Pollutant source term estimation in a two-dimensional street canyon using physics-informed neural networks

Hideki Kikumoto [a], Yichen Wang [b], Bingchao Zhang [c]

[a]*The University of Tokyo, Meguro-Ku, Tokyo, Japan, kkmt@iis.u-tokyo.ac.jp*
[b]*The University of Tokyo, Meguro-Ku, Tokyo, Japan, wyc23@iis.u-tokyo.ac.jp*
[c]*The Hong Kong University of Science and Technology, Kowloon, Hong Kong, China, zhangbc@ust.hk*

SUMMARY

This study investigates the application of physics-informed neural networks (PINNs) for pollutant source term estimation (STE) in a two-dimensional street canyon. We propose a two-step PINN framework to reconstruct flow fields and estimate source distributions from sparsely measured mean velocity and concentration data. Three output constraints for the source term were evaluated: unconstrained, ReLU, and Softplus. Based on a study using a large-eddy simulation dataset, the results demonstrate that enforcing non-negativity via the Softplus function significantly enhances solution stability. Specifically, this constraint effectively eliminated unphysical negative values and spatial artifacts observed in other cases. This yielded a source distribution close to the ground truth with a total emission estimation error of approximately 30%. Unlike conventional methods that are restricted to point sources, our approach successfully visualizes the source's spatial distribution without prior geometric assumptions. These findings confirm the potential of PINNs for flexible inverse analysis in urban environments.

Keywords: *urban street canyon, pollutant dispersion, source term estimation, physics-informed neural networks*

1. INTRODUCTION

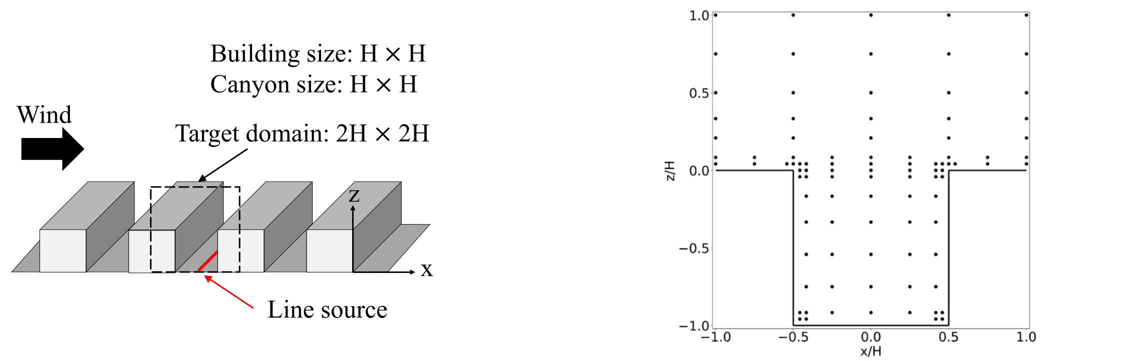
Accurate source term estimation (STE) from measured concentrations is critical for understanding air quality issues and developing effective control strategies. Conventionally, STE has been approached by integrating computational fluid dynamics simulations with optimization or Bayesian inference methods (Hutchinson et al., 2017). A major limitation of these conventional approaches, however, is their reliance on point-source assumptions, which restricts their scalability to complex source geometries (Jia & Kikumoto, 2021).

Recently, physics-informed neural networks (PINNs) have emerged as a powerful paradigm for assimilating data into physical models (Karniadakis et al., 2021). Among their many capabilities, PINNs are particularly promising for handling flexible inverse problems (Raissi et al., 2019). In this study, we explore the potential of PINNs for STE applications. As a preliminary investigation, we demonstrate the efficacy of a PINNs-based STE approach applied to pollutant dispersion in a simplified two-dimensional street canyon.

2. METHODOLOGY

2.1. Street canyon and sensing data configurations

Based on a large-eddy simulation (LES) study (Zhang et al., 2022), a two-dimensional street canyon with dimensions of $H \times H$ ($H = 0.12$ m), as shown in Fig. 1(a), was investigated. A tracer line source extending in the spanwise direction is located at the center of the canyon bottom. We assumed that the tracer is emitted at a constant rate from this line source (but unknown to STE) within a steady flow field, resulting in the formation of a concentration field. Virtual measurements of wind velocity and concentration were extracted from the LES at discrete points, as depicted in Fig. 1(b). The objective of this study is to perform STE to inversely determine the distribution of the tracer source term from these measurement values.



(a) Configuration of the canyon and domain

(b) Sensing arrangement for velocity and concentration (black dots)

Figure 1: Schematic of the two-dimensional street canyon model and sensing arrangement.

2.2. PINN framework for source term estimation

The authors have previously proposed a method for reconstructing detailed flow fields from sparse measurement data using PINNs (Wang et al., 2026). Fig. 2 illustrates the two-step framework for PINN-based STE used in this study. In Step 1, the distributions of mean velocity $\bar{\boldsymbol{u}}$ (= $(\bar{u}, \bar{w})$), mean pressure $\bar{p}$, and eddy viscosity ν_t are estimated using sparsely measured mean velocities ($\bar{\boldsymbol{u}}$), the continuity equation, and the Reynolds-averaged Navier-Stokes (RANS) equations (Eq. (1)) based on the gradient diffusion approximation. Subsequently, in Step 2, the distributions of mean concentration $\bar{c}$ and the concentration source term q are estimated via the concentration transport equation (Eq. (2)), utilizing the densely predicted $\bar{\boldsymbol{u}}$ and ν_t by the model established in Step 1 and the sparsely measured mean concentrations ($\bar{c}$). Here, Re, Sc, and Sc_t denote the Reynolds number (= 7680), the Schmidt number (= 1), and the turbulent Schmidt number (= 0.5), respectively.

$$\nabla \cdot \bar{\boldsymbol{u}} = 0, \qquad \bar{\boldsymbol{u}} \cdot \nabla \bar{\boldsymbol{u}} = -\nabla \bar{p} + \nabla \cdot [(1/Re + \nu_t)(\nabla \bar{\boldsymbol{u}} + (\nabla \bar{\boldsymbol{u}})^T)] \tag{1}$$

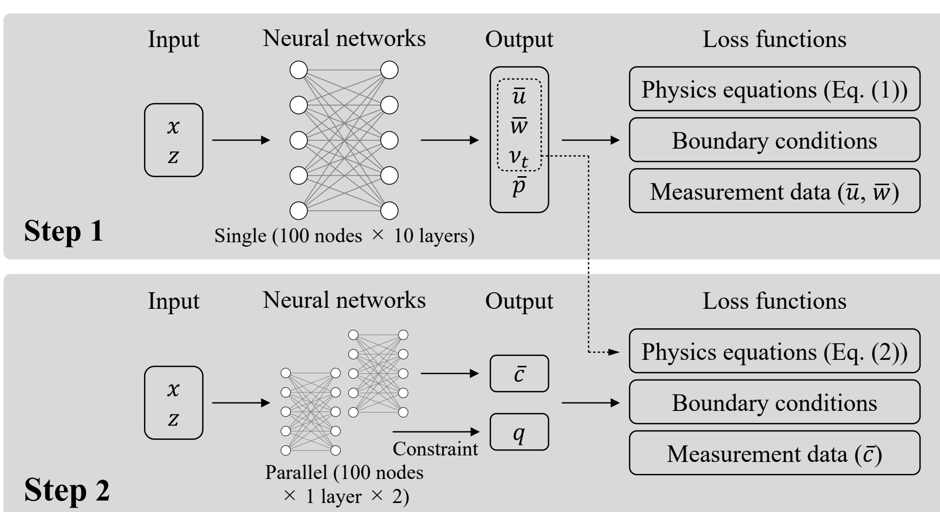


Figure 2: Framework of the two-step PINN model for source term estimation.

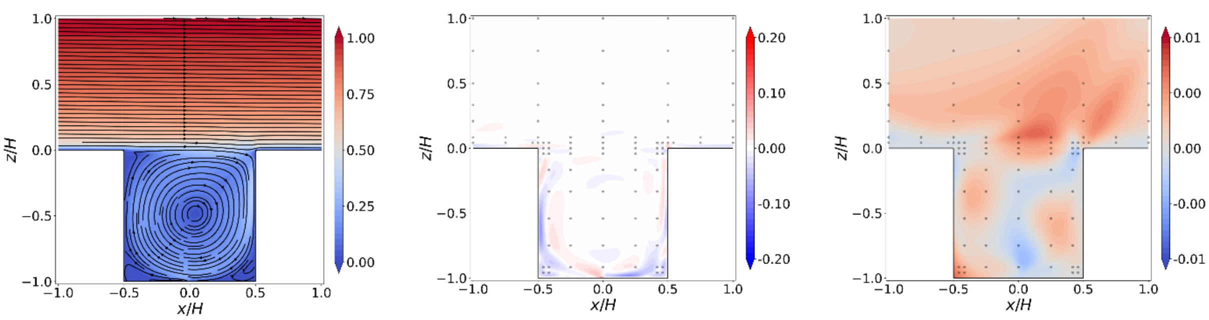
$$\overline{\boldsymbol{u}} \cdot \nabla \bar{c} = \nabla \cdot [(1/ReSc + \nu_t/Sc_t)\nabla \bar{c}] + q \tag{2}$$

The PINN model was implemented using DeepXDE (Lu et al., 2021), an open-source Python library. The neural network architectures are illustrated in Fig. 2. Since the source term is physically non-negative, three cases were considered for the output node of q: an unconstrained case and two cases enforcing non-negativity using the ReLU and Softplus (sharpness parameter = 8) functions. In Step 1, the training process involved $3{\times}10^4$ iterations of the Adam optimizer followed by $2{\times}10^4$ iterations of the L-BFGS. In Step 2, only the L-BFGS was utilized.

3. RESULTS

Fig. 3 shows the mean-flow and eddy-viscosity distributions reconstructed in Step 1 from sparse velocity data. Although slightly larger errors are observed near the ground and building surfaces, the reconstructed flow field shows good agreement with the ground truth (original LES data).

Fig. 4 presents the distributions of the pollutant source term, q, estimated in Step 2. In the case where no constraint was applied to q, non-zero values were widely distributed across the entire domain, including unphysical negative values. When ReLU was used as a constraint, q became concentrated near the center of the canyon floor, the true source location, but non-zero values were also observed in unrealistic locations, such as near the roof level. In the Softplus case, these artifacts were eliminated, and the distribution was estimated to be closer to the true source location. The total emission rates obtained by spatially integrating these q distributions (where the ground truth is 1) were 1.44, 1.75, and 1.29 for the No-constraint, ReLU, and Softplus cases, respectively.

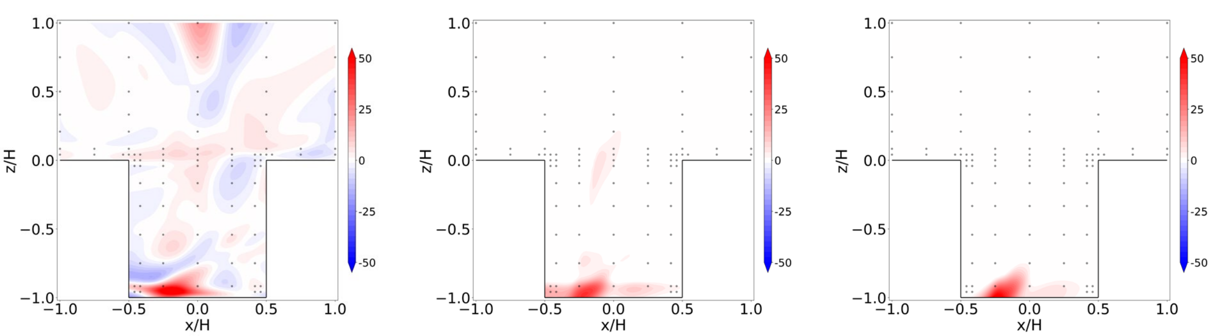


(a) Mean flow (magnitude and streamlines)

(b) Error in mean flow magnitude (reconstructed - truth)

(c) Eddy viscosity

Figure 3: Reconstructed fields and error in Step 1. The gray dots indicate the locations of the velocity data.

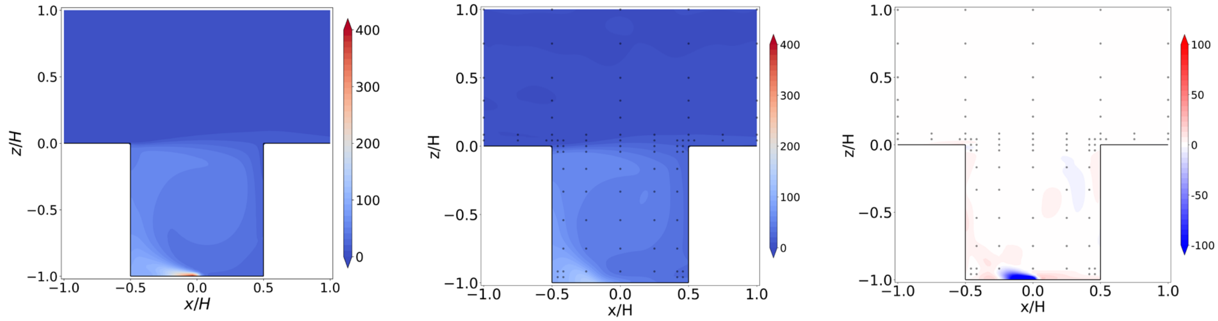


(a) No constraint

(b) ReLU

(c) Softplus

Figure 4: Distributions of reconstructed source term q in Step 2.



(a) Truth (LES) (b) Reconstructed (Softplus) (c) Error (reconstructed - truth)

Figure 5: Distributions of true and reconstructed concentration $\bar{c}$ in Step 2.

Fig. 5 shows the concentration distribution reconstructed in the Softplus case, corresponding to the source term described above. The overall distribution pattern agrees well with the ground truth. While some underestimation is observed near the source, this is likely attributable to the sparsity of measurement data and the inherent limitations of the eddy diffusivity model (Eq. (2)) accuracy.

4. CONCLUSION

To assess the potential of PINNs for STE, this study examined pollutant dispersion from a source at the bottom center of a two-dimensional street canyon. A two-step PINN model was employed to reconstruct the source distribution from sparse velocity and concentration measurements. The results indicate that constraints on the source term output significantly affect solution stability. Specifically, enforcing non-negativity via the Softplus function yielded a source distribution that closely matched the ground truth with a total emission estimation error of approximately 30%.

ACKNOWLEDGEMENTS

This work was partially supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI Grant Numbers 23K26257 and 24K21642.

REFERENCES

Hutchinson, M., Oh, H., & Chen, W.-H. (2017). A review of source term estimation methods for atmospheric dispersion events using static or mobile sensors. *Information Fusion*, *36*, 130-148. https://doi.org/10.1016/j.inffus.2016.11.010

Jia, H., & Kikumoto, H. (2021). Line source estimation of environmental pollutants using super-Gaussian geometry model and bayesian inference. *Environmental Research*, *194*, 110706. https://doi.org/10.1016/j.envres.2020.110706

Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., & Yang, L. (2021). Physics-informed machine learning. *Nature Reviews Physics*, *3*(6), 422-440. https://doi.org/10.1038/s42254-021-00314-5

Lu, L., Meng, X., Mao, Z., & Karniadakis, G. E. (2021). DeepXDE: A Deep Learning Library for Solving Differential Equations. *SIAM Review*, *63*(1), 208-228. https://doi.org/10.1137/19m1274067

Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, *378*, 686-707. https://doi.org/10.1016/j.jcp.2018.10.045

Wang, Y., Zhang, B., & Kikumoto, H. (2026). Two-step High-resolution Reconstruction of Mean Flow Field and Reynolds Stress Distributions using Physics-Informed Neural Networks in a Two-Dimensional Street Canyon. *Building and Environment*, *288*. https://doi.org/10.1016/j.buildenv.2025.113958

Zhang, B., Ooka, R., & Kikumoto, H. (2022). Spatiotemporal Spectral Analysis of Turbulent Structures and Pollutant Removal in Two-Dimensional Street Canyon. *Boundary-Layer Meteorology*, *185*(1), 63-91. https://doi.org/10.1007/s10546-022-00724-7