Global Environmental Science
www.sciltp.com/journals/ges

*Article*

# Multi-Factor Coupled Deep Learning Framework for Predicting Toxic Gas Diffusion Patterns in Urban Areas

Mingtao Ding [1,†], Xuqiang Shao [2,3,†], Xinhang Jing [2,4], Xiaofan Liu [2], Sixuan Chen [2], Zijia Zhao [2], Jiancai Huang [2,4] and Zhijian Liu [1,*]

[1] Department of Power Engineering, North China Electric Power University, Baoding 071003, China
[2] Department of Computer Science, North China Electric Power University, Baoding 071003, China
[3] Hebei Key Laboratory of Knowledge Computing for Energy & Power, Baoding 071003, China
[4] Engineering Research Center of Intelligent Computing for Complex Energy Systems, Ministry of Education, Baoding 071003, China
* Correspondence: zhijianliu@ncepu.edu.cn
† These authors contributed equally to this work.

**How To Cite:** Ding, M.; Shao, X.; Jing, X.; et al. Multi-Factor Coupled Deep Learning Framework for Predicting Toxic Gas Diffusion Patterns in Urban Areas. *Glob. Environ. Sci.* **2026**, *2*(4), 553–576. https://doi.org/10.53941/ges.2026.100034

## Publication History

Received: 30 December 2025
Revised: 16 June 2026
Accepted: 16 November 2026
Published: 25 November 2026

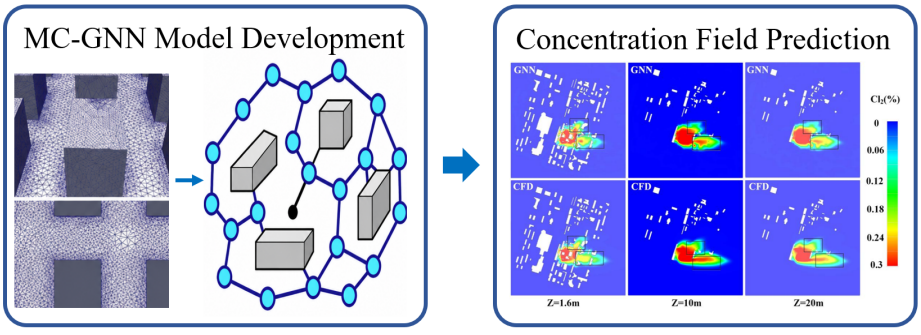
## Keywords

toxic gas diffusion;
multiple factors;
deep learning;
graph neural network;
exposure risk

## Highlights

- An MC-GNN model integrating multiple factors is proposed for urban toxic gas diffusion prediction.
- MC-GNN accurately predicts urban toxic gas diffusion while substantially reducing computational time.
- Lateral distance analysis of high-risk regions.

**Abstract:** Toxic gas leakage accidents in urban environments pose serious threats to public health, environmental safety, economic activity, and social stability, underscoring the urgent need for effective monitoring and prevention strategies. At present, the diffusion of toxic gases is typically analyzed using computational fluid dynamics (CFD) simulations. However, toxic gas diffusion in urban settings is jointly influenced by multiple factors, including building layouts, meteorological conditions, leakage concentrations, and leakage rates, thereby making it difficult for traditional numerical approaches to comprehensively account for these influences. Moreover, CFD simulations inherently suffer from high computational cost and long simulation time. To address these limitations, this study proposes a multi-factor coupled graph neural network model. The graph architecture effectively represents the unstructured flow field data produced by CFD simulations, and the incorporation of meteorological data, leakage-source parameters, and inlet boundary conditions enables comprehensive prediction of toxic gas diffusion under multi-factor coupling. Furthermore, a dose–response model is introduced to assess exposure risks, and dimensional analysis is employed to evaluate the lateral distance of high-risk regions. The π theorem is used to derive an analytical formula for lateral distance prediction, and its validity is verified using 40 simulation cases. The results demonstrate that the proposed predictive framework improves computational speed by nearly one order of magnitude compared with traditional CFD simulations, achieving an average $R^2$ of up to 0.84 for concentration field prediction. Through dimensional analysis, the π theorem–based lateral-distance prediction formula achieves an $R^2$ of 0.96, with absolute errors ranging from 0.1% to 5.03%.




**Copyright:** © 2026 by the authors. This is an open access article under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).
**Publisher's Note**: Scilight stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## 1. Introduction

Currently, urbanization is progressing rapidly, making urban areas the most densely populated places. According to statistics, more than 50% of the world's population lives in cities on average, and the urban population in China accounts for 63.89%. At the same time, the improvement of industrialization has led to frequent leaks of toxic gases such as hydrogen sulfide, methyl isocyanate, and chlorine. According to statistical data, a total of 364 chemical accidents occurred nationwide from 2021 to 2023, resulting in 452 deaths. Among them, poisoning and asphyxiation accidents accounted for 78 cases and 116 fatalities, representing 21.4% of all accidents and 25.7% of the total deaths, respectively, making them the accident types with the highest number of incidents and fatalities. It can be seen that accidental leakage of toxic gases in densely populated urban environments often causes serious casualties. Therefore, it is crucial to grasp the toxic gas diffusion law in a timely and accurate manner after the accident, which can not only conduct urban safety risk assessment, but also effectively guide emergency rescue operations, reduce the harm caused by the accident, and thus strengthen the resilience of the city.

At present, the commonly used approaches for studying toxic gas dispersion mainly include field experiments [1], scale model experiments [2,3], and numerical simulation [4,5]. Field experiments are conducted by recreating scenarios consistent with the scale of actual leakage accidents, and the experimental subjects typically involve gases such as $SO_2$, $CO_2$, LNG, $Cl_2$, $NH_3$, and their mixtures. Representative experiments include the Maplin Sands experiment [6–8], the Kit Fox experiment [9], and the Jack Rabbit series [1,10]. Field experiments can provide highly reliable and realistic data; however, they require extensive preparation, involve long experimental cycles, pose considerable safety risks, and suffer from poor repeatability. Consequently, conducting field experiments on a large scale in urban environments remains challenging. The wind tunnel method employs scaled physical models to reproduce large-scale scenarios under controlled atmospheric conditions, thereby enabling small-scale simulations of toxic gas dispersion. This technique has been widely applied in studies involving single buildings [11], building arrays [12], and even realistic urban environments [13]. Variations in urban morphology, building geometry [14,15], street aspect ratios, and meteorological parameters can all significantly influence dispersion behavior. For instance, Tan et al. [16] conducted comparative experiments using the wind tunnel platform at the Tianjin Fire Research Institute, considering the combined effects of wind speed, wind direction, temperature, and humidity on heavy gas dispersion in urban settings. Compared with field experiments, wind tunnel scale model experiments can simplify the experimental process and reduce operational risks to a certain extent. Nevertheless, the inevitable simplification of complex urban environments in scaled models limits the representativeness of the results and reduces their consistency with real-world conditions. Numerical simulation does not require costly experimental equipment or dedicated test sites, and enables multi-condition studies directly through computational resources. Parameters such as leakage intensity, wind speed and direction, temperature, humidity, and building layout can be modified flexibly, facilitating multi-scenario analysis and forecasting. For instance, prior studies have explored pollutant behavior in idealized building arrays [17], investigated the influence of boundary shapes on gas diffusion [18], and simulated diffusion scenarios in real urban layouts [19]. Such versatility makes CFD particularly valuable for understanding the complex dynamics of toxic gas diffusion. Despite its strengths, CFD simulations don't consider the combined effects of multiple factors. For example, previous work [20,21] has investigated the effects of wind speed and direction by adjusting inflow boundary parameters, yet it lacks a comprehensive consideration of multiple factors. The diffusionFoam solver, developed in previous work [22], was coupled with WRF to incorporate multiple meteorological factors and enhance simulation accuracy. Nevertheless, CFD simulation methods are computationally intensive and time-consuming. Moreover, for large-scale urban areas, high-resolution turbulence simulations require substantial computational resources and long computation time, making it difficult to provide rapid feedback for emergency toxic gas leakage incidents.

With the rapid development of deep learning technologies, constructing appropriate deep learning models to replace traditional numerical simulations for fast prediction of wind fields and toxic gas concentration fields in urban environments has gained increasing attention. Predictive models for toxic gas diffusion rely on observational data and CFD simulation data. Observational data, such as sensor measurements, provide real-world information but lack comprehensive spatiotemporal coverage [23–25]. In contrast, CFD simulations offer high-resolution, full-field spatiotemporal distributions of toxic gas concentration field, making them a crucial resource for training deep learning models. In CFD-driven methods, Convolutional Neural Networks (CNNs) [26] are commonly applied due to their strong feature extraction capabilities. CNNs can conduct both steady-state flow field prediction [27,28] and transient flow field [29–31] prediction, achieving a prediction speed several orders of magnitude faster than CFD solvers. However, CNNs are limited to regular grids, and cannot be generalized to more commonly used unstructured meshes. Graph Neural Networks (GNNs) [32] address these limitations by effectively representing unstructured data. For example,

 https://doi.org/10.53941/ges.2026.100034

the work by Zhao et al. [33] demonstrated the application of GNNs to steady-state toxic gas diffusion, directly correcting CFD outputs towards a high-fidelity target in a single inference step and bypassing the limitations of steady RANS flaws. Additionally, a multi-objective GNN model was proposed in previous work [34] as CFD surrogate to rapidly predict the transient dispersion of airborne pollutants while achieving significant GPU memory reduction. However, current GNN methods struggled to model the coupled effects of multiple factors. To address this gap, we propose a Multi-Factors Coupled Graph Neural Network (MC-GNN) model, which leverages the flexibility of GNNs to handle complex spatial data structures. This model replaces CFD simulations for urban toxic gas diffusion prediction, effectively learning intricate relationships within unstructured flow field data.

To provide critical data support for post-accident emergency response and thereby minimize casualties, this study further conducts exposure risk assessment based on the concentration field of toxic gas diffusion. Based on the rapidly predicted toxic gas concentration field data, conducting fast and accurate exposure risk analysis is crucial for effective emergency response and urban safety planning. Current methods commonly employ models such as the Wells-Riley model [35] and dose–response model [36] to assess exposure risk. The Wells-Riley model commonly evaluates the risk of airborne diseases [37,38]. The dose-response model can describe and evaluate the relationship between exposure doses and biological effects [39,40], enabling the evaluation of toxic gas exposure risks, determination of safety limits, and guidance for risk management. Therefore, we use a dose-response model to evaluate exposure risk, defining the furthest position from the source with a mortality rate not exceeding 40% within a specified timeframe as the diffusion edge. Dimensional analysis is applied to determine the lateral distance of high exposure risk areas.

In summary, the rapid and accurate prediction of toxic gas diffusion and exposure risk based on the concentration field can effectively mitigate the harm caused by toxic gas leakage accidents. The key challenges lie in constructing high-accuracy time ordered flow field datasets and designing precise deep neural network models. This study utilizes a numerical simulation solver considering multiple factors to generate training data and develops a deep learning model for toxic gas diffusion prediction. Simulations are conducted for the Beijing Chemical Experimental Plant and nearby urban areas under varying inflow boundary conditions to rapidly predict diffusion scenarios. Additionally, exposure risks are evaluated using a dose-response model, and the lateral distances of high exposure risk areas are analyzed through dimensional analysis. Overall, this research provides a fast and effective method for predicting toxic gas concentration fields in urban environments, offering important contributions for enhancing urban resilience and emergency response.

## 2. Methods

### *2.1. Multi-Factor Coupled Numerical Simulation Method*

#### 2.1.1. Inflow Boundary Conditions for Numerical Simulation Considering Multiple Factors

To ensure that the study aligns more closely with real-world conditions, multiple influencing factors on toxic gas dispersion are considered, including meteorological factors such as wind speed, wind direction, and temperature, as well as leakage source concentration and leakage rate. A one-way coupling method between WRF and OpenFOAM is adopted, in which the meteorological conditions obtained from WRF simulations—such as wind speed, wind direction, and temperature—are combined with different leakage source concentrations and leakage rates and applied as inflow boundary conditions in OpenFOAM for numerical simulation.

The WRF model utilizes a three-level nested grid technique with the central simulation area set in Chaoyang District, Beijing (116.48° E, 39.89° N). We selected representative dates and times to capture seasonal and diurnal variations that could affect gas diffusion: 11 October 2022, 22 December 2022, 28 April 2023, and 22 June 2023, at 11:50 and 23:50 Beijing time. The selected dates represent typical seasonal conditions (spring, summer, autumn, and winter), while the selected times (11:50 and 23:50 Beijing time) correspond to representative daytime and nighttime atmospheric states. This selection allows the dataset to capture seasonal variations in temperature and wind conditions, as well as diurnal differences in atmospheric stability that may affect gas dispersion. These dates span diverse meteorological conditions, thereby providing a robust set of scenarios for our analysis.

The WRF simulations provide meteorological outputs, including wind direction angles, average wind speeds, and near-ground temperatures. The corresponding temperatures (converted to Celsius) are 12.43 °C, −4.72 °C, 15.85 °C, and 32.85 °C, while wind direction angles are 85.48°, 162.87°, 98.74°, and 152.08°, respectively. The specific information is shown in Table 1. The selected dates represent different seasonal meteorological conditions in the study area. By including variations in temperature and wind speed, the dataset captures a diverse range of atmospheric conditions that influence toxic gas dispersion patterns, thereby improving the generalization capability of the proposed model.

The variation of wind speed with height was fitted using a polynomial approach, in which different polynomial orders were explored. It was observed that lower-order polynomials performed less effectively in certain regions, while higher-order polynomials may introduce unnecessary oscillations. Therefore, a quintic

polynomial was selected as a practical compromise to best capture the wind speed. The equations used for fitting the wind speed are given as follows:

$$y = -3.34\times10^{-12}\cdot z^5 + 4.64\times10^{-9}\cdot z^4 - 2.2\times10^{-6}\cdot z^3 + 4.13\times10^{-4}\cdot z - 0.0193\cdot z + 2.44 \quad (1)$$

$$y = -1.43\times10^{-12}\cdot z^5 + 1.73\times10^{-9}\cdot z^4 - 5.2\times10^{-7}\cdot z^3 - 5.03\times10^{-5}\cdot z^2 + 0.0383\cdot z + 2.081 \quad (2)$$

$$y = -4.38\times10^{-12}\cdot z^5 + 6.01\times10^{-9}\cdot z^4 - 2.46\times10^{-6}\cdot z^3 + 3.93\times10^{-4}\cdot z - 0.0073\cdot z + 0.4107 \quad (3)$$

$$y = 3.27\times10^{-11}\cdot z^5 - 3.01\times10^{-8}\cdot z^4 + 1.04\times10^{-5}\cdot z^3 - 2.0\times10^{-3}\cdot z^2 + 0.2448\cdot z + 2.04 \quad (4)$$

where y represents wind speed and $z$ represents the height above the ground. The fitting wind profile results are shown in Figure S1.

**Table 1.** Meteorological information of typical days.

| Date | Temperature (℃) | Wind Direction Angle (°) |
|---|---|---|
| 2022-10-11 | 12.43 | 85.48 |
| 2022-12-22 | −4.72 | 162.87 |
| 2023-04-28 | 15.85 | 98.74 |
| 2023-06-22 | 32.85 | 152.08 |

The wind speed, wind direction, and temperature data obtained from WRF simulations are used to generate boundary conditions required for CFD numerical simulations. The generated inflow boundary conditions are presented in Table 2.

Among the meteorological factors that influence the dispersion of toxic gases, the wind direction factor is derived from the wind speed calculated by WRF meteorological simulation. By rotating the building model at the corresponding angle, the resulting wind direction angle is obtained, as depicted in Figure S2.

**Table 2.** Boundary conditions for CFD numerical simulation.

| Parameter | Inlet Boundary | Outlet Boundary | Walls (Buildings & Floors) | Sides (Top & Side) |
|---|---|---|---|---|
| velocity (*U*) | WRF gets the wind profile | zeroGradient | noSlip | Symmetry |
| pressure (*p*) | zero gradient | fixedValue | zeroGradient | Symmetry |
| turbulent energy (*k*) | calculated according to U | inletOutlet | kqRWallFunction | Symmetry |
| turbulent dissipation rate ($\varepsilon$) | calculated according to U | inletOutlet | epsilonWallFunction | Symmetry |
| temperature (*T*) | WRF to get the temperature | inletOutlet | fixedValue | Symmetry |
| turbulent viscosity ($\nu_t$) | calculated | Calculated | nutkWallFunction | Symmetry |

Meanwhile, source properties such as leakage concentration and leakage rate were systematically varied. Source concentration was specified as mass fractions of 0.1, 0.5, and 1, representing different severity levels of gas leakage, while leakage rates were set at 5 m/s, 10 m/s, and 15 m/s, simulating different leakage intensities.

To summarize, computational fluid dynamics (CFD) simulations incorporated these boundary conditions and source properties to generate a comprehensive dataset capturing the spatiotemporal diffusion behavior of toxic gases in a real urban environment. This approach ensures the dataset reflects diverse environmental and source conditions, enabling robust training and validation of subsequent deep learning models.

### 2.1.2. Numerical Solver

In this study, we analyze turbulence using the RNG $k-\varepsilon$ model and opt for standard wall functions near building walls. The control equations are solved by employing the finite volume method. The general form of the control equations for RNG $k-\varepsilon$ is outlined below:

$$\frac{\partial}{\partial t}(\rho\varphi) + \nabla\cdot(\rho\varphi\vec{\mathbf{V}}) = \nabla\cdot(\Gamma_\varphi\nabla_\varphi) + S_\varphi \quad (5)$$

where $\rho$ denotes air density, $\vec{\mathbf{V}}$ signifies the air velocity vector, $\varphi$ represents one of the three velocity components, $\Gamma_\varphi$ stands for the effective diffusion coefficient, and $S_\varphi$ denotes the source term.

Gas concentration computation is conducted in this study using the diffusionFoam solver [22]. The primary tasks of the diffusionFoam solver include solving the continuity equation (Equation (6)), momentum equation (Equations (7) and (8)), component transport equation (Equation (9)), and energy equation (Equation (10)).

$$\frac{\partial\rho}{\partial t} + \nabla\cdot(\rho\mathbf{U}) = 0 \quad (6)$$

$$\frac{\partial}{\partial t}(\rho\mathbf{U}) + \nabla(\rho\mathbf{U}\mathbf{U}) = -\nabla p + \nabla\tau \quad (7)$$

$$\tau = \mu[(\nabla\mathbf{U} + \nabla\mathbf{U}^T) - (2/3)\cdot\nabla\cdot\mathbf{U}I] \quad (8)$$

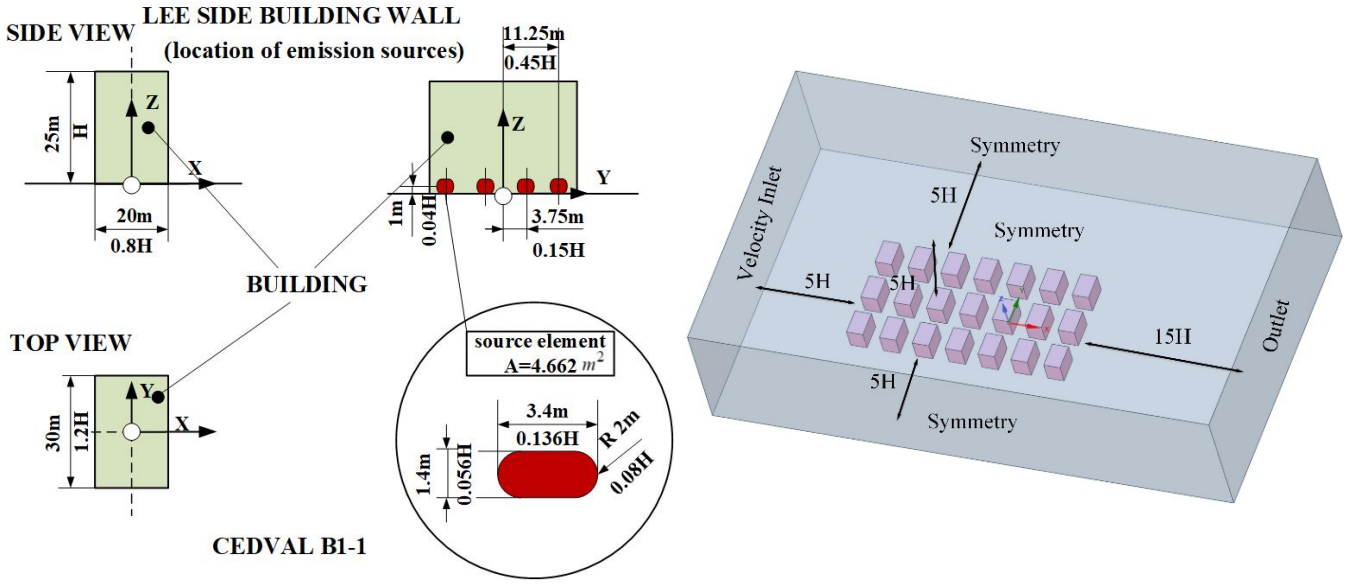
$$\frac{\partial}{\partial t}(\rho Y_i) + \nabla(\rho \mathbf{U} Y_i) = \nabla \mu_{eff} \nabla Y_i \quad (9)$$

$$\frac{\partial}{\partial t}(\rho h) + \nabla(\rho \mathbf{U} h) + \frac{\partial}{\partial t}(\rho K) + \nabla(\rho \mathbf{U} K) - \frac{\partial p}{\partial t} = \nabla \alpha_{eff} \nabla h \quad (10)$$

where $\rho$ signifies air density, $t$ stands for time, $\mathbf{U}$ represents the velocity vector, $p$ indicates pressure, $\mu$ refers to the dynamic viscosity of the fluid mixture, and $\mu_{eff}$ represents the effective dynamic viscosity derived from the turbulence model. Within the component transport equation, $Y_i$ denotes the mass fraction of the $i$-th gas, while the energy equation solves for internal enthalpy $h$, with $K$ and $\alpha_{eff}$ respectively representing kinetic energy and effective thermal diffusion coefficient.

### 2.1.3. Solver Validation

To verify the accuracy of the numerical simulations employed in this paper, simulations were conducted based on the wind tunnel experiment B1-1 from the CEDVAL project (Figure 1a).




(**a**) (**b**)

**Figure 1.** CEDVAL B1-1 Wind Tunnel. (**a**) Wind tunnel layout diagram. (**b**) CEDVAL B1-1 Wind Tunnel Calculation Domain.

The experiment took place in the Blasius wind tunnel [41], located at the Institute of Meteorology, University of Hamburg. The buildings were scaled down to a ratio of 1:200 for modeling purposes. Four $CO_2$ release sources were strategically placed near the buildings, as illustrated in Figure 1a. The concentration was quantified in a dimensionless manner, as described by Equation (11):

$$K_c = \frac{C_{measured}}{C_{source}} \frac{U_{ref} H^2}{Q} \quad (11)$$

where $C_{measured}$ denotes the concentration of the measured tracer gas (unit: ppm), $C_{source}$ signifies the concentration of the tracer gas at the release source (unit: ppm), $U_{ref}$ represents the reference wind speed value measured at a height of 0.66m (unit: m/s), $H$ indicates the height of the tallest building ($H$ = 0.125 m), and $Q$ denotes the release rate of the source (unit: $m^3/s$).

In order to ensure thorough dispersion of the fluid in the air, as stipulated by the CFD practice guidelines [42], the upstream, downstream, lateral, and height components of the computational domain are respectively set to 5H, 15H, 5H, and 5H, as illustrated in Figure 1b. The inlet boundary conditions for the domain are specified using profiles for mean wind speed, turbulent kinetic energy, and turbulent dissipation rate, given by Equations (12)–(14).

$$U_{(Z)} = \frac{u^*}{K} \ln\left(\frac{z + z_0}{z_0}\right) \quad (12)$$

$$k = \sqrt{C_1 \ln(z + z_0) + C_2} \quad (13)$$

$$\varepsilon = \frac{u^* \sqrt{C_\mu}}{K(z + z_0)} \sqrt{C_1 \ln(z + z_0) + C_2} \quad (14)$$

where $U(z)$ signifies the average wind speed at height $z$, $z_0$ denotes the roughness length (0.00075 m), and $U^*$ stands for the friction velocity (0.37772 m/s). Constants $C_1$ and $C_2$ are 0.025 and 0.41 respectively; $C_\mu$ remains constant at 0.09; $K$ represents the von Karman constant, valued at 0.4187 according to Ai and Mak [43].

In order to conduct grid sensitivity analysis, this study set up three different grid systems with minimum grid sizes of 0.006 m, 0.004 m, and 0.003 m, respectively.

Correspondingly, the number of grid units in these grid systems is 1841290 (coarse), 3821297 (medium), and 6593461 (fine), respectively. As shown in Figure S3, Figure S3a and Figure S3b correspond to the coarse-grid resolutions of the 3D map and the 0.01 m height-section map, respectively. Figure S3c,d presents the mid-grid resolutions of the 3D image and the 0.01 m height section. Figure S3e,f illustrates the fine-grid resolutions of the 3D map and the 0.01 m height-section map.

A comparison of the simulation results obtained using coarse (RNG_C), medium (RNG_M), and fine (RNG_F) grids is presented in Figure 2. The numerical results indicate that the medium grid system (minimum grid size 0.004 m) achieved accurate results, with a near-wall $y^+$ value of 18.78 (ensuring $y^+ > 15$, which aligns with standard wall function requirements). The fine grid ($y^+$ = 10.75) and coarse grid ($y^+$ = 42.68) were less effective in balancing computational efficiency and accuracy. Consequently, the medium grid system was selected for subsequent simulations due to its optimal balance.

A comparison was conducted between the wind tunnel experimental data from CEDVAL B1-1 and the results of wind speed numerical simulations using a medium-scale grid system. As shown in Figure 3, the wind speeds obtained from numerical simulations using the medium-scale grid exhibit an overall trend consistent with those measured in the wind tunnel experiments. The difference between the two datasets is minimal, further confirming their good agreement. In Figure 3a, except for z = 0.025 and z = 0.1, the numerical simulation predicts wind speed more accurately. The numerical simulation in Figure 3b has a higher prediction of wind speed at y = 0.105 to y = 0.165, while the predictions of wind speed at other points are more accurate. The numerical simulation in Figure 3c has a lower prediction of wind speed at x = 0.03, while the predictions for wind speed at other points are higher than the predicted values. Taking all factors into consideration, the medium grid system accurately predicted the wind speed distribution of the building's wake.

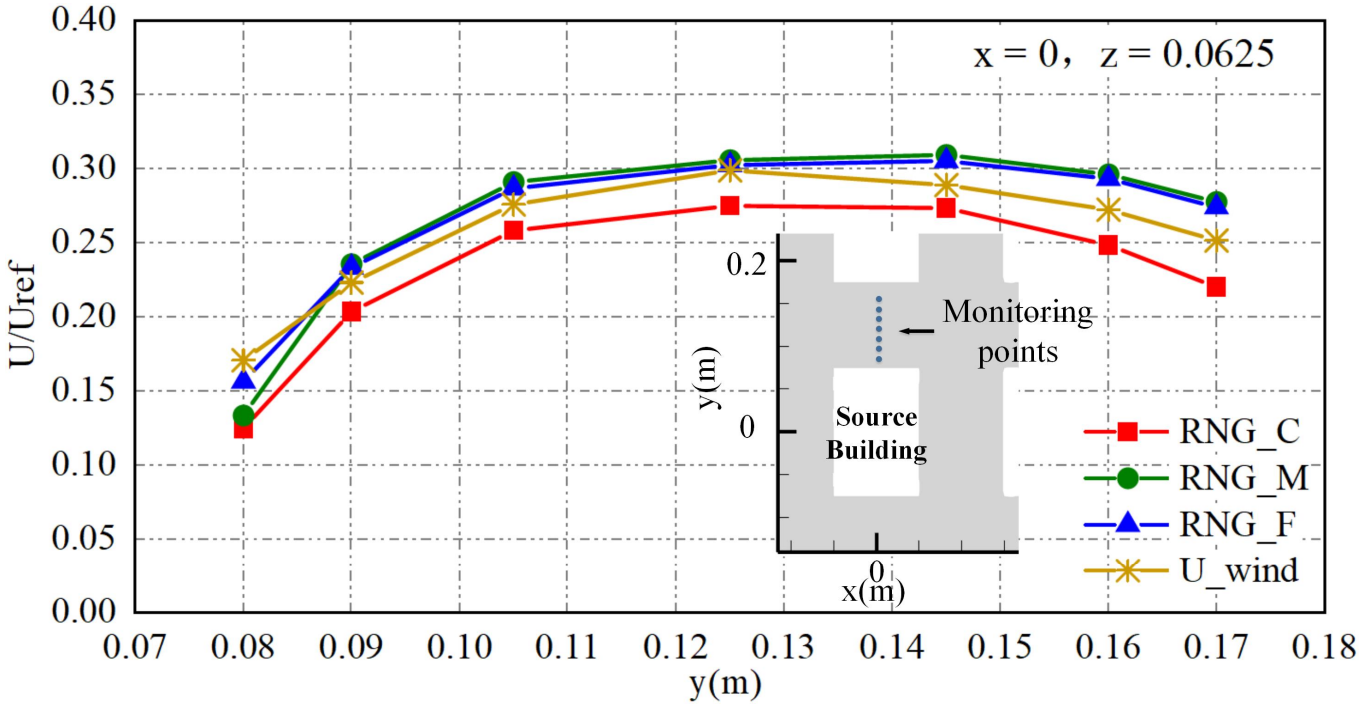
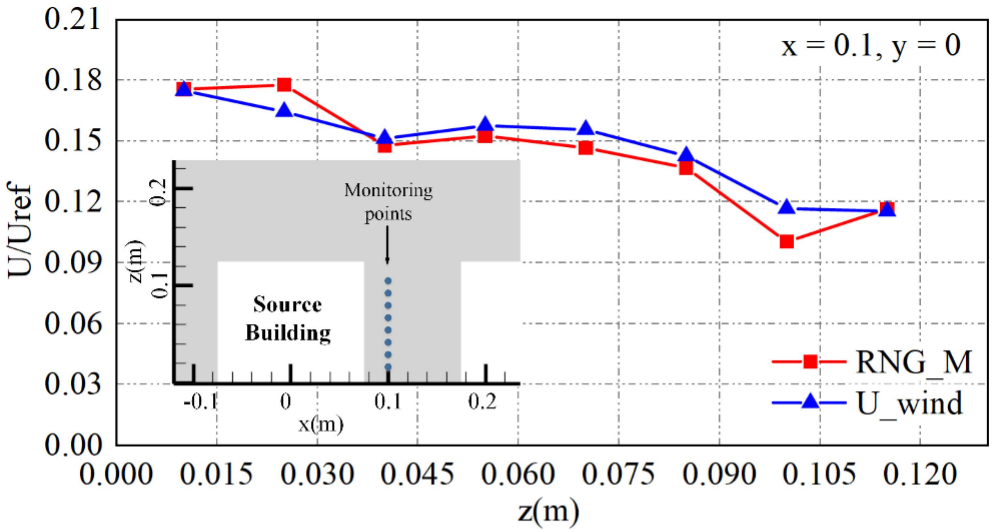


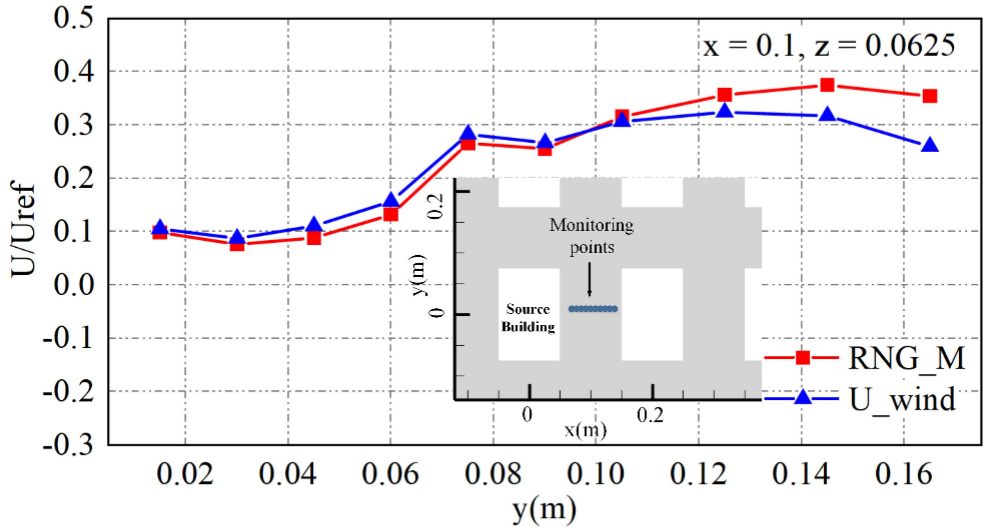

**Figure 2.** Comparison of wind tunnel data with simulated RNG $k-\varepsilon$ models of coarse (RNG_C), medium (RNG_M), and fine (RNG_F) grid systems.

The simulated $CO_2$ concentration distribution from the wind tunnel experiment is shown in Figure 4, with measurement points taken at a height of 2 m above ground level. As the distance from the release source increases, the discrepancy between the simulation results and the experimental data gradually decreases. At most measurement locations, the predicted concentrations are in good agreement with the corresponding wind tunnel measurements. As shown in Figure 4, the variation trend of predicted concentrations is consistent with that of wind tunnel experimental results. Thus, both qualitative trend analysis and quantitative numerical analysis confirm that the model constructed in this study meets the accuracy requirements for predicting wind fields and toxic gas concentration fields in urban environments, making it suitable for the current research.

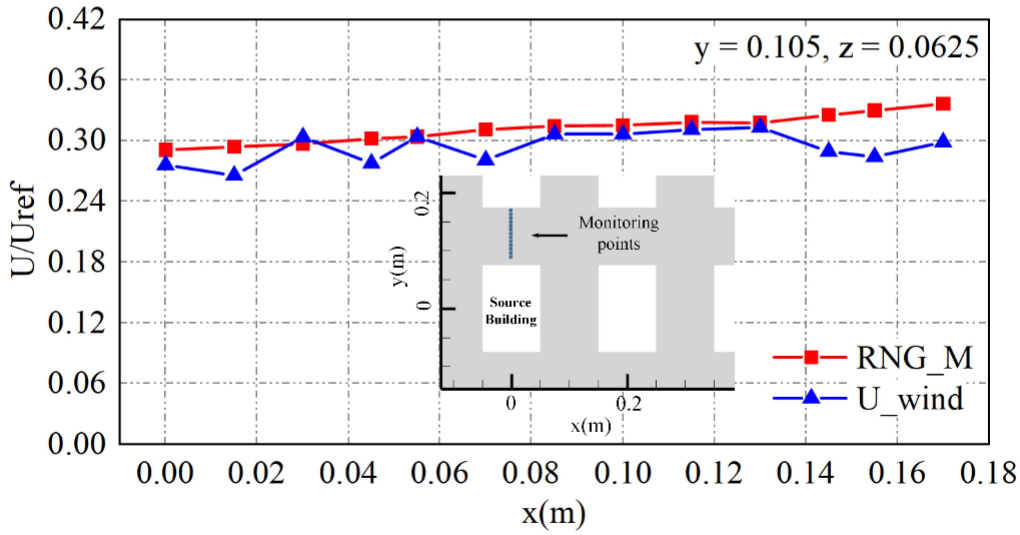
 https://doi.org/10.53941/ges.2026.100034




(a) Comparison of wind speeds at x=0.1 and y=0




(b) Comparison of wind speeds at x=0.1 and z=0.0625




(c) Comparison of wind speeds at y=0.105 and z=0.0625

**Figure 3.** Analysis of the average wind speed distribution at different sampling points.

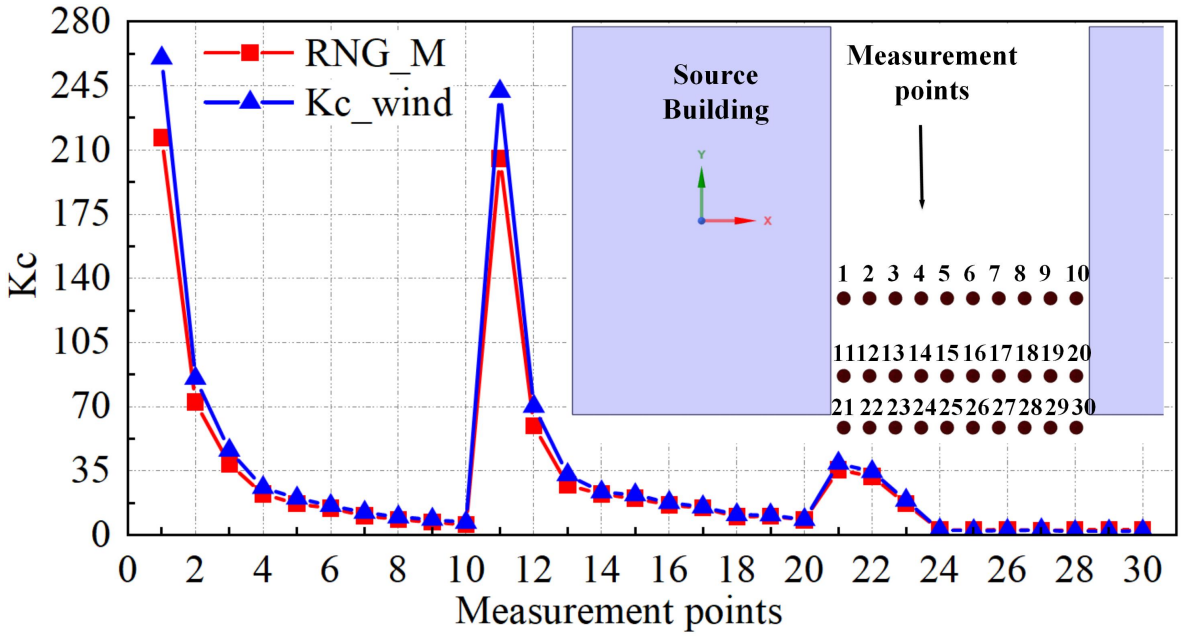



**Figure 4.** Comparison of concentrations at different monitoring points.

### *2.2. Graph Neural Network Architecture*

#### 2.2.1. Toxic Gas Diffusion Dataset Generation Considering Multiple Factors

This study selected a real urban area in Chaoyang District, Beijing, namely the Beijing Chemical Experimental Factory and nearby urban areas, as the research objects for numerical simulation. The simulation area spans 1020 m × 620 m, featuring diverse building structures, with the tallest building measuring 84 m. Figure 5a illustrates the layout, highlighting the factory (red dot), dense high-rise buildings to the east (green box), residential areas (red box), and a shopping center to the north (black dashed box). This area is densely populated, with numerous residential and activity zones, making it a significant and practical choice for the simulation. The fluid domain and boundary conditions for CFD simulation are shown in Figure 5b. The negative direction of the x-axis is the flow direction, the z-axis is the vertical direction, and the y-axis is the horizontal direction.

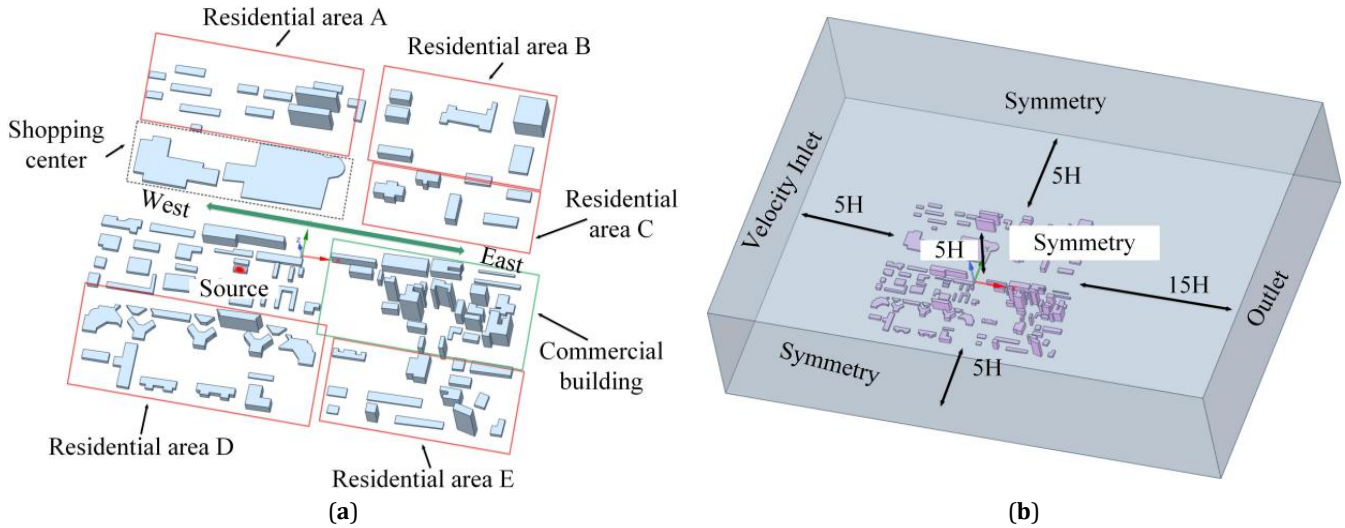



**Figure 5.** CFD simulation area and boundary condition settings. (**a**) CFD simulation area. (**b**) CFD simulates fluid domains and boundary settings.

The diffusionFoam solver was employed to generate the training dataset for the neural networks. The dataset was constructed using wind speed, wind direction, and temperature fields obtained from WRF simulations conducted on 11 October 2022, 22 December 2022, 28 April 2023, and 22 June 2023. These meteorological fields were subsequently used to define the inflow boundary conditions for the experimental settings involving the Beijing Chemical Plant and its surrounding urban area. The input parameters of the numerical simulations include wind speed ( $V$ ), wind direction ( $D_\theta$ ),

temperature ( $T$ ), leakage concentration ( $C$ ), and leakage rate ( $Q_v$ ). Wind speed included four different profiles generated from the WRF simulations, while wind direction was characterized by five angles: 0°, 85.48°, 162.87°, 98.74°, and 152.08°. The temperature variations comprised four values: 12.43 °C, −4.72 °C, 15.85 °C, and 32.85 °C. Additionally, leak concentrations were set at 0.1, 0.5, and 1 (mass fraction), and leak rates at 5 m/s, 10 m/s, and 15 m/s. A total of 360 scenarios were simulated, each consisting of 500 consecutive time steps. The training dataset comprised 80% of the data, while the remaining 20% was the validation dataset.

The data generated in this study is unique, as it is controlled by distinct initial conditions of wind speed ( $V$ ), wind direction ( $D_\theta$ ), temperature ( $T$ ), leak source concentration ( $C$ , represented as a mass fraction), and leak rate ( $Q_v$ ). As a result, it can be confirmed that none of the data in the validation set overlaps with the training set. Figure S4 illustrates nine sets of scenarios within the dataset. The nine scenarios shown in Figure S4 represent representative combinations of environmental conditions and leakage source parameters rather than the full cross-combination of all variables. It can be clearly observed from Figure S4 that the built environment remains identical across the nine cases, whereas the chlorine dispersion pattern varies with different inflow boundary conditions.

### 2.2.2. Data Preprocessing

This study used Multilayer Perceptron (MLP) to process training and prediction data, and the method flowchart is shown in Figure 6.

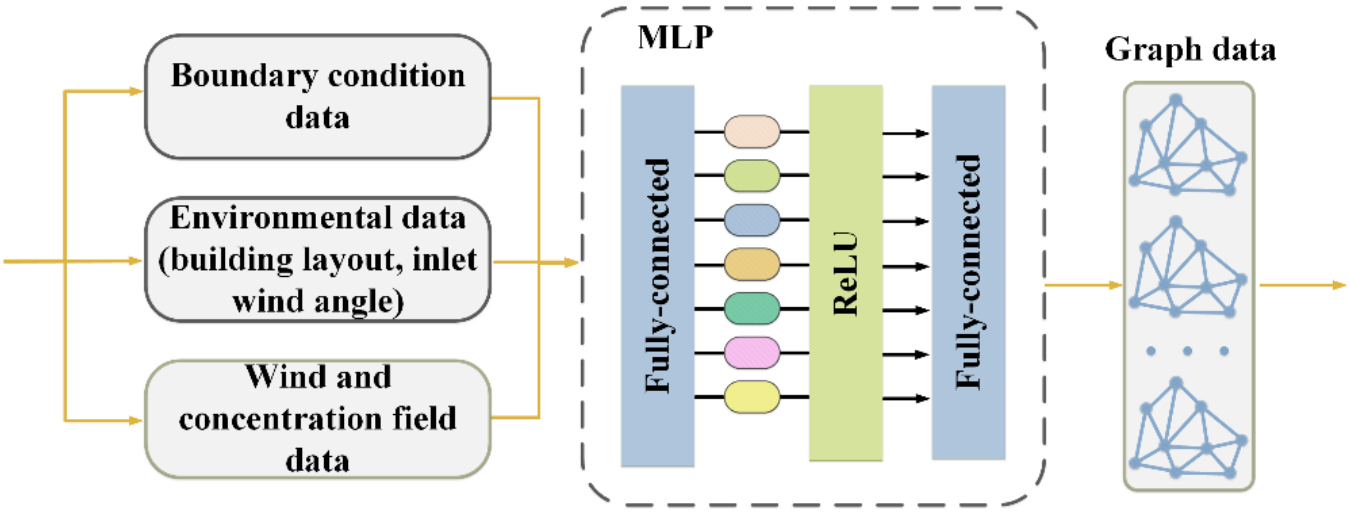



**Figure 6.** Data preprocessing flowchart.

This study employs MLP to address the various factors influencing the dispersion of toxic gases in urban environments. First, the factors affecting toxic gas dispersion are categorized as follows: (1) Wind field and toxic gas concentration data obtained through CFD numerical simulation; (2) Architectural environmental factors, including building layout, unstructured grid information generated by CFD, and wind direction conditions calculated using WRF to simulate wind speed; (3) Inflow boundary conditions obtained through WRF meteorological simulation, encompassing inflow wind profiles, inflow temperatures, and gas boundary condition information such as pressure ($p$), turbulent kinetic energy ($k$), and turbulent dissipation rate ( $\varepsilon$ ).

The MLP is employed to process different types of data. The different types of data are processed in parallel branches using dedicated MLPs to construct the specific components of the graph. Specifically, the field data obtained from CFD numerical simulations are processed by the MLP to generate graph-structured representations. The architectural layout information, namely the unstructured mesh data, is also processed by the MLP to construct corresponding graph-structured data. In addition, different labels are assigned according to the inflow boundary conditions to distinguish the meanings of various boundary conditions, and these labels are further processed by the MLP to generate graph-structured label data. Through these procedures, the data preprocessing is completed.

### 2.2.3. A Graph Neural Network Model Coupled with Multiple Factors

The neural network training dataset processed by MLP method is a graph-structured dataset. In this dataset, the dataset at time $t$ is represented by constructing a graph $G=(V,E)$, where each node $i \in V$ corresponds to a grid cell $C_i$, thus the size of the graph is $|V|=N$. The set of edges $E$ is defined by the adjacency relationships between grid cells. If two grid cells $C_i$ and $C_j$ are neighbors, both directional edges $(i,j)$ and $(j,i)$ are

included in the edge set. At each time step *t*, each node $G=(V,E)$ utilizes the local state vector $\mathbf{u}_{i,t}$ as its attribute. Consequently, the data for each scene in the dataset can be depicted as a sequential graph series $(G,(Y_0,...,Y_T,Y_{T+1}))$, and the objective of the MC-GNN model is to iteratively forecast the concentration field $(Y_0,...,Y_T,Y_{T+1})$.

The MC-GNN model utilized in this study employs an encoder-processor-decoder architecture [44], as illustrated in Figure 7. In particular, the method comprises three key components: encoder, processor, and decoder.

In this framework, the unstructured CFD mesh is represented as a graph structure. Each computational cell corresponds to a node, and the adjacency relationships between neighboring cells define the graph edges. Node features include gas concentration, wind velocity, and spatial coordinates. Among these, gas concentration and wind velocity are treated as dynamic state variables, while spatial coordinates are considered static within each simulation scenario. Edge features represent geometric relationships between connected cells. In addition, global environmental parameters, including meteorological conditions and leakage source characteristics, are incorporated into the node feature representation and shared across all nodes, enabling the model to capture global contextual information during message passing.

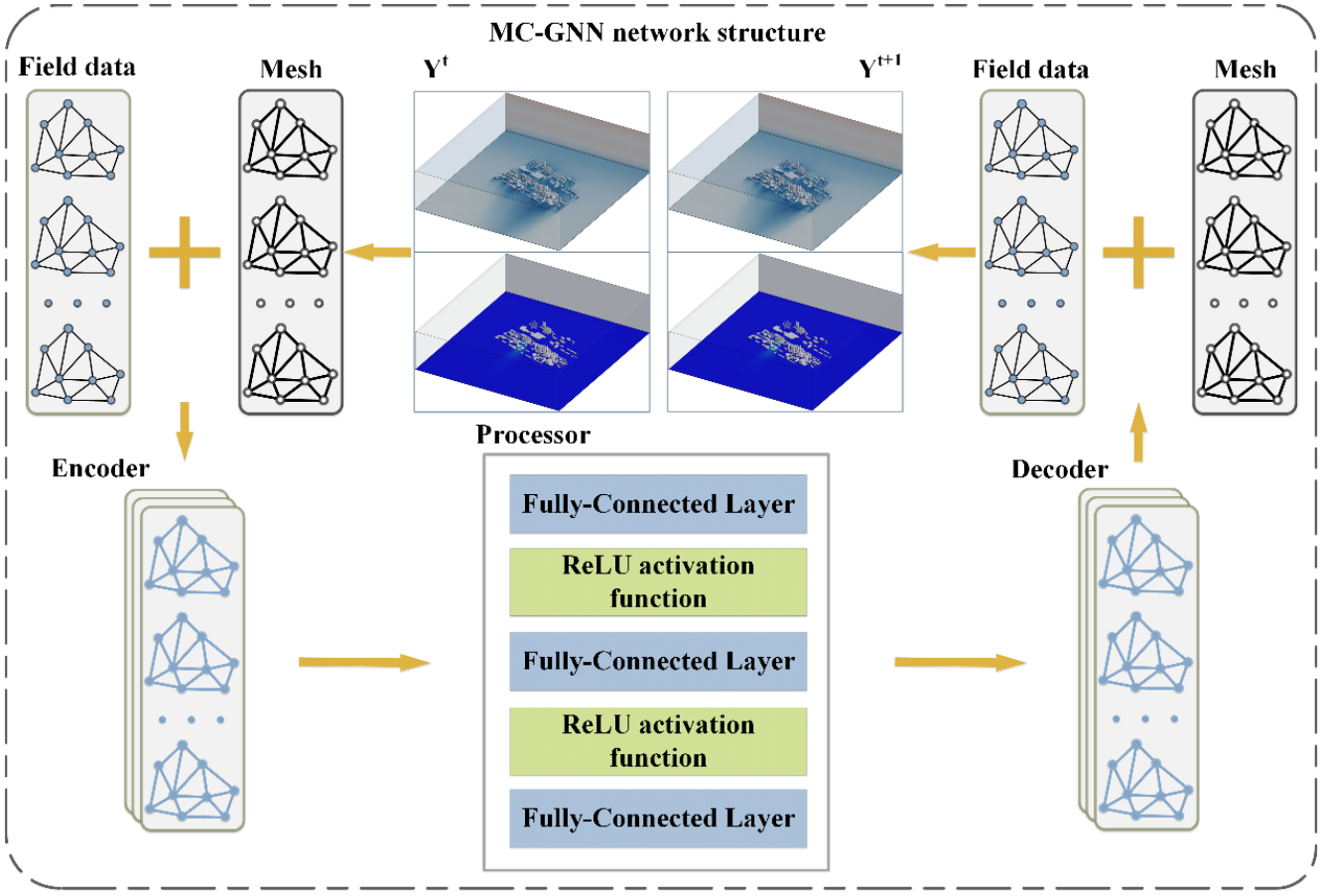



**Figure 7.** GNN network structure.

The encoder processes input data, such as wind speed fields, toxic gas concentration fields, building information, and inflow boundary conditions $Y_t$, into a graphical structure $G=(V,E)$. Where, $V$ represents the set of network nodes within the simulation model, while $E$ delineates the network's edges. Features are translated into graph nodes and edges: relative displacement vectors $\mathbf{u}_{ij}=\mathbf{u}_i-\mathbf{u}_j$ and their magnitudes $|\mathbf{u}_{ij}|$ are transformed into edge attributes $\mathbf{e}_{ij}\in E$, while any remaining dynamic characteristics $\mathbf{q}_i$ are encoded as node attributes $\mathbf{v}_{ij}\in\mathbf{V}$. The node state vector contains the physical variables that evolve during the dispersion process. In this study, the dynamic node variables include gas concentration and wind velocity components obtained from the CFD simulation results. Prior to model training, all physical variables are normalized using min–max normalization to ensure numerical stability during training. The normalized variables are then used as node attributes in the graph neural network. These attributes undergo encoding via node and edge encoders, respectively, utilizing MLP, each with a dimensionality of 128.

 https://doi.org/10.53941/ges.2026.100034

The processor is composed of L identical Message Passing Blocks (MPBs) [45]. This module depends on the output of the preceding adjacent block in the processor sequence to update the edge features $\mathbf{e}_{ij}$ within the network, and it updates the node features $\mathbf{v}_{ij}$ by aggregating all the updated edge features. This operation can be mathematically expressed as:

$$f^E\left(\mathbf{e}_{ij}, \mathbf{v}_i, \mathbf{v}_j\right) \rightarrow \mathbf{e}'_{ij} \quad (15)$$

$$f^V\left(\mathbf{v}_i, \sum_j \mathbf{e}'_{ij}\right) \rightarrow \mathbf{v}'_i \quad (16)$$

where $f^E$ and $f^V$ are implemented using MLP.

MPB updates node features and edge features by assembling graph data processed from node features, edge features, and edge indices. As shown in Figure 8, MPB processes graph data in two parts: first, the graph data is input into EdgeBlocks (EBs) for processing. Second, the graph data is input into NodeBlocks (NBs) for processing. The processor consists of multiple MPBs, with each MPB representing one layer of the processor. Each layer of the processor iterates on the basis of the previous layer, and residual connections are used between layers.

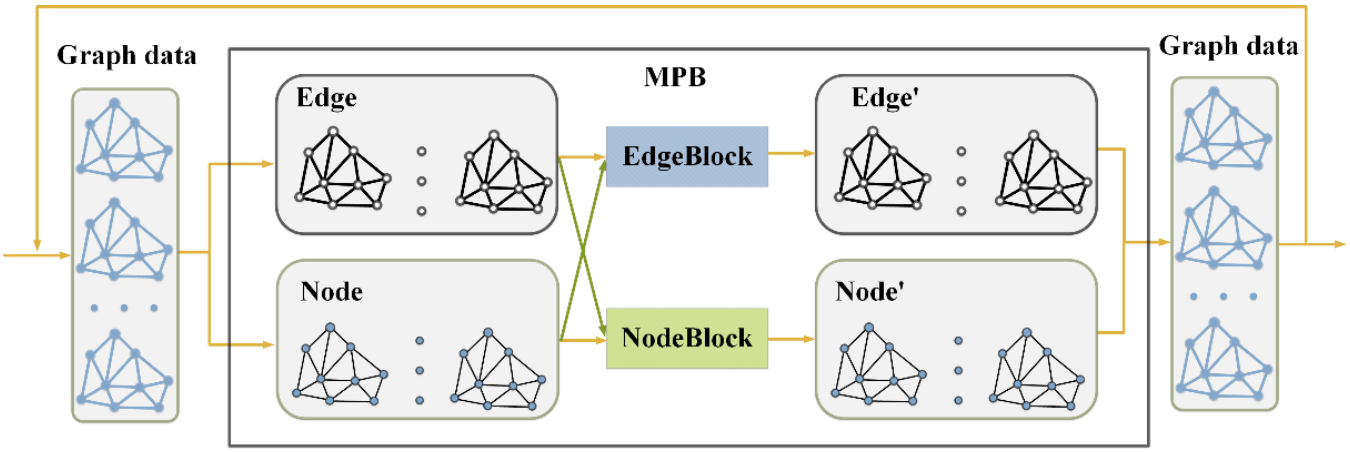



**Figure 8.** MPB Process Diagram.

The decoder employs MLP to convert processed latent node features $\mathbf{v}_i$ into one or multiple output features $\mathbf{p}_i$. These output features $\mathbf{p}_i$ can be seen as derivatives of $\mathbf{q}_i$, amalgamating several output features, and employing forward Euler integration of $\Delta t=1$ to compute the subsequent dynamic momentum $\mathbf{q}_i^{t+1}$, denoted as $\mathbf{q}_i^{t+1} = \mathbf{p}_i + \mathbf{q}_i^t$, ultimately generating the grid $M^{t+1}$ for the succeeding time point. During training, the MC-GNN model adopts a one-step supervised learning strategy. At each time step *t*, the model takes the current graph state as input and predicts the next state *t* + 1. The loss function is computed between the predicted state and the corresponding CFD simulation results. Specifically, the predicted state at time *t* + 1is fed back into the model as the input for the next step to obtain the state at *t* + 2. This process is repeated iteratively to generate long-term predictions of the toxic gas concentration field.

Although the training datasets in this study were generated from a representative urban area, the MC-GNN model incorporates boundary-condition and environmental features and is constructed on graph representations of unstructured meshes. Therefore, the framework has potential for adaptation to other urban layouts with similar flow characteristics. However, when the model is applied to cities with substantially different urban morphologies, such as different building densities or street configurations, its prediction accuracy may be affected if these geometrical and environmental features are not sufficiently represented in the training dataset. For deployment in new environments, partial retraining or fine-tuning using additional CFD-generated data can be applied, significantly reducing the computational cost compared with building a full new CFD dataset.

### *2.3. Exposure Analysis*

#### 2.3.1. Exposure Risk Calculation Based on Dose-Response Model

The dose-response model is often used to analyze the consequences of acute exposure. The concentration of toxic gas and exposure time are two important variables for the dose-response model. This study calculated the toxic dose based on the concentration of toxic gases at the measurement points and exposure time and estimated the mortality rate using a dose-response model.

The Probit equation can be used to quantify exposure levels to various toxic substances in the air. The Probit variable Y is defined as follows.

$$Y = \mathrm{A} + \mathrm{B}\ln(\mathrm{D}) \quad (17)$$

Among them, A and B are constants, with values of −5.3 and 0.5, respectively, $\mathrm{D}$ representing the toxic

dose. From the beginning of the leak to any point in time, the toxic exposure dose $D$ at each monitoring point is shown in Formula (18):

$$D=\int_{tend}^{t0} C^n dt \approx \sum_{i=1}^{k} C^n \cdot \Delta t_i \quad (18)$$

where $C$ is the concentration (in ppm or mg/m$^3$), $n$ is the concentration index, and the concentration index of chlorine gas is 2.75; $t$ is the exposure time, and $i$ is the time increment.

The Probit variable is related to the mortality rate, so the mortality rate $P$ of chlorine gas at a specific concentration can be calculated:

$$P=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{Y-5}\exp(-\frac{x^2}{2})dx \quad (19)$$

where $x$ is the integral variable. This study investigated the mortality rate of pedestrians at a height of 1.6 m, identified areas with a mortality rate exceeding 40% as high-exposure risk areas, and conducted a high-exposure risk analysis. The steps for conducting a high exposure risk analysis are the following: (1) Divide the simulated area into sub-areas and calculate the regional exposure risk; (2) Calculate the size of high exposure risk areas based on the obtained regional exposure risks, and conduct high exposure risk analysis.

### 2.3.2. Horizontal Distance Analysis of High Exposure Risk

In this study, the lateral distance of high-exposure risk areas is analyzed. The position that is farthest from the release source and has a fatality rate of 40% at 500 time steps is termed the diffusion edge, and the longest distance of the lateral diffusion edge is defined as the lateral distance of the high-fatality region, abbreviated as "$L_{HFR}$". This study analyzes the lateral distance of pedestrian height (1.6 m).

Dimensional analysis is employed to establish the relationships among the different physical quantities, thereby providing both qualitative insights and quantitative descriptions of the associated physical phenomena. In the International System of Units, fundamental physical quantities consist of length ($L$), mass ($M$), time ($V$), electric current ($I$), thermodynamic temperature ($\theta$), amount of substance ($N$), and luminous intensity ($J$), representing the fundamental dimensions. Besides these seven base dimensions, the dimensions of other physical quantities can be derived by combining these base dimensions, termed as derived dimensions.

Commonly used dimensional analysis methods include the Rayleigh method and Buckingham's theorem, also known as the π theorem, with the π theorem being more general in applicability. The $\pi$ theorem is particularly valuable for analyzing the structural relationships among various physical quantities. Accordingly, the $\pi$ theorem is employed in this study to derive the formula for calculating horizontal distances. For a particular physical phenomenon, if there are $n$ variables $x$ related to each other as functions, it can be expressed as:

$$f\left(x_1,x_2,...,x_n\right)=0 \quad (20)$$

if there is $m$ independent fundamental physical quantity, the physical process can be expressed by $n-m$ dimensionless numbers to establish a new functional relationship, as follows:

$$F\left(\pi_1,\pi_2,...,\pi_{n-m}\right)=0 \quad (21)$$

In this study, the $L_{HFR}$ formulation is derived through dimensional analysis, incorporating a total of ten physical quantities, including: the distance from the leak source along the diffusion direction to the building edge ($\varphi$), lateral distance ($L_{HFR}$), wind speed ($V$), wind direction ($D_\theta$), temperature ($T$), leak source concentration ($C$, expressed as mass fraction), leak rate ($Q_v$), dynamic viscosity ($\mu$), thermal conductivity ($\lambda$), and gravitational acceleration ($g$). The dimensionless relationship can be expressed as follows:

$$f\left(V,T,\varphi,\mu,L_{HFR},Q_v,g,\lambda,C,D_\theta\right) \quad (22)$$

where $V$, $T$, $\varphi$, and $\mu$ represent fundamental physical quantities, $C$ and $D_\theta$ denote dimensionless quantities. Other physical quantities can be expressed in terms of these fundamental physical quantities. Thus, the relationship can be deduced from $\pi$ Theorem as follows:

$$F\left(\pi_1,\pi_2,\pi_3,\pi_4,\pi_5,\pi_6,\pi_7,\pi_8,\pi_9,\pi_{10}\right)=0 \quad (23)$$

where $\pi_1=\pi_2=\pi_3=\pi_4=1$, $\pi_9=C$, $\pi_{10}=D_\theta$, and $\pi_5,\pi_6,\pi_7,\pi_8$ represent other physical quantities. The relationship between $\pi_5,\pi_6,\pi_7,\pi_8$ can be derived from $\pi$ theorem as follows:

$$\begin{cases}\pi_5=\dfrac{L_{HFR}}{\varphi}\\ \pi_6=\dfrac{Q_v}{V}\\ \pi_7=\dfrac{g}{V^2\varphi^{-1}}\\ \pi_8=\dfrac{\lambda}{\mu V^2T^{-1}}\end{cases} \quad (24)$$

Through transformation, the result relationship is as follows:

$$\frac{L_{HFR}}{\varphi}=a\times\left(\frac{Q_v}{V}\right)^b\times\left(\frac{g}{V^2\varphi^{-1}}\right)^c\times\left(\frac{\lambda}{\mu V^2T^{-1}}\right)^d\times C^e\times D_\theta{}^f \quad (25)$$

where $a, b, c, d$ and $f$ are empirical constants determined through regression analysis.

#### 2.3.3. Analysis of Factors Influencing High Exposure Risk Lateral Distance

Based on the MC-GNN predictions, the diffusion patterns of chlorine gas under various inflow boundary conditions are fitted. The coefficients obtained from fitting the simulated data are summarized as follows:

$$\begin{cases} a = 0.08 \\ b = 0.07 \\ c = -2.19 \\ d = 1.52 \\ e = 0.58 \\ f = -0.07 \end{cases} \tag{26}$$

By substituting the coefficients into Formula (26), an exponential formula can be obtained, as follows:

$$\begin{cases} L_{HFR} = 0.08 \times V^{1.27} \times T^{1.52} \times \varphi^{-1.19} \times \mu^{-1.52} \times Q_v^{\ 0.07} \times g^{-2.19} \times \lambda^{1.52} \times C^{0.58} \times D_\theta^{\ -0.07} \\ R^2 = 0.96 \end{cases} \tag{27}$$

where $V$ represents the average wind speed at a height of 1.6 m above the inlet, calculated based on the wind profile. $T$ denotes the average temperature at the inlet, while $\varphi$ stands for the distance from the leak source along the dispersion direction to the building edge. $\mu$ refers to dynamic viscosity, and $Q_v$ indicates the leakage velocity at the leak point (measured in m/s). $g$ represents the gravitational acceleration, and $\lambda$ signifies the thermal conductivity coefficient. $C$ represents the concentration of the leak source expressed in mass fraction. $D_\theta$ denotes the wind direction angle, and $L_{HFR}$ stands for the lateral distance. Various factors influence the lateral distance of high fatality areas. Formula (27) illustrates that among the fundamental physical quantities, wind speed $V$ and temperature $T$ significantly affect the lateral distance of high fatality zones.

## 3. Results and Discussion

### *3.1. Comparative Analysis of MC-GNN Prediction and CFD Simulation*

This section compares the MC-GNN model's predictions with CFD simulation results for the Beijing Chemical Plant and surrounding urban areas. Figure S5 shows 90 monitoring points distributed across the simulation area, with A depicting vertical heights and B illustrating horizontal positions. Points are grouped by height: 1.6 m (1–30), 10 m (31–60), and 20 m (61–90). These points are strategically placed near emission sources, densely populated areas, and open spaces to assess the impact of different layouts and environmental conditions on toxic gas diffusion.

This study compares the MC-GNN model's predictions with CFD simulation results for concentration cloud maps and concentration values at monitoring points, under four wind direction scenarios: 85.48°, 162.87°, 98.74°, and 152.08°. Figures 9, 11, 13 and 15 depict the comparison of concentration cloud maps at heights of 1.6 m, 10 m, and 20 m in the simulation area at 50 s for each scenario. Identical color scales are applied to both the GNN and CFD plots to facilitate a direct visual comparison. Meanwhile, Figures 10, 12, 14 and 16 present concentration comparisons at monitoring points.

As shown in Figure 9, when the wind direction is 85.48°, the MC-GNN model shows reduced prediction accuracy at the edges of the chlorine diffusion, particularly in the regions marked by the black boxes. This is primarily because the concentrations near the diffusion boundary are extremely low, with most values below 0.06. Such low values are highly sensitive to minor variations in flow patterns and numerical noise, making them inherently difficult to predict and consequently leading to relatively lower prediction accuracy in these regions. In addition, the plume-edge region usually contains sparse non-zero concentration values, and its contribution to the overall training loss may be relatively limited. As a result, small deviations in the predicted flow direction or plume position can lead to visible differences near the diffusion boundary. In the central part of the dispersion zone at heights of 10 m and 20 m, the MC-GNN model overestimates the chlorine concentration, which indicates that its predictive accuracy is also limited in high concentration regions. However, in most areas of the simulation domain, especially at the height of 1.6 m, the MC-GNN model developed in this study provides relatively accurate predictions.

As shown in Figure 10, the concentration comparisons at the local monitoring points are consistent with the concentration contours presented in Figure 9. In the high-concentration central region, such as at points 3, 4, and 5, as well as in the low-concentration boundary areas, the MC-GNN model provides less accurate predictions of chlorine concentration. The maximum deviation at the less accurate points is 0.097. These results indicate that the MC-GNN model generally delivers reliable predictions at the monitoring points, with the difference between the CFD-simulated concentration and the MC-GNN prediction remaining within 0.097.

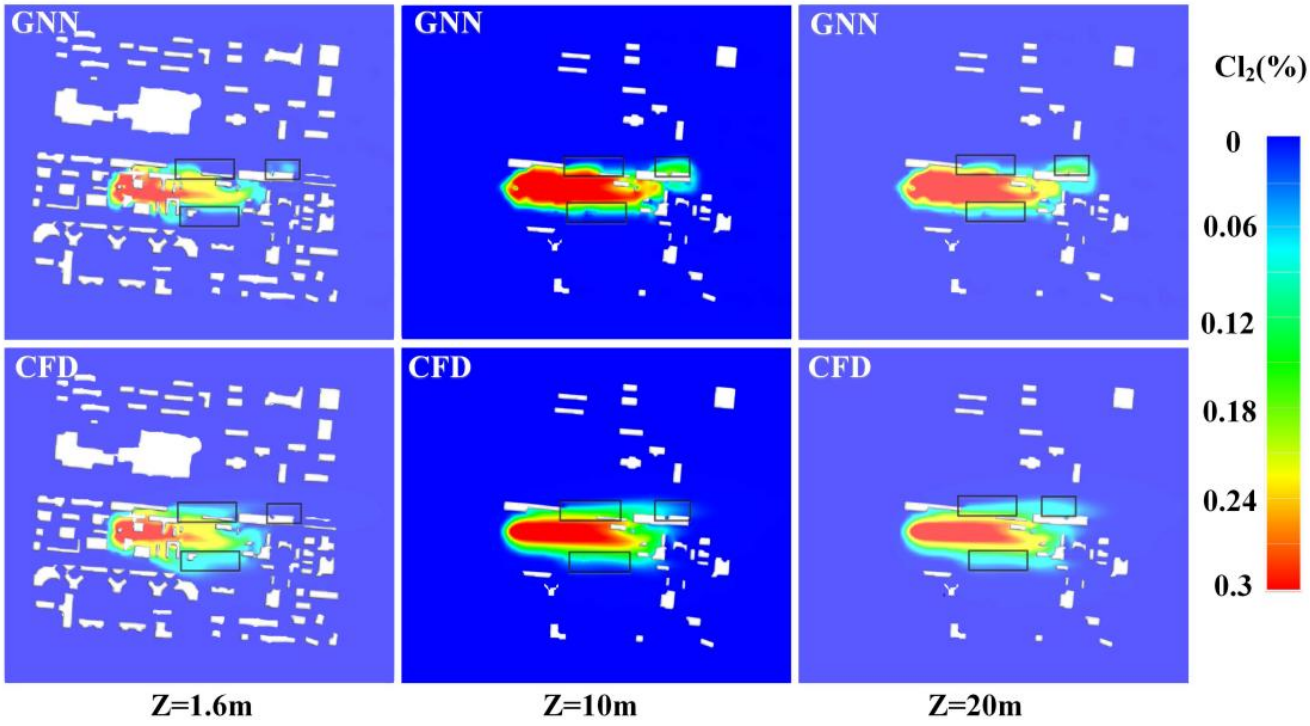



**Figure 9.** Comparison of concentration cloud maps with wind direction of 85.48°.

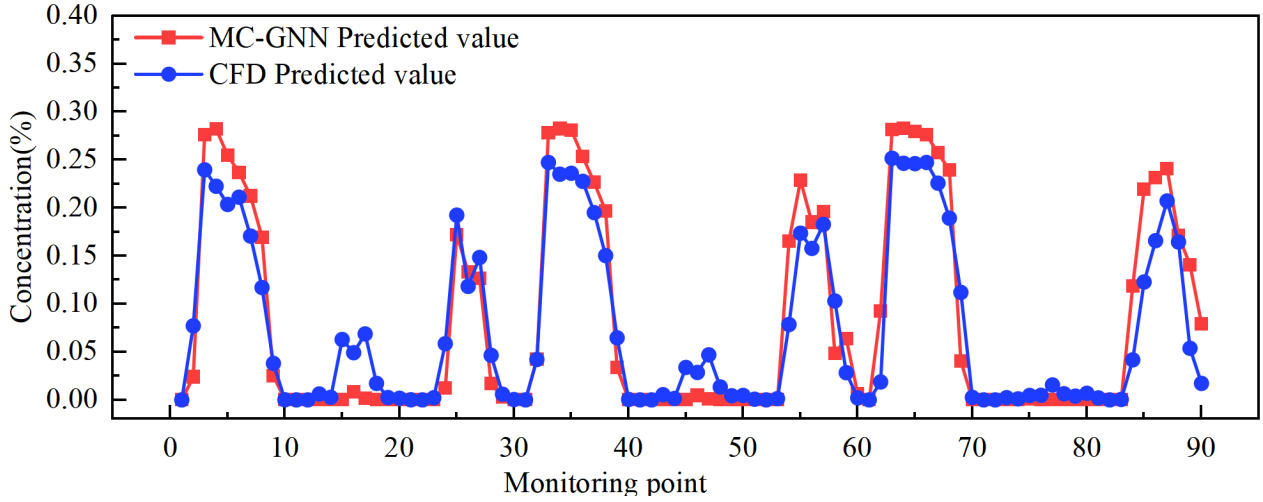



**Figure 10.** Comparison of concentrations at monitoring points with wind direction of 85.48°.

As shown in Figure 11, when the wind direction is 162.87°, the MC-GNN model provides less accurate predictions in both the edge and central regions of chlorine diffusion, as indicated by the black boxes in the figure. In particular, at a height of 1.6 m, in the lower black box region, building obstructions reduce the local chlorine concentration, resulting in less accurate predictions by the MC-GNN model. At heights of 10 m and 20 m, in the lower black box region, with only a single tall building causing obstruction, chlorine diffuses with the wind and is less influenced by the buildings. However, the MC-GNN model has not fully captured the diffusion pattern of chlorine in this area, resulting in predictions that are less accurate than those at a height of 1.6 m. Overall, across most parts of the simulation domain, the MC-GNN model employed in this study achieves relatively accurate predictions.

According to Figure 12, the concentration predicted by MC-GNN at the monitoring points closely resembles the concentration simulated by CFD. Notably, at the monitoring point with a height of 1.6 m, the numerical simulation from CFD and the prediction results from MC-GNN exhibit minimal disparity. However, at the heights of 10 m and 20 m, the concentration comparisons at monitoring points 54, 55, and 84 reveal significant differences, with the maximum variance reaching 0.036. Consequently, the MC-GNN model's prediction at the 1.6 m height proves to be more accurate. Overall, with a maximum difference of 0.036, the MC-GNN model provides precise simulation for this scenario.

 https://doi.org/10.53941/ges.2026.100034

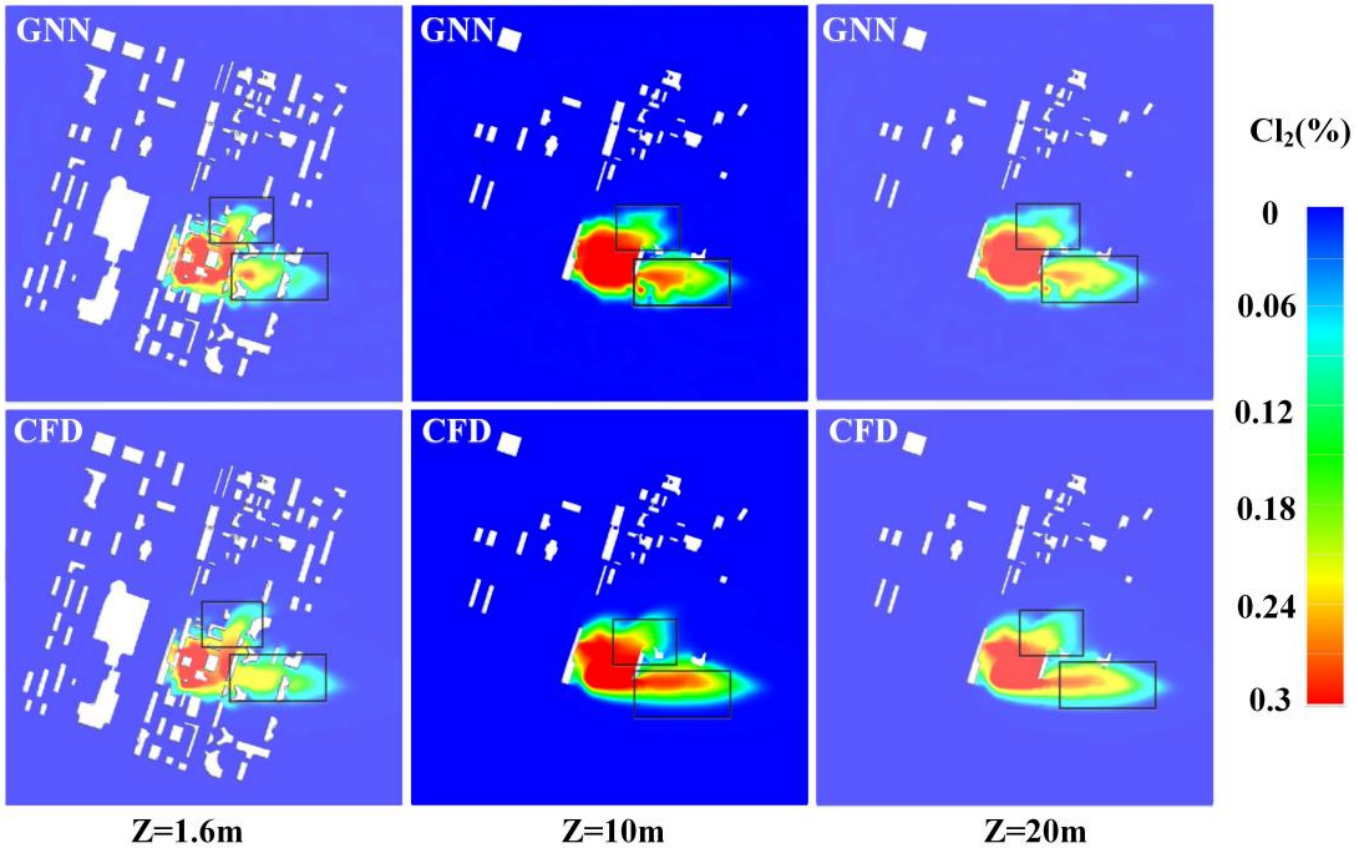



**Figure 11.** Comparison of concentration cloud maps with wind direction of 162.87°.

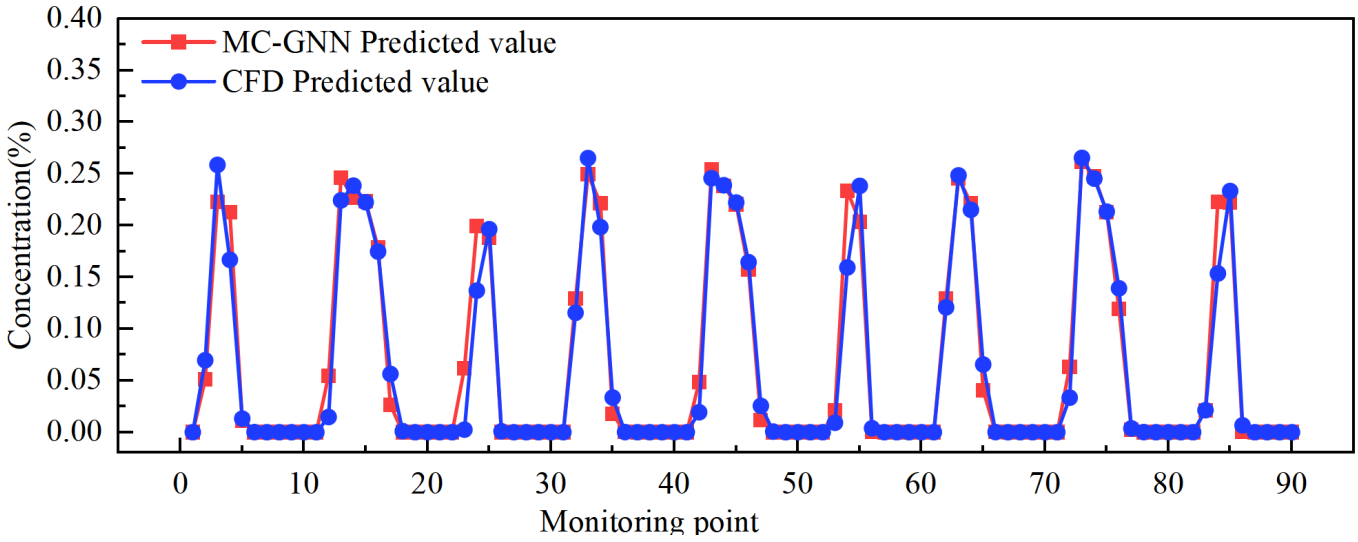



**Figure 12.** Comparison of Concentration at Monitoring Points with Wind Direction of 162.87°.

A comparison between the chlorine diffusion predicted by the MC-GNN model and the CFD simulations for a wind direction of 98.74° is shown in Figure 13. Chlorine primarily diffuses with the wind; however, the CFD results indicate limited diffusion along the east–west streets despite higher wind speeds. The MC-GNN model, on the other hand, overestimates the influence of wind, predicting higher concentrations in the upper black-box region, corresponding to the east–west street area, particularly at heights of 10 m and 20 m compared to 1.6 m. A similar overestimation is observed in the lower black-box region, where sparse buildings and elevated wind speeds contribute to discrepancies. Nevertheless, across most of the simulation domain, the MC-GNN model provides relatively accurate predictions.

Based on the observations from Figure 14, it can be inferred that at a height of 1.6 m, monitoring points 8, 9, 28, and 29 are located at the edge of the diffusion area, where the MC-GNN model predicts relatively low concentrations. In contrast, monitoring points 13 and 14

are situated in the open area beneath the lower black box in Figure 13, and points 24, 25, 26, and 27 are near the east–west street zone, where the MC-GNN model predicts higher concentrations. At heights of 10 m and 20 m, substantial discrepancies are observed for monitoring points 55–60 and 84–88, which lie within the black box region of Figure 13, with the MC-GNN model predicting higher values and a maximum difference of 0.16. Overall, these results suggest that the MC-GNN model tends to overestimate the influence of wind speed on chlorine gas concentration, leading to less accurate predictions in areas with higher wind speeds. Nevertheless, in other regions, the model provides comparatively accurate predictions.

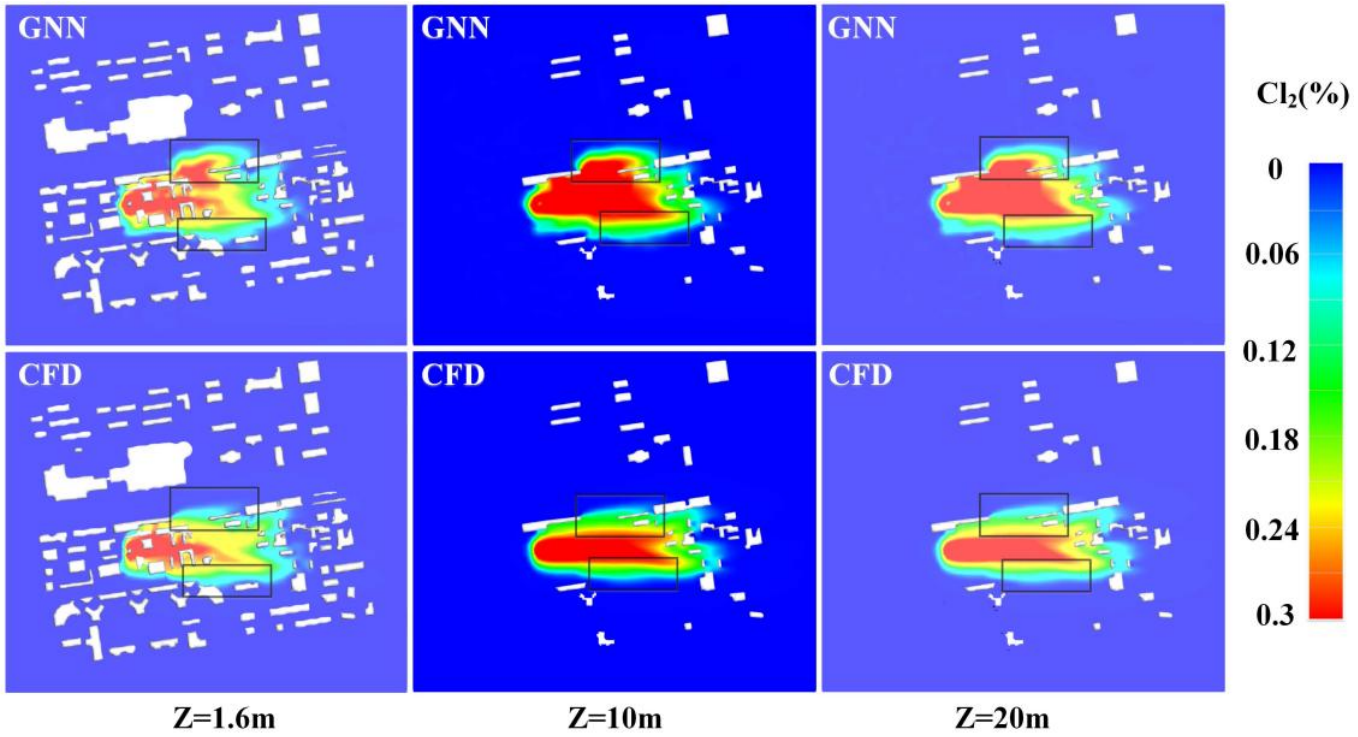



**Figure 13.** Comparison of concentration cloud maps with wind direction of 98.74°.

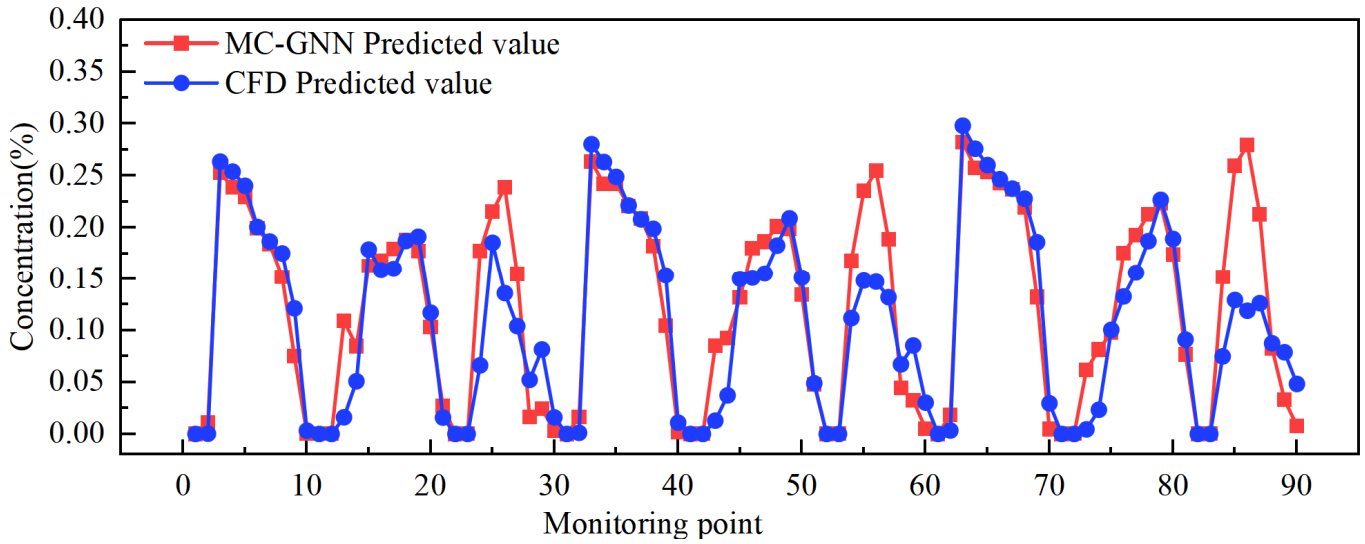



**Figure 14.** Comparison of Concentration at Monitoring Points with Wind Direction of 98.74°.

In Figure 15, the MC-GNN model shows inaccuracies in predicting chlorine concentrations near the diffusion periphery beneath the black box, where it struggles with lower concentrations. However, it performs better in predicting higher concentrations at the central positions within the diffusion area. In the red box, where buildings obstruct the diffusion, the model fails to fully capture the reduction in chlorine concentrations, leading to inaccuracies. Despite these limitations, the model demonstrates a high level of accuracy across most of the simulated areas

The chlorine concentrations predicted by the MC-GNN model, as shown in Figure 16, generally align closely with those obtained from the CFD simulations. However,

at monitoring points located in high-concentration zones, specifically in the central region of the diffusion area, the MC-GNN model predicts higher chlorine concentrations, consistent with the concentration contour map shown in Figure 15. At monitoring points 24 and 25, which are located between two buildings where wind speeds are low and chlorine tends to accumulate, the MC-GNN model underestimates the concentration, highlighting its limitations in simulating diffusion under such conditions. The maximum discrepancy between the MC-GNN and CFD models at these points is 0.13, whereas the differences at other monitoring points are minimal. Overall, the MC-GNN model provides relatively accurate predictions across most of the simulation domain.

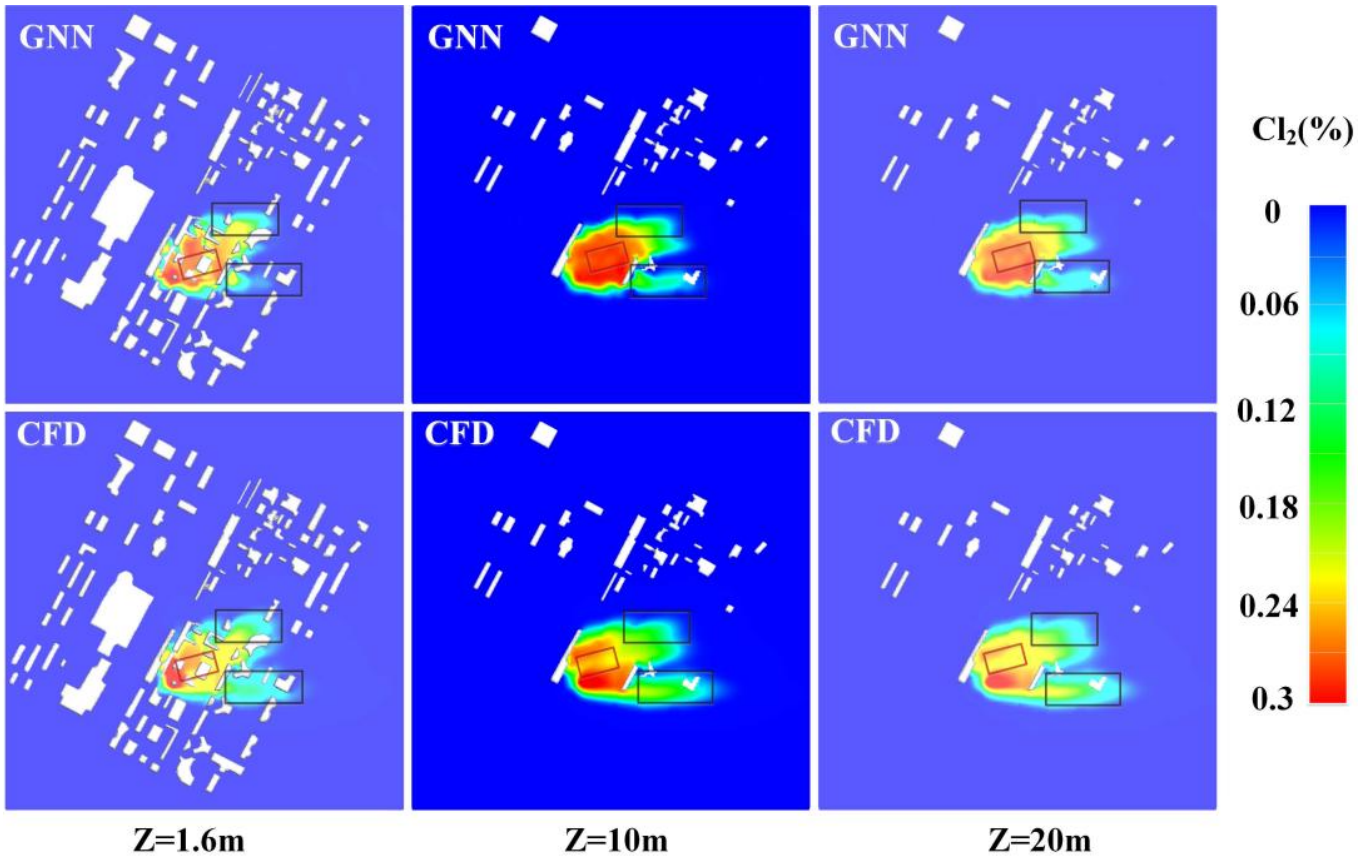



**Figure 15.** Comparison of concentration cloud maps with wind direction of 152.08°.

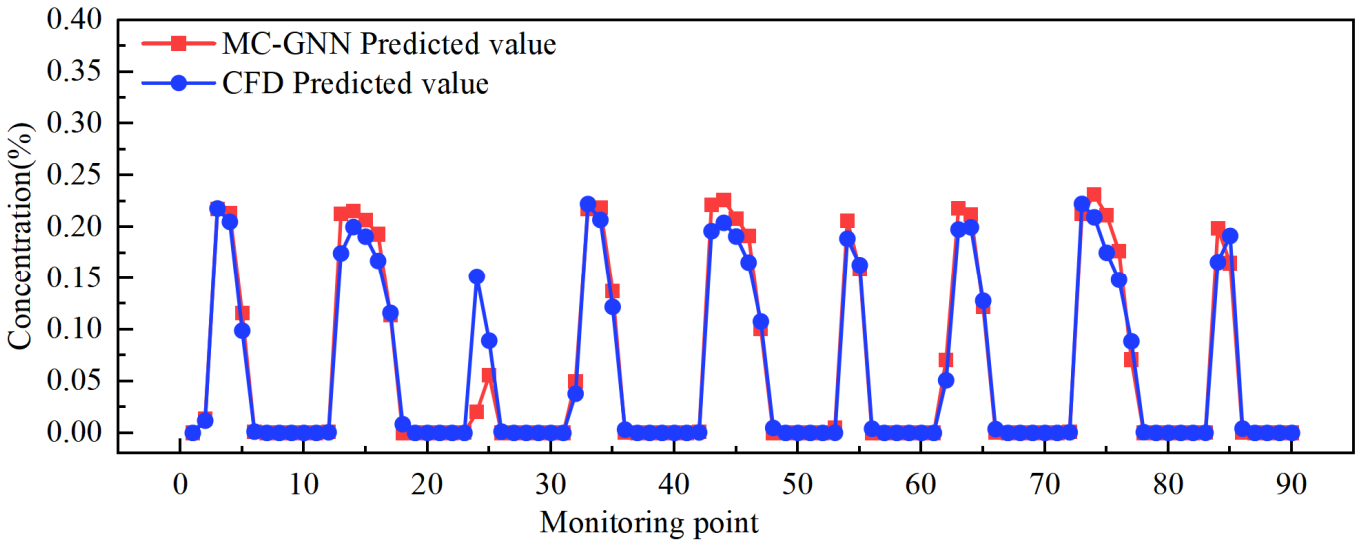



**Figure 16.** Comparison of Concentration at Monitoring Points with Wind Direction of 152.08°.

A comparison of MC-GNN predictions with CFD simulation results across four scenarios indicates that the MC-GNN model achieves higher accuracy for wind directions of 162.87° and 152.08°, while its accuracy decreases for wind directions of 85.48° and 98.74°. This is because, at wind directions of 85.48° and 98.74°, the wind

is nearly aligned with the orientation of the buildings, reducing the obstruction effect of the buildings on chlorine diffusion. As a result, chlorine disperses more rapidly and over a wider area. Furthermore, in the case of a wind direction of 98.74°, the near-surface inflow wind speeds are relatively low. Statistical analysis indicates that this type of data is relatively underrepresented in the training dataset, leading to lower prediction accuracy. Future work will focus on enriching the training dataset by incorporating a more diverse set of simulation cases to improve model robustness.

Overall, the MC-GNN model underperforms in regions with extremely high or low concentrations. Specifically, for wind directions of 85.48° and 98.74°, the high concentration areas are more extensive, which further challenges the predictive capability of the model. At the monitoring points, the maximum deviation of chlorine gas concentration between the MC-GNN model and the CFD simulation is 0.16, with most deviations falling below 0.05. This indicates that, despite some limitations, the MC-GNN model delivers relatively accurate predictions for both vertical and horizontal chlorine gas dispersion patterns.

### *3.2. Prediction Accuracy and Speed Evaluation of Toxic Gas Concentration Field*

In this study, the performance of the proposed MC-GNN model is evaluated using the root mean squared error (*RMSE*) and the $R^2$. The corresponding computational formulations for *RMSE* and $R^2$ are given as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_i (\hat{y}_i - y_i)^2} \tag{28}$$

$$R^2 = 1 - \frac{\sum_i (\hat{y}_i - y_i)^2}{\sum_i (\bar{y}_i - y_i)^2} \tag{29}$$

where $\hat{y}_i$ represents the predicted value of the node variable, $y_i$ stands for the true value of the node variable, and $\bar{y}_i$ denotes the average of the true values of the node variables. *RMSE* ranges from 0 to positive infinity, with smaller values indicating better predictive performance of the model. $R^2$ ranges from 0 to 1, with values closer to 1 indicating better predictive performance of the model.

In Figure 17, the results of *RMSE* and $R^2$ for monitoring points and the entire simulation area at the 500th time step (50 s) are depicted. It is evident from the figure that the $R^2$ values of the monitoring points are higher than those of the entire simulation area, especially when the wind direction is 162.87° and 152.08°, where the disparity in $R^2$ between the monitoring points and the entire simulation area is more pronounced. This is because the monitoring points we selected are uniformly distributed near the diffusion source, with more points located to the right of the leakage source. Consequently, when the wind direction is 85.48° and 98.74°, the monitoring points are evenly distributed in the diffusion area, and the *RMSE* and $R^2$ values of the selected monitoring points closely approximate those of the entire simulation area. Thus, in these wind directions, the analysis of the monitoring points can effectively substitute for the analysis of the entire simulation area. However, for wind directions of 162.87° and 152.08°, owing to the wind direction, the chlorine gas diffuses towards the lower right side of the leakage source, leading to the uneven distribution of the selected monitoring points in the diffusion area, resulting in an increased number of monitoring points with a chlorine gas concentration of 0, thereby causing higher $R^2$ values for the monitoring points. Ultimately, statistical analysis reveals that the *RMSE* of the monitoring points remains within the range of 0.018 to 0.045, the $R^2$ between 0.74 and 0.96; while the *RMSE* of the entire simulation area ranges from 0.027 to 0.051, and the $R^2$ from 0.73 to 0.92.

This article evaluates and analyzes the prediction speed of the MC-GNN model, primarily by comparing the computational efficiency of the MC-GNN model, the GNN + Self-Attention (GNN-SA) model, and the CFD model in predicting a single time step. The experimental setup includes a GPU (NVIDIA RTX 4090, 24 GB memory; NVIDIA Corporation, Santa Clara, CA, USA) and a CPU (Intel I9-13900KF; Intel Corporation, Santa Clara, CA, USA). After analysis, the comparison of efficiency using the MC-GNN model, GNN-SA model, and simulating with the CFD model is illustrated in Figure 18. The figure displays the average computation time of a single time step (unit: ms) under four groups of simulation experiments with different inflow conditions and grid resolutions, as well as the acceleration ratio of the MC-GNN model and GNN-SA model compared to the CFD model in the four groups of simulation scenarios. Figure 18 indicates that the average acceleration ratio of the GNN-SA model in the four groups of simulation scenarios is 1.79, and the average acceleration ratio of the MC-GNN model in the four groups of simulation scenarios is 5.45. This is due to the GNN-SA model utilizing an attention mechanism to learn the relationship weights between nodes, which entails calculating attention coefficients between nodes and their neighboring nodes during the information aggregation process at each node. Particularly in scenarios with large graph structures and high node degrees, this will increase time consumption. In summary, the proposed MC-GNN model demonstrates superior speed, with the MC-GNN model improving the speed by almost an order of magnitude compared to the CFD model, making it more suitable for predicting the diffusion of toxic gases in urban environments with multiple coupled factors.

It should be noted that the MC-GNN model requires an offline training phase using CFD-generated datasets.

 https://doi.org/10.53941/ges.2026.100034

Once trained, the model can perform inference rapidly, significantly reducing computational time compared with conventional CFD simulations, which is particularly beneficial for emergency response applications. For deployment in cities with distinctly different urban morphologies, additional CFD simulations may be required for transfer learning or partial retraining to maintain high prediction accuracy. This approach substantially reduces the computational cost compared with constructing an entirely new CFD dataset. In future work, the transferability of the proposed MC-GNN framework across diverse urban morphologies will be systematically investigated, including studies on transfer learning strategies and the minimum amount of CFD data required for model adaptation.

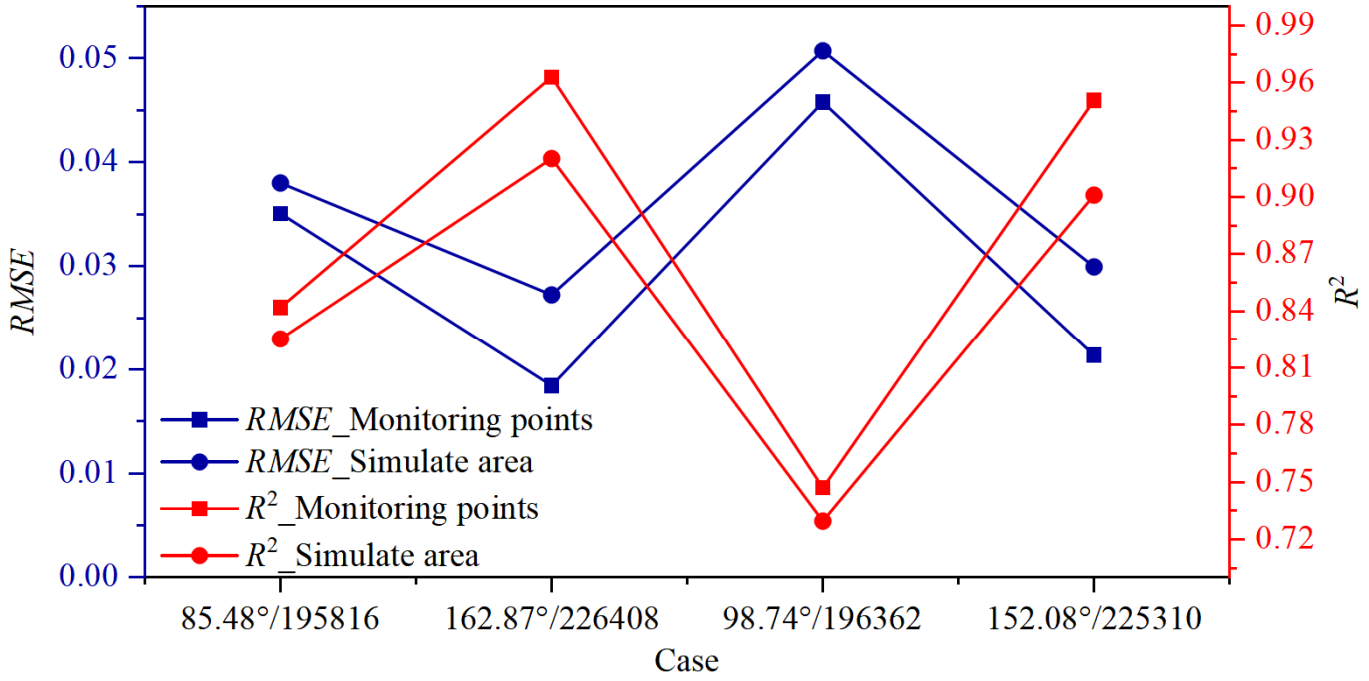



**Figure 17.** Comparison between *RMSE* and $R^2$ for monitoring points and the entire simulation area at the 500th time step (50 s).

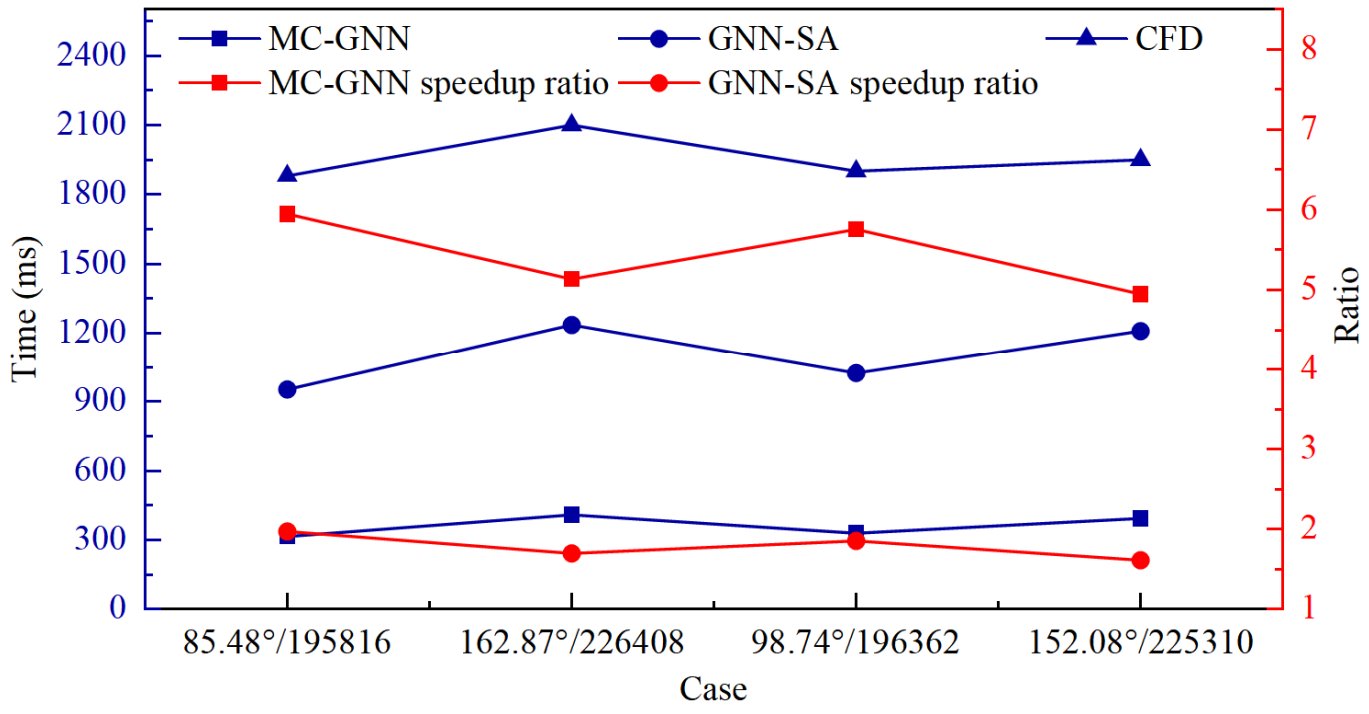



**Figure 18.** Comparison of computational efficiency between MC-GNN model, GNN-SA model, and CFD model.

### *3.3. Analysis of Factors Influencing High Exposure Risk*

This study simulated 40 different inflow conditions and conducted calculations $L_{HFR}$ based on Formula (27). The parameter selection for different simulation experiments is shown in Table 3.

**Table 3.** Selection of simulation experiment parameters.

| n | $D_\theta$ (°) | $V$ (m/s) | $T$ (K) | $Q_v$ (m/s) | $C$ | n | $D_\theta$ (°) | $V$ (m/s) | $T$ (K) | $Q_v$ (m/s) | $C$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 360(0) | 10 | 285 | 5 | 1 | 21 | 360(0) | 10 | 268 | 10 | 0.5 |
| 2 | | 5 | | | | 22 | | 5 | | | |
| 3 | | 7.5 | | | | 23 | | 7.5 | | | |
| 4 | | 12.5 | | | | 24 | | 12.5 | | | |
| 5 | 85.48 | 10 | 285 | 10 | 0.5 | 25 | 85.48 | 10 | 268 | 15 | 0.1 |
| 6 | | 5 | | | | 26 | | 5 | | | |
| 7 | | 7.5 | | | | 27 | | 7.5 | | | |
| 8 | | 12.5 | | | | 28 | | 12.5 | | | |
| 9 | 162.87 | 10 | 285 | 15 | 0.1 | 29 | 162.87 | 10 | 268 | 5 | 1 |
| 10 | | 5 | | | | 30 | | 5 | | | |
| 11 | | 7.5 | | | | 31 | | 7.5 | | | |
| 12 | | 12.5 | | | | 32 | | 12.5 | | | |
| 13 | 98.74 | 10 | 289 | 5 | 1 | 33 | 98.74 | 10 | 306 | 10 | 0.5 |
| 14 | | 5 | | | | 34 | | 5 | | | |
| 15 | | 7.5 | | | | 35 | | 7.5 | | | |
| 16 | | 12.5 | | | | 36 | | 12.5 | | | |
| 17 | 152.08 | 10 | 289 | 10 | 0.5 | 37 | 152.08 | 10 | 306 | 15 | 0.1 |
| 18 | | 5 | | | | 38 | | 5 | | | |
| 19 | | 7.5 | | | | 39 | | 7.5 | | | |
| 20 | | 12.5 | | | | 40 | | 12.5 | | | |

Six representative case studies are selected for demonstration, and the corresponding diffusion results are presented in Figure 19. The figure presents the chlorine concentration cloud maps simulated at 50 s using the diffusionFoam solver under three different wind directions and six distinct inflow boundary conditions. The diversity of inflow boundary conditions leads to varying chlorine concentration outcomes and different high-fatality areas, resulting in differing lateral distances from the high-fatality regions. Hence, we will employ Formula (27) to compute the lateral distance and compare the results with those from CFD numerical simulations to validate the effectiveness of the lateral distance formula derived from the $\pi$ theorem.

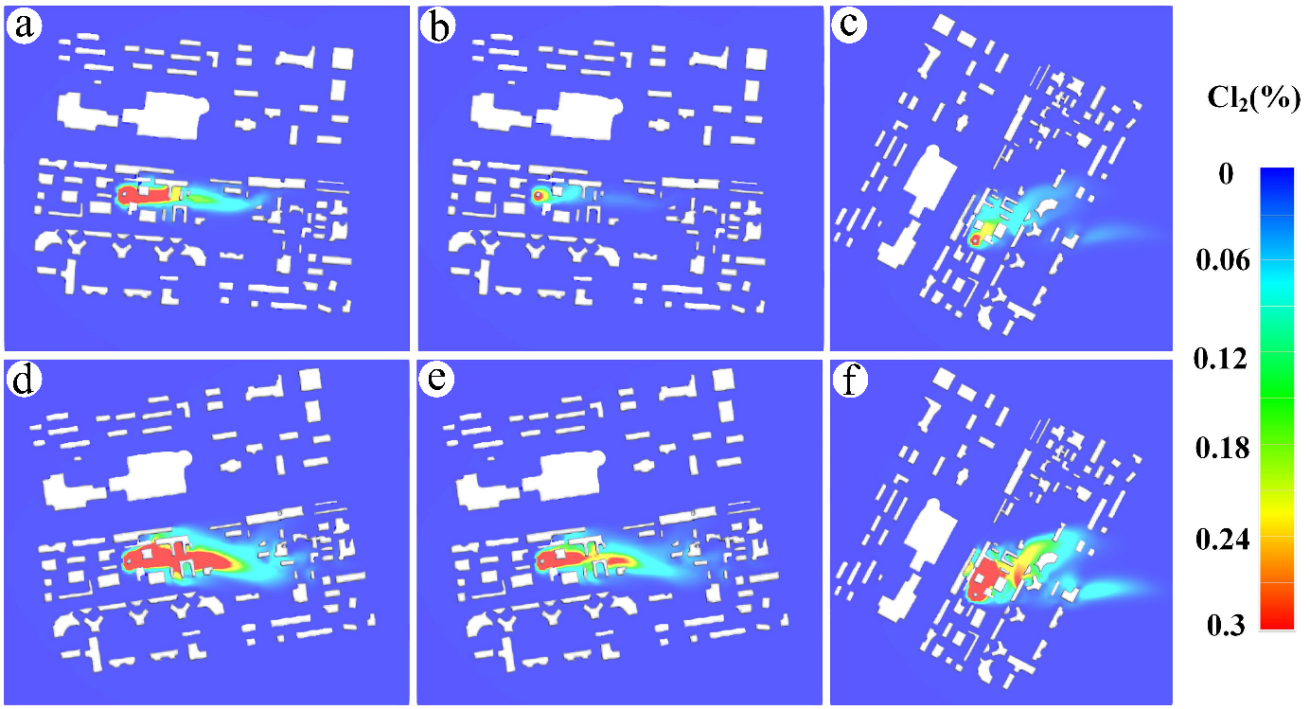



**Figure 19.** Results of chlorine diffusion experiment. (**a**–**f**) Six different combinations of wind directions and inflow boundary conditions.

 https://doi.org/10.53941/ges.2026.100034

In various experiments, Table 4 shows the comparison between the simulated values of $L_{HFR}$ and the calculated values obtained using Formula (27), as well as the absolute error between the simulated and calculated values. The results obtained from Formula (27) closely match the numerical simulation results, with an $R^2$ value of 0.96, indicating the accuracy of the formula. The maximum absolute error between the calculated values from Formula (27) and the simulated values is 5.03%, while the minimum absolute error is 0.1%. The relatively larger error observed in a few cases may be related to complex local flow structures caused by building-induced turbulence and specific combinations of meteorological conditions. However, in most cases, the predicted LHFR values are still in good overall agreement with the simulated ones. This demonstrates that Formula (27) can accurately predict $L_{HFR}$ and has practical potential.

**Table 4.** Comparison between simulated and calculated values of $L_{HFR}$.

| n | $L_{HFR}$ Calculated Values (m) | $L_{HFR}$ Simulated Values (m) | Absolute Error | n | $L_{HFR}$ Calculated Values (m) | $L_{HFR}$ Simulated Values (m) | Absolute Error |
|---|---|---|---|---|---|---|---|
| 1 | 261.759 | 260.749 | 0.39% | 21 | 167.364 | 164.7516 | 1.57% |
| 2 | 143.102 | 140.748 | 1.67% | 22 | 63.302 | 65.831 | 3.84% |
| 3 | 227.281 | 220.759 | 2.96% | 23 | 101.000 | 100.647 | 0.35% |
| 4 | 372.123 | 375.31 | 0.85% | 24 | 254.490 | 250.749 | 1.49% |
| 5 | 250.155 | 249.915 | 0.10% | 25 | 67.387 | 65.377 | 3.07% |
| 6 | 97.143 | 99.998 | 2.86% | 26 | 34.089 | 35.03 | 2.69% |
| 7 | 145.837 | 153.266 | 4.85% | 27 | 39.702 | 40.05 | 0.87% |
| 8 | 299.007 | 309.83 | 3.49% | 28 | 82.741 | 85.113 | 2.79% |
| 9 | 103.923 | 99.273 | 4.68% | 29 | 408.431 | 430.066 | 5.03% |
| 10 | 53.093 | 52.337 | 1.44% | 30 | 141.067 | 135.619 | 4.02% |
| 11 | 78.118 | 80.925 | 3.47% | 31 | 169.552 | 162.226 | 4.52% |
| 12 | 82.971 | 80.576 | 2.97% | 32 | 515.691 | 531.209 | 2.92% |
| 13 | 310.279 | 300.55 | 3.24% | 33 | 245.483 | 250.56 | 2.03% |
| 14 | 158.660 | 160.575 | 1.19% | 34 | 77.645 | 75.587 | 2.72% |
| 15 | 235.318 | 240.645 | 2.21% | 35 | 173.414 | 181.018 | 4.20% |
| 16 | 401.934 | 385.98 | 4.13% | 36 | 312.634 | 310.55 | 0.67% |
| 17 | 264.615 | 270.727 | 2.26% | 37 | 115.684 | 121.106 | 4.48% |
| 18 | 189.725 | 191.782 | 1.07% | 38 | 37.969 | 37.243 | 1.95% |
| 19 | 193.630 | 201.269 | 3.80% | 39 | 74.279 | 75.228 | 1.26% |
| 20 | 410.310 | 420.177 | 2.35% | 40 | 163.585 | 164.949 | 0.83% |

## 4. Conclusions

To address sudden toxic gas leakage incidents in urban environments, this study developed a GNN-based predictive model to forecast the diffusion patterns of toxic gases across urban areas. In addition, the exposure risk within the simulated region was assessed using a dose–response model, and a high-risk exposure analysis was conducted. The main findings are summarized as follows:

(1) To address the low accuracy and slow computational speed of traditional CFD simulations in predicting large-scale urban toxic gas dispersion, this study developed the MC-GNN model. The network employs a graph neural network architecture, which can more effectively represent the unstructured grid data generated by CFD simulations. Moreover, the model integrates meteorological factors, building information, leakage source characteristics, and other related information as inputs, enabling it to account for multiple factors influencing toxic gas diffusion in urban environments.

(2) Using the MC-GNN model to predict toxic gas concentration fields in urban environments, analysis of several typical cases demonstrated that, compared with CFD-based simulations, the MC-GNN model achieved average *RMSE* of 0.036 and $R^2$ of 0.84, indicating high prediction accuracy. In terms of computational efficiency, the prediction speed was improved by nearly one order of magnitude.

(3) Based on the toxic gas concentration fields predicted by the MC-GNN model, exposure risk assessments were performed using the dose–response model. Additionally, dimensional analysis was applied to evaluate the lateral distances of high-risk regions, where the π theorem was employed to derive the corresponding lateral distance formula. The validity of this formula was verified through 40 simulation experiments, yielding an $R^2$ of 0.96 and absolute errors ranging from 0.1% to 5.03%.

In summary, this study provides an effective framework for the rapid and accurate prediction of urban toxic gas diffusion and exposure risk, offering valuable guidance for emergency response and urban safety planning. However, this study still has certain limitations. The current dataset was generated from a representative urban area, and the transferability of the trained MC-GNN

model to cities with substantially different urban morphologies has not yet been systematically evaluated. For deployment in new urban environments, the model may require fine-tuning with a limited amount of CFD-generated data from the target area to maintain prediction accuracy. Future work will focus on constructing more diverse multi-city datasets and investigating transfer-learning strategies to improve the robustness and cross-city generalization capability of the proposed framework.

## Supplementary Materials

The additional data and information can be downloaded at: https://media.sciltp.com/articles/others/2611231451219607/GES-25120206-SI.pdf. Figure S1: WRF fitting wind profile results. Figure S2: Schematic diagram of different inlet wind angles. Figure S3: Schematic diagram of three grid resolutions. Figure S4: Partial result graph of concentration field dataset. Figure S5: Schematic diagram of monitoring points.

## Author Contributions

M.D.: Conceptualization, Project administration, Supervision, Writing—Original Draft, Writing—Review & Editing, Methodology, Software; X.S.: Conceptualization, Project administration, Funding acquisition, Supervision, Writing—Original Draft, Writing—Review & Editing, Methodology; X.J.: Writing—Original Draft, Methodology, Software, Visualization, Validation; X.L.: Writing—Original Draft, Software, Visualization, Validation; S.C.: Software, Visualization, Validation; Z.Z.: Software, Visualization, Validation; J.H.: Methodology, Validation; Z.L.: Project administration, Funding acquisition, Supervision, Writing—Original Draft, Writing—Review & Editing. All authors have read and agreed to the published version of the manuscript.

## Funding

This work was supported by the National Key R&D Program of China (No. 2023YFC2605300), the Fundamental Research Funds for the Central Universities (No. 2024JG007), the National Natural Science Foundation of China (No. 42122058), the Hebei Provincial Natural Science Foundation (No. F2020502014 and E2021502046) and the Ministry of Education of the People's Republic of China (No. 2021MS095).

## Institutional Review Board Statement

Not applicable.

## Informed Consent Statement

Not applicable.

## Data Availability Statement

The data are available from the corresponding author upon reasonable request.

## Conflicts of Interest

The authors declare no conflict of interest.

## Use of AI and AI-Assisted Technologies

No AI tools were utilized for this paper.

## References

1. Gant, S.E.; Weil, J.C.; Monache, L.D.; et al. Dense gas dispersion model development and testing for the Jack Rabbit II phase 1 chlorine release experiments. *Atmos. Environ.* **2018**, *192*, 218–240. Correction in *Atmos. Environ.* **2019**, *196*, 179.
2. Zheng, D.; Zhang, J.; Liu, Z.; et al. Effects of Pipeline Pressure on Diffusion Characteristics of Leaked Natural Gas in Tunnel Space. *ACS Omega* **2023**, *8*, 10235–10241.
3. Zhou, Z.Q.; Zhao, W.W.; Zhang, Y.L.; et al. Study on the Diffusion Law of Heavy Gas Leakage in Complex Scenarios Based on Scaled-Down Experiments. *ACS Omega* **2024**, *9*, 31533–31545.
4. Wu, B.; Zhao, R.; Meng, G.; et al. A numerical study on CO migration after blasting in high-altitude tunnel by inclined shaft. *Sci. Rep.* **2022**, *12*, 14696.
5. Feng, X.; Jiang, Z.; Zhang, G.; et al. Study on CO diffusion law and concentration distribution function under ventilation after blasting in high-altitude tunnel. *J. Wind. Eng. Ind. Aerodyn.* **2022**, *220*, 104871.
6. Hirst, W.J.S.; Eyre, J.A. Maplin Sands Experiments 1980: Combustion of Large LNG and Refrigerated Liquid Propane Spills on the Sea. In *Heavy Gas and Risk Assessment—II*; Springer: Dordrecht, The Netherlands, 1983; pp. 211–224.
7. Puttock, J.S.; Colenbrander, G.W.; Blackmore, D.R. Maplin Sands Experiments 1980: Dispersion Results from Continuous Releases of Refrigerated Liquid Propane. In *Heavy Gas and Risk Assessment—II*; Springer: Dordrecht, The Netherlands, 1983; pp. 147–161.
8. Colenbrander, G.W.; Puttock, J.S. Maplin Sands Experiments 1980: Interpretation and Modelling of Liquefied Gas Spills onto the Sea. In *Atmospheric Dispersion of Heavy Gases and Small Particles*; Springer: Berlin/Heidelberg, Germany, 1984; pp. 277–295.
9. Hanna, S.R.; Steinberg, K.W. Overview of Petroleum Environmental Research Forum (PERF) dense gas dispersion modeling project. *Atmos. Environ.* **2001**, *35*, 2223–2229.
10. Hanna, S.; Britter, R.; Argenta, E.; et al. The Jack Rabbit chlorine release experiments: Implications of dense gas removal from a depression and downwind concentrations. *J. Hazard. Mater.* **2012**, *213–214*, 406–412.
11. Thompson, R.S. Building amplification factors for sources near buildings: A wind-tunnel study. *Atmos. Environ. Part A. Gen. Top.* **1993**, *27*, 2313–2325.
12. Chen, B.; Liu, S.; Miao, Y.; et al. Construction and validation of an urban area flow and dispersion model on building scales. *Acta Meteorol. Sin.* **2013**, *27*, 923–941.

13. Janour, Z.; Jurcakova, K.; Brych, K.; et al. Potential risks at an industrial site: A wind tunnel study. *Process Saf. Environ. Prot.* **2010**, *88*, 185–190.
14. Salim, S.M.; Buccolieri, R.; Chan, A.; et al. Numerical simulation of atmospheric pollutant dispersion in an urban street canyon: Comparison between RANS and LES. *J. Wind. Eng. Ind. Aerodyn.* **2011**, *99*, 103–113.
15. Tominaga, Y.; Akabayashi, S.I.; Kitahara, T.; et al. Air flow around isolated gable-roof buildings with different roof pitches: Wind tunnel experiments and CFD simulations. *Build. Environ.* **2015**, *84*, 204–213.
16. Tan, W.; Wang, K.; Li, C.; et al. Experimental and numerical study on the dispersion of heavy gases in urban environments. *Process Saf. Environ. Prot.* **2018**, *116*, 640–653.
17. Dai, Y.; Mak, C.M.; Ai, Z.; et al. Evaluation of computational and physical parameters influencing CFD simulations of pollutant dispersion in building arrays. *Build. Environ.* **2018**, *137*, 90–107.
18. Hang, J.; Chen, L.; Lin, Y.; et al. The impact of semi-open settings on ventilation in idealized building arrays. *Urban Clim.* **2018**, *25*, 196–217.
19. Pontiggia, M.; Derudi, M.; Alba, M.; et al. Hazardous gas releases in urban areas: Assessment of consequences through CFD modelling. *J. Hazard. Mater.* **2010**, *176*, 589–596.
20. Foroutan, H.; Tang, W.; Heist, D.K.; et al. Numerical analysis of pollutant dispersion around elongated buildings: An embedded large eddy simulation approach. *Atmos. Environ.* **2018**, *187*, 117–130.
21. Tominaga, Y.; Stathopoulos, T. Steady and unsteady RANS simulations of pollutant dispersion around isolated cubical buildings: Effect of large-scale fluctuations on the concentration field. *J. Wind. Eng. Ind. Aerodyn.* **2017**, *165*, 23–33.
22. Shao, X.; Zhao, Z.; Liu, Z.; et al. Exposure risk and emergency evacuation modeling of toxic gas leakage in urban areas under the influence of multiple meteorological factors. *Environ. Pollut.* **2023**, *333*, 122044.
23. Wang, B.; Chen, B.; Zhao, J. The real-time estimation of hazardous gas dispersion by the integration of gas detectors, neural network and gas dispersion models. *J. Hazard. Mater.* **2015**, *300*, 433–442.
24. Jiao, Z.; Ji, C.; Sun, Y.; et al. Deep learning based quantitative property-consequence relationship (QPCR) models for toxic dispersion prediction. *Process Saf. Environ. Prot.* **2021**, *152*, 352–360.
25. Wang, Y.; Luo, Z.; Luo, J. Research on predicting the diffusion of toxic heavy gas sulfur dioxide by applying a hybrid deep learning model to real case data. *Sci. Total Environ.* **2023**, *901*, 166506.
26. Li, Z.; Liu, F.; Yang, W.; et al. A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**, *33*, 6999–7019.
27. Guo, X.; Li, W.; Iorio, F. Convolutional Neural Networks for Steady Flow Approximation. In Proceedings of the KDD'16: The 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 481–490.
28. Sekar, V.; Jiang, Q.; Shu, C.; et al. Fast flow field prediction over airfoils using deep learning approach. *Phys. Fluids* **2019**, *31*, 057103.
29. Kochkov, D.; Smith, J.A.; Alieva, A.; et al. Machine learning–accelerated computational fluid dynamics. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2101784118.
30. Miyanawala, T.P.; Jaiman, R.K. An Efficient Deep Learning Technique for the Navier-Stokes Equations: Application to Unsteady Wake Flow Dynamics. *arXiv* **2017**, arXiv:1710.09099.
31. Xiang, S.; Fu, X.; Zhou, J.; et al. Non-intrusive reduced order model of urban airflow with dynamic boundary conditions. *Build. Environ.* **2021**, *187*, 107397.
32. Alet, F.; Jeewajee, A.K.; Bauzá, M.; et al. Graph Element Networks: Adaptive, structured computation and memory. In Proceedings of the International Conference on Machine Learning, Long Beach, CA, USA, 10–15 June 2019; pp. 212–222.
33. Zhao, R.; Liu, S.; Liu, J.; et al. A two-stage CFD-GNN approach for efficient steady-state prediction of urban airflow and airborne contaminant dispersion. *Sustain. Cities Soc.* **2024**, *112*, 105607.
34. Shao, X.; Zhang, S.; Liu, X.; et al. Rapid prediction for the transient dispersion of leaked airborne pollutant in urban environment based on graph neural networks. *J. Hazard. Mater.* **2024**, *478*, 135517.
35. Wang, F.; Zhang, T.T.; You, R.; et al. Evaluation of infection probability of Covid-19 in different types of airliner cabins. *Build. Environ.* **2023**, *234*, 110159.
36. Zhang, X.; Wang, J. Dose-response Relation Deduced for Coronaviruses from Coronavirus Disease 2019, Severe Acute Respiratory Syndrome, and Middle East Respiratory Syndrome: Meta-analysis Results and its Application for Infection Risk Assessment of Aerosol Transmission. *Clin. Infect. Dis.* **2021**, *73*, e241–e245.
37. Liu, M.; Liu, J.; Cao, Q.; et al. Evaluation of different air distribution systems in a commercial airliner cabin in terms of comfort and COVID-19 infection risk. *Build. Environ.* **2022**, *208*, 108590. Correction in *Build. Environ.* **2022**, *214*, 108939.
38. Liu, Z.; Cao, H.; Hu, C.; et al. Modeling the infection risk and emergency evacuation from bioaerosol leakage around an urban vaccine factory. *npj Clim. Atmos. Sci.* **2023**, *6*, 6.
39. Schubach, S. A measure of human sensitivity in acute inhalation toxicity. *J. Loss Prev. Process Ind.* **1997**, *10*, 309–315.
40. Yang, D.; Chen, G.; Shi, J.; et al. A novel approach for hazardous area identification of toxic gas leakage accidents on offshore facilities. *Ocean. Eng.* **2020**, *217*, 107926.
41. Leitl, B.; Schatzmann, M. *Compilation of Experimental Data for Validation Purposes*; Meteorology Institute, Hamburg University: Hamburg, Germany, 1998.
42. Franke, J.; Hellsten, A.; Schlunzen, K.H.; et al. The COST 732 Best Practice Guideline for CFD simulation of flows in the

urban environment: A summary. *Int. J. Environ. Pollut.* **2011**, *44*, 419–427.

43. Ai, Z.T.; Mak, C.M. CFD simulation of flow in a long street canyon under a perpendicular wind direction: Evaluation of three computational settings. *Build. Environ.* **2017**, *114*, 293–306.
44. Pfaff, T.; Fortunato, M.; Sanchez Gonzalez, A.; et al. Learning Mesh-Based Simulation with Graph Networks. In Proceedings of the International Conference on Learning Representations, Virtual, 3–7 May 2021.
45. Sánchez-González, Á.; Heess, N.; Springenberg, J.T.; et al. Graph networks as learnable physics engines for inference and control. In Proceedings of the International Conference on Machine Learning, Stockholm, Sweden, 10–15 July 2018; Volume 80, pp. 4470–4479.