^2}(\boldsymbol{d}-A\boldsymbol{s})+\Sigma^{-1}\boldsymbol{\mu}\right)\ ,\ \left(\frac{1}{\sigma^2}\mathbb{I}+\Sigma^{-1}\right)^{-1}\right).$$

3 M-MALA-within-Gibbs Pseudocode

Algorithm 1: M-MALA-within-Gibbs

Input: number of iterations: L , initialize variables: $\boldsymbol{\theta}^{(0)} = [s, \tilde{x}, \tilde{y}, a_H, a_V, b_H, b_V], \sigma^2, \boldsymbol{\beta}$, data: $\boldsymbol{d}$, log-likelihood function: $\log(p(\boldsymbol{\theta}|\boldsymbol{d}))$, step-size: $\zeta^{(0)}$.

Output: Samples for: $\boldsymbol{\theta}, \sigma^2, \boldsymbol{\beta}$, log-likelihoods, acceptance rates.

for $l = 1, 2, 3, \cdots, L$ **do**

  # M-MALA

    Evaluate: $\log(p(\boldsymbol{\theta}^{(l-1)}|\boldsymbol{d})), \nabla \log(p(\boldsymbol{\theta}^{(l-1)}|\boldsymbol{d})), \nabla^2 \log(p(\boldsymbol{\theta}^{(l-1)}|\boldsymbol{d}))$, and $\boldsymbol{A}^{(l-1)}$

    Propose samples: $\boldsymbol{\theta}^* \sim \mathrm{N}_n\left(\boldsymbol{\theta}^{(l-1)} + 0.5\zeta^{(l-1)} G^{-1} \nabla \log(p(\boldsymbol{\theta}^{(l-1)}|\boldsymbol{d})), \zeta^{(l-1)} G^{-1}\right)$

    Evaluate: $\log(p(\boldsymbol{\theta}^*|\boldsymbol{d})), \nabla \log(p(\boldsymbol{\theta}^*|\boldsymbol{d})), \nabla^2 \log(p(\boldsymbol{\theta}^*|\boldsymbol{d}))$, and $\boldsymbol{A}^*$

    acceptance-prob $= \log(p(\boldsymbol{\theta}^*|\boldsymbol{d})) - \log(p(\boldsymbol{\theta}^{(l-1)}|\boldsymbol{d})) + q(\boldsymbol{\theta}^{(l-1)}|\boldsymbol{\theta}^*) - q(\boldsymbol{\theta}^*|\boldsymbol{\theta}^{(l-1)})$

    $u \sim \mathrm{Uniform}(0, 1)$

    if $\log(u) <$ *acceptance-prob* **then**

      $\boldsymbol{\theta}^{(l)} = \boldsymbol{\theta}^*$

      $\boldsymbol{A}^{(l)} = \boldsymbol{A}^*$

      sum-accept += 1

    else

      $\boldsymbol{\theta}^{(l)} = \boldsymbol{\theta}^{(l-1)}$

      $\boldsymbol{A}^{(l)} = \boldsymbol{A}^{(l-1)}$

      sum-accept += 0

    end

    $\text{acceptance-rate}^{(l)} = (0.01) \times \text{acceptance-rate}^{(l-1)} + (0.99) \times \frac{\text{sum-accept}}{l}$

    Update step-size $\zeta^{(l)}$ using $\text{acceptance-rate}^{(l)}$

  # Gibbs

    $\sigma^{2(l)} \sim \text{Inv-Gamma}\left(\frac{n_{obs}}{2} + a\ ,\ b + \frac{\sum^{n_{obs}} (\boldsymbol{d} - \boldsymbol{\beta}^{(l-1)} - \boldsymbol{A}^{(l)} \boldsymbol{s}^{(l)})^2}{2}\right)$

    $\boldsymbol{\beta}^{(l)} \sim \mathrm{N}\left(\left(\frac{1}{\sigma^{2(l)}}\mathbb{I} + \Sigma^{-1}\right)^{-1} \left(\frac{1}{\sigma^{2(l)}}(\boldsymbol{d} - \boldsymbol{A}^{(l)}\boldsymbol{s}^{(l)}) + \Sigma^{-1}\boldsymbol{\mu}\right)\ , \left(\frac{1}{\sigma^{2(l)}}\mathbb{I} + \Sigma^{-1}\right)^{-1}\right)$

  return $\boldsymbol{\theta}^{(l)}, \sigma^{2(l)}, \boldsymbol{\beta}^{(l)}$, $\text{acceptance-rate}^{(l)}, \log(p(\boldsymbol{\theta}^{(l)}|\boldsymbol{d}))$.

end

Note: The code allows to either keep the step-size ζ fixed throughout the MCMC or to tune it during burn-in to achieve an optimal acceptance rate of 70% [Girolami and Calderhead, 2011] then fixing it at the end of the burn-in period.

Additionally, using Table 4 in Supplementary Materials B.3 we define $\boldsymbol{\Sigma_s} = \sigma_s^2\mathbb{I}$, $\boldsymbol{\Sigma_{xy}} = [\sigma_x^2, \sigma_y^2]\mathbb{I}$, $\boldsymbol{\Sigma_a} = \sigma_a^2\mathbb{I}$, and $\boldsymbol{\Sigma_b} = \sigma_b^2\mathbb{I}$.

$$p(\boldsymbol{\theta} \mid \boldsymbol{d}) = \mathrm{N}\left(\boldsymbol{d} \mid \boldsymbol{As} + \boldsymbol{\beta}, \sigma^2\mathbb{I}\right) \times \left\{ \mathrm{N}\left(\boldsymbol{s} \mid \boldsymbol{\mu}_s, \boldsymbol{\Sigma}_s\right) \times \mathrm{N}\left(\begin{bmatrix}\tilde{\boldsymbol{x}} \\ \tilde{\boldsymbol{y}}\end{bmatrix} \mid \boldsymbol{\mu}_{xy}, \boldsymbol{\Sigma}_{xy}\right) \right.$$
$$\left. \times \mathrm{N}\left(\begin{bmatrix}a_H \\ a_V\end{bmatrix} \mid \boldsymbol{\mu}_a, \boldsymbol{\Sigma}_a\right) \times \mathrm{N}\left(\begin{bmatrix}b_H \\ b_V\end{bmatrix} \mid \boldsymbol{\mu}_b, \boldsymbol{\Sigma}_b\right) \right\},$$

$$q(\boldsymbol{\theta}^* \mid \boldsymbol{\theta}^{(l-1)}) = \mathrm{N}\left(\boldsymbol{\theta}^* \mid \boldsymbol{\theta}^{(l-1)} + 0.5\zeta^{(l-1)} G^{-1} \nabla \log(p(\boldsymbol{\theta}^{(l-1)} \mid \boldsymbol{d})), \zeta^{(l-1)} G^{-1}\right).$$

References

Girolami, M. and Calderhead, B. (2011). Riemann manifold Langevin and Hamiltonian Monte Carlo methods. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 73(2):123–214.

Stockie, J. M. (2011). The mathematics of atmospheric dispersion modeling. *SIAM Review*, 53(2):349–372.