Generalized spatiotemporally-decoupled framework for reconstructing the source of non-constant atmospheric radionuclide releases

Yuhan Xu[1], Sheng Fang [1,*], Xinwen Dong[1], Shuhan Zhuang[1]

[1]Institute of Nuclear and New Energy Technology, Collaborative Innovation Centre of Advanced Nuclear Energy Technology, Key Laboratory of Advanced Reactor Engineering and Safety of Ministry of Education, Tsinghua University, Beijing 100084, China

**Correspondence to*: Sheng Fang (fangsheng@tsinghua.edu.cn)

This supplementary material (9 pages) includes 1 Note, 5 Figures, 1 Tables and 2 References.

Note S1: Bayesian method for source reconstruction

The Bayesian method is combined with Markov chain Monte Carlo sampling to estimate the source location and release rate simultaneously. Bayes' theorem can be expressed as:

$$p(\mathbf{s}|\boldsymbol{\mu}) = \frac{p(\boldsymbol{\mu}|\mathbf{s})p(\mathbf{s})}{p(\boldsymbol{\mu})} \propto p(\boldsymbol{\mu}|\mathbf{s})p(\mathbf{s}) \ , \tag{S1}$$

where $\mathbf{s}$ is the parameter vector containing source parameters and $\boldsymbol{\mu}$ is an observation vector. $p(\mathbf{s})$ describes the probability distribution of prior knowledge on $\mathbf{s}$ and $p(\boldsymbol{\mu}|\mathbf{s})$ is the likelihood function quantifying the goodness of fit between the simulations and observations. Consistent with general approaches, we define $p(\mathbf{s})$ as the uniform distribution bounded by the lower and upper limits of the source parameters. Referring to (Dumont Le Brazidec et al., 2021), $p(\boldsymbol{\mu}|\mathbf{s})$ is defined as the log-Cauchy distribution:

$$p(\boldsymbol{\mu}|\mathbf{s}) = -\sum_{i=1}^{O} ln\left\{r + \left[ln(\mu_i + \mu_t) - ln\left(\left(\mathbf{F}(\mathbf{s})\right)_i + \mu_t\right)\right]^2\right\} + \frac{1}{2}ln(r) \ , \tag{S2}$$

where r is a covariance parameter which forms the covariance matrix $\mathbf{R} = r\mathbf{I}$ and μ_t is a positive threshold that ensures the logarithm is defined properly for zero measurements or simulations.

However, the release rate is time-varied, so it is not realistic to define the prior distribution of the release rate in every time step. Hence, we incorporate the source inversion method into this process, which involves calculating the corresponding release rate for the sampled source location and then obtaining Eq. (S3) using the sampled location and calculated release rate. We apply a traditional Tikhonov method to calculate the release rate as follows:

$$\mathbf{q}(\mathbf{r}) = argmin\left(\frac{1}{2}\|\boldsymbol{\mu} - \mathbf{A}(\mathbf{r})\mathbf{q}\|_2^2 + \frac{1}{2}\lambda^2\|\mathbf{q}\|_2^2\right) , \tag{S3}$$

where $\mathbf{q}(\mathbf{r})$ refers to the estimated release rate vector under the source location $\mathbf{r}$ and $\mathbf{A}(\mathbf{r})$ refers to the source–receptor matrix given $\mathbf{r}$. λ is a regularization parameter that is automatically selected by generalized cross-validation (Hansen and O'Leary, 1993).

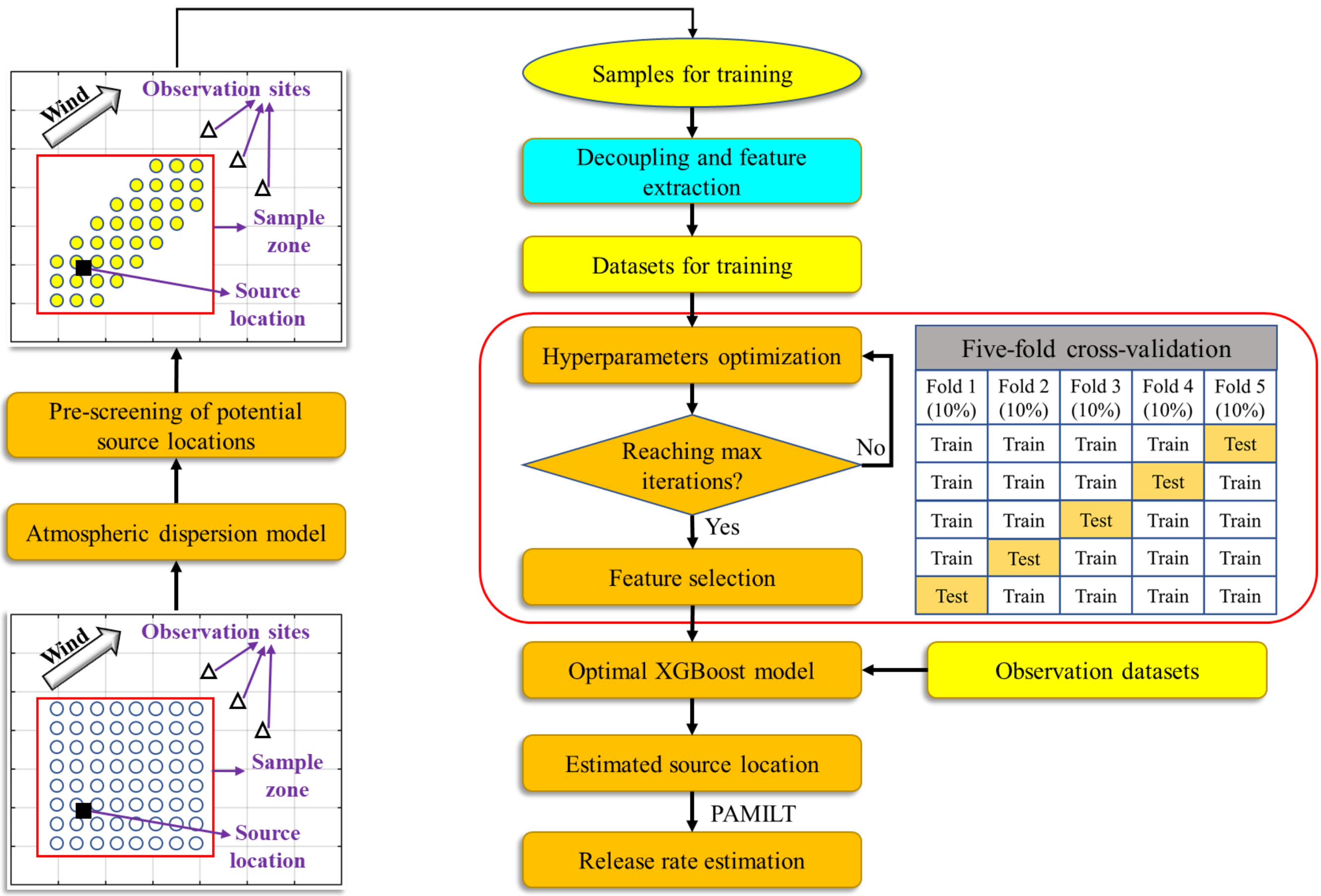


Figure S1. Flowchart of the proposed spatiotemporally decoupled source reconstruction method.

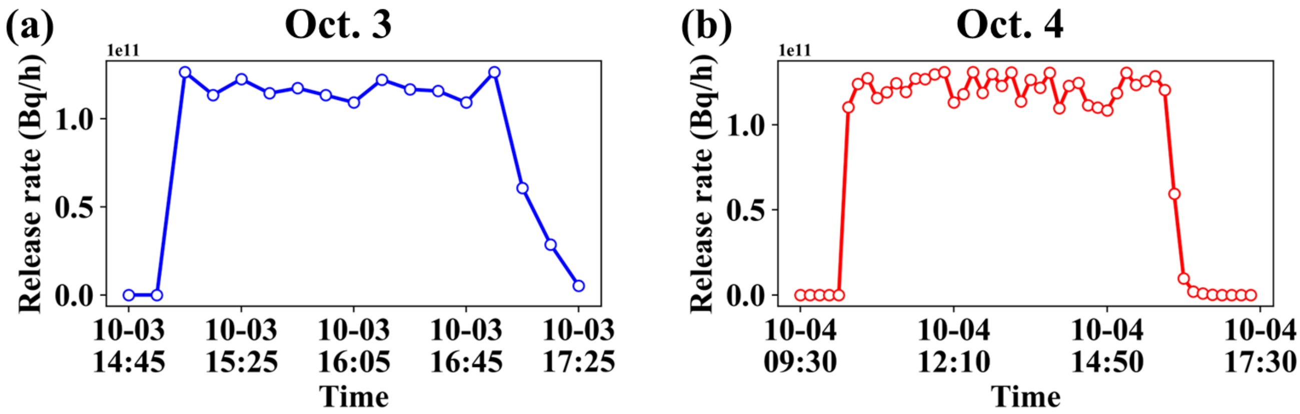
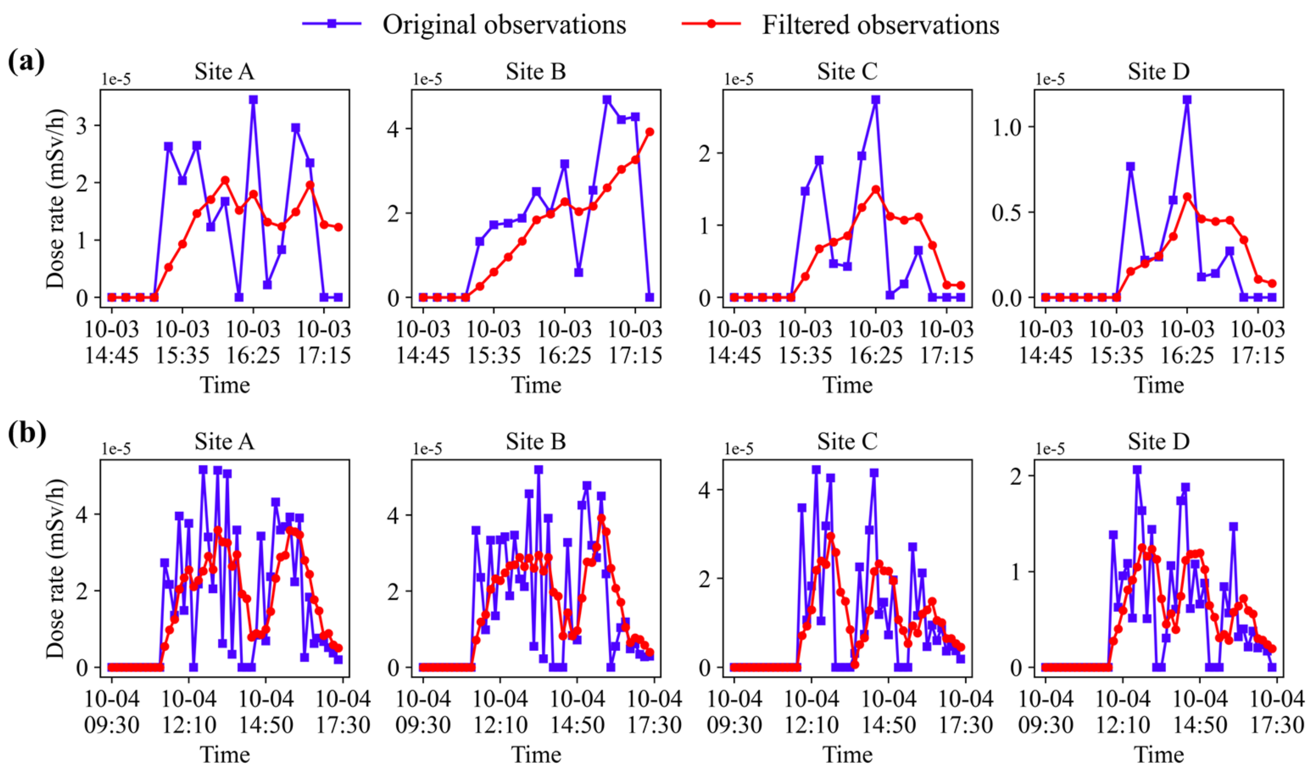


Figure S2. Synthetic release rates for generating synthetic observations. (a) Oct. 3; (b) Oct. 4.



41 **Figure S3.** Observations before and after filtering at observation sites. (a) Oct. 3; (b) Oct. 4.

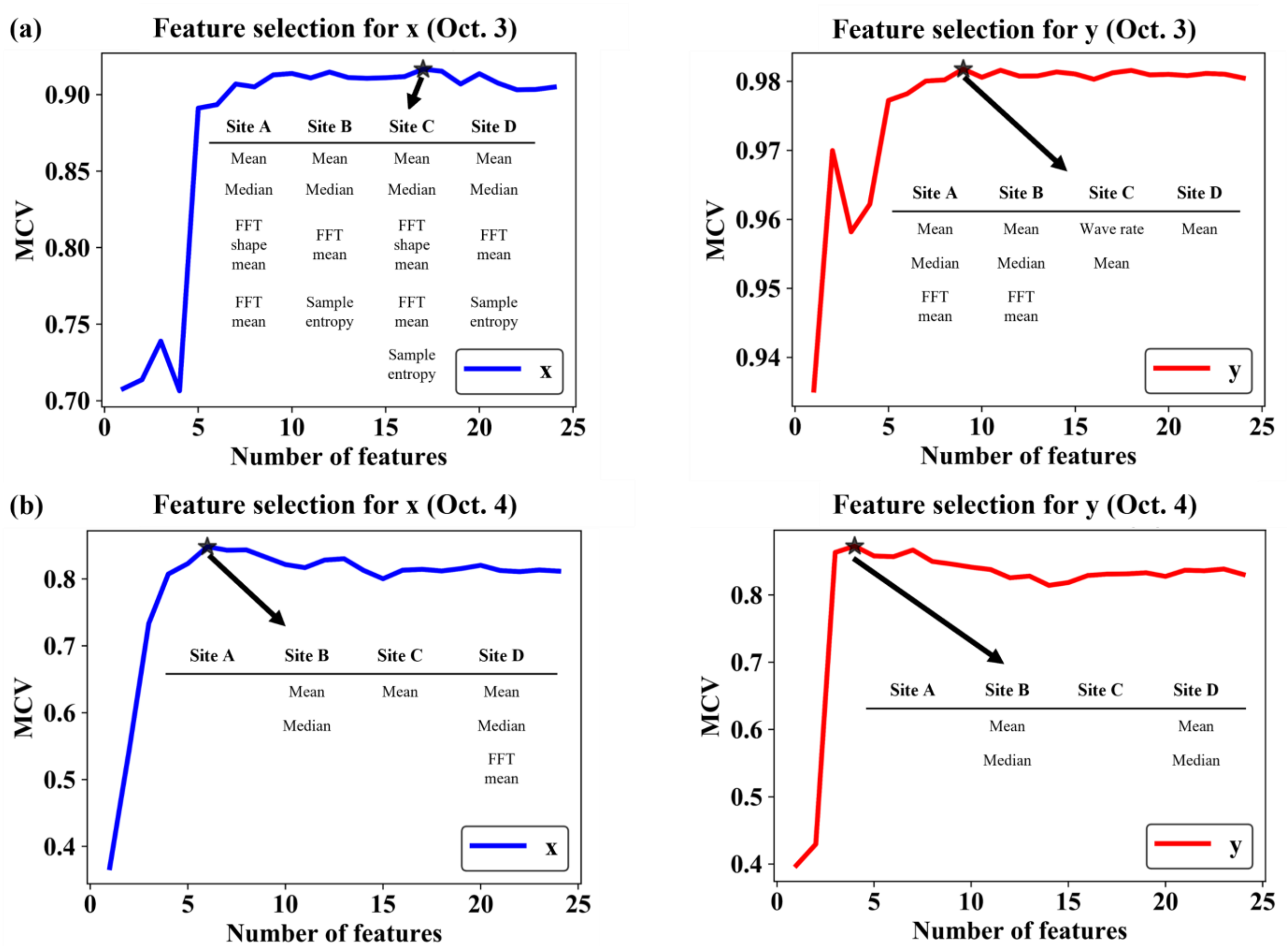


Figure S4. Results of feature selection in x and y directions. (a) Oct. 3; (b) Oct. 4. The black stars denote the optimal number of features. The table inserted in each subgraph lists the selected features for each observation site.

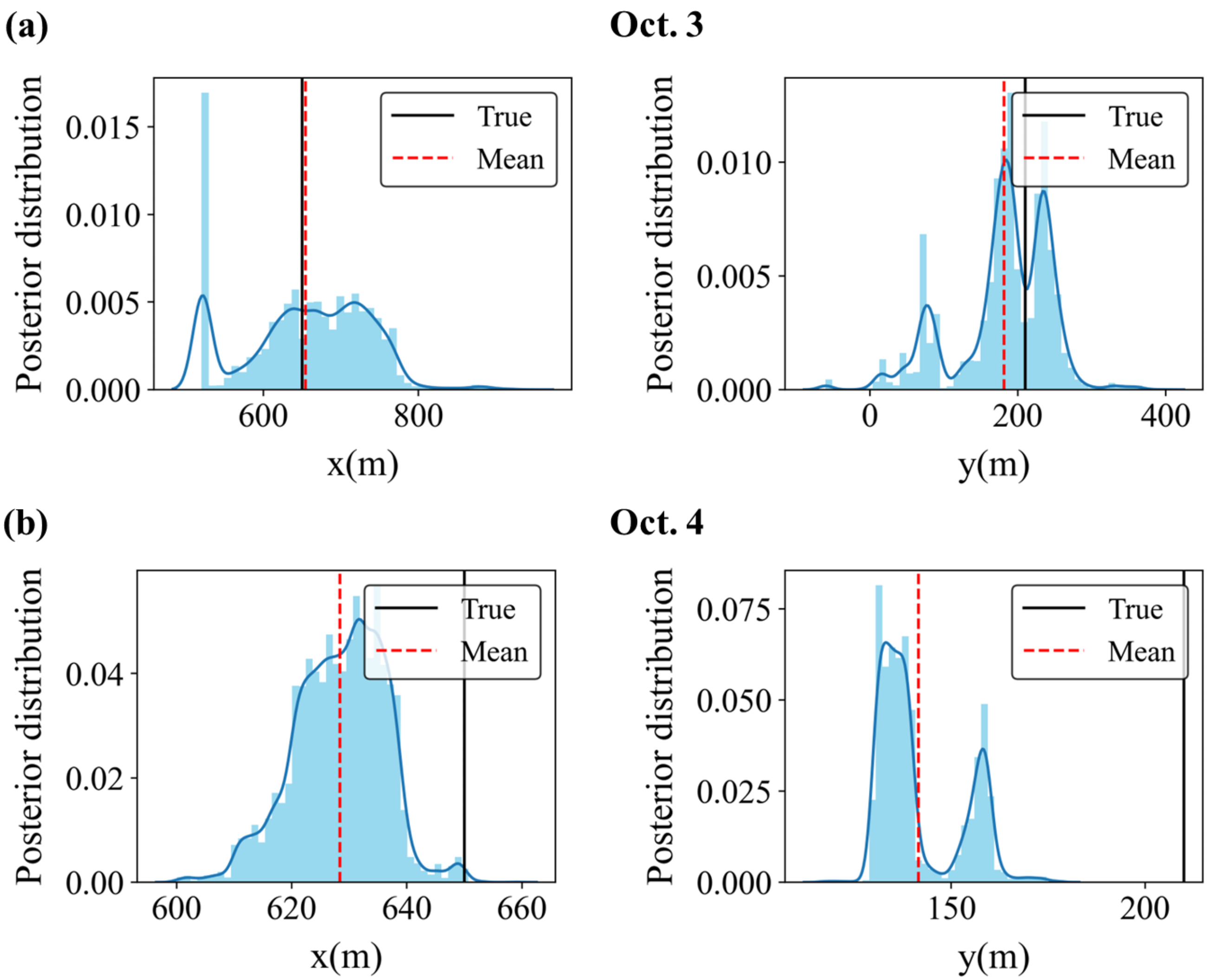


Figure S5. Posterior distributions of source location parameters. (a) Oct. 3; (b) Oct. 4. The black solid lines denote the true location parameters and the dashed lines denote the mean estimates of all posterior samples.

Table S1. Hyperparameter optimization results.

Optimization results		Experiment	
		Oct. 3	Oct. 4
Hyperparameters	*max_depth* ([3,8])	6	7
	learning_rate ([0.05,0.3])	0.05747	0.05199
	n_estimators ([50,300])	291	199
	min_child_weight ([2,10])	5	10
	subsample ([0.5,1])	0.64408	0.69456
	colsample_bytree ([0.01,1])	0.63517	0.99868
	reg_lambda ([0.01,5])	2.44598	4.27874
	gamma ([0.01,1])	0.99061	0.85555
Optimal GC		0.01299	0.05104

References

Dumont Le Brazidec, J., Bocquet, M., Saunier, O., and Roustan, Y.: Quantification of uncertainties in the assessment of an atmospheric release source applied to the autumn 2017 106Ru event, Atmos. Chem. Phys., 21, 13247–13267, https://doi.org/10.5194/acp-21-13247-2021, 2021.

Hansen, P. C. and O'Leary, D. P.: The Use of the L-Curve in the Regularization of Discrete Ill-Posed Problems, SIAM J. Sci. Comput., 14, 1487–1503, https://doi.org/10.1137/0914086, 1993.