Transactions on Environment, Energy and Earth Sciences
ISSN: 2960-1924, eISSN: 2960-2300 | Volume 7
GTREE 2026

Language Model Empowered Cross-Scale Framework for Urban Air Pollution Monitoring, Diffusion Analysis and Early Warning

Jingyi Qian *

Shanghai Star River Bilingual School, Shanghai, China

* Corresponding Author Email: jeanqian0702@163.com

Abstract. Modern urban morphology and local meteorological conditions create significant variations in pollutant dispersion across scales. Under global climate change and rapid urbanization, air pollution threatens public health and sustainable development, yet traditional monitoring fails to address both micro-level protection and macro-level governance—especially within the "street canyon–residential area" microenvironment where daily human activity concentrates. To bridge this gap, this study first develops an LLM-enhanced CFD system with atmospheric stability metrics to model pollutant dispersion from street canyons to residential areas. Since micro-level research cannot cover large regions, we further propose a context-aware human-machine collaborative drone framework integrating multimodal sensing, XGBoost prediction, and natural language interaction for autonomous surveillance, source localization, and short-term alerts. Results show that under stable conditions, peak CO in street canyons exceeds 80 μg/m³—over three times higher than under unstable conditions (5–35μg/m³)—with high-concentration plumes traveling long distances to downstream residential areas. The drone system achieves 0.34 m localization accuracy, a 1-hour CO prediction RMSE of 2.3 μg/m³, 98% natural language interpretation accuracy, and sub-3-minute response to sudden pollution events. This cross-scale framework unifies micro-mechanisms and macro-monitoring, offering a practical solution for urban air quality management and healthy city construction.

Keywords: Large Language Model, Agentic Drone, Atmospheric Stability, CFD.

1. Introduction

1.1. New Challenges in Urban Multi-Scale Environmental Monitoring

Air pollution poses a major public health threat, yet a critical mismatch persists monitoring occur at the city level, while exposure happens at the human level. This creates two challenges. At the macro scale, fixed stations are too sparse to detect pollution in industrial zones, transport hubs, or sudden hotspots. At the micro scale, urban street canyons—narrow corridors between high-rises—trap vehicle exhaust near the ground, endangering nearby residents. As shown in Shanghai (Figure 1), both Puxi and Pudong have many high-rise structures forming such canyons, with large residential communities often located next to them.

Motor vehicles frequently traverse street canyons, making exhaust emissions a primary pollution source in these microenvironments. The shielding effect of canyon-shaped buildings impedes airflow, causing pollutants to accumulate within canyons and downstream residential areas, posing public health risks. At the macro level, traditional fixed monitoring stations—spaced kilometers apart—cannot cover key sources like industrial parks or transport hubs, nor respond swiftly to sudden incidents, and urban complexity worsens these blind spots. Significant spatiotemporal discrepancies often arise between local air quality measurements and city-wide reports, especially where high-density buildings and intricate transport networks coexist. Pollutants such as CO, PM2.5, and NOx show strong variability tied to traffic density, street canyon effects, ventilation corridors, and micro-emission sources. These gaps highlight a critical need for adaptive, high-resolution monitoring systems that connect macro-level zero-carbon goals with the dynamic realities of complex urban micro-environments.

Content from this work may be used under the terms of CC BY-NC 4.0 licence (https://creativecommons.org/licenses/by-nc/4.0/).
Published by Warwick Evans Publishing.



Figure 1. The street canyons and residential areas in the Shanghai urban area. a) The Perspective of Puxi, Shanghai； b) A Perspective from Shanghai Pudong.

1.2. Literature Review: From Microscopic Pollution Spread Studies to Macroscopic UAV Monitoring

The multi-scale nature of urban pollution requires both broad regional perception and detailed mechanistic analysis. However, existing research on macro-scale drone monitoring and micro-scale fluid dynamics has long evolved independently, lacking an integrated cross-scale framework.

1.2.1. Micro-pollution dispersion: Integrated application of CFD simulation and LLM

Research on urban pollutant dispersion mainly employs field measurements, wind tunnel experiments, and CFD simulations [1–3], with key influencing factors including wind speed [4], direction [5], turbulence intensity [6], and thermal effects [7]. Thermal stratification—unstable, neutral, or stable—directly regulates turbulent transport: unstable conditions enhance vertical mixing [8], while stable stratification suppresses turbulence and traps pollutants near the surface [9]. Studies on isolated buildings and idealized street canyons have shown that thermal effects alter roof-level diffusion [10], wind speed and thermal forcing interact to determine lateral diffusion [11], and thermal stratification can modify vortex patterns and ventilation rates, potentially transforming a single vortex into a dual-vortex configuration [12,13]. However, most studies have focused on isolated buildings or 2D street canyons, overlooking the critical pathway from street canyons to adjacent residential areas, and have failed to elucidate how different stability states specifically influence this transport process and resulting exposure risks. Comprehensive comparative analyses across all stability scenarios remain scarce. Moreover, CFD simulations are computationally intensive, typically requiring hours or days per run, making them inadequate for real-time monitoring and early warning. Therefore, integrating prior physical knowledge from CFD with the semantic reasoning capabilities of large language models to develop lightweight, interpretable tools for microenvironmental pollution analysis represents a critical approach to overcoming current technological limitations.

1.2.2. Macro-scale Environmental Monitoring: Current Status of UAV Applications

The traditional urban environmental monitoring system, centered on fixed ground-based stations, provides city-scale air quality trends but suffers from low spatial resolution, limited coverage, and inability to respond to sudden incidents [14]; the average spacing between stations worldwide exceeds 5 km with coverage below 20% in key source areas—for example, only one station exists within Shanghai's 3.5 km^2 Minhang Economic and Technological Development Zone. Unmanned aerial vehicles (UAVs) have emerged as a promising solution due to their flexibility, rapid deployment, high spatial resolution, and low cost, with quadcopter UAVs enabling meter-level in-situ monitoring [15] and over 300 research projects now using UAVs to detect six or more key pollutants (PM2.5, PM10, CO, NO_2, SO_2, O_3), some integrating infrared thermal imaging for diffuse source identification [16]. In application, drones have been widely used for industrial park inspections, traffic monitoring, emergency response, and urban background surveys—e.g., Zhang et al. (2024) identified three

unregistered fugitive VOC sources with sub-meter accuracy [17], and Li et al. (2017) found near-ground CO within 10 m was over 40% higher than at 30 m altitude [18]. Despite these advances, existing drone systems face three major bottlenecks: rigid path planning that cannot adapt to real-time data; delayed data processing due to "airborne collection–ground offline analysis," preventing real-time alerts; and a high operational threshold requiring expert skills, limiting routine use.

Existing research has three major gaps: (1) no systematic study of the "street canyon–residential area" layout and atmospheric stability effects on cross-zone transport; (2) no intelligent drone system capable of real-time, autonomous decisions based on sensor data and natural language commands; and (3) no integrated framework bridging micro-scale physics and macro-scale sensing. This study addresses these gaps by establishing a bidirectional macro-micro analytical framework and, for the first time, using large language models to both interpret CFD physical insights and control autonomous drone missions.

2. Research Methods for Multiscale Monitoring Systems

2.1. Microscale: CFD Simulation Under the Influence of Atmospheric Stability

2.1.1. Atmospheric stability

Atmospheric stability describes the atmosphere's tendency to suppress or enhance vertical motion, influencing weather, cloud formation, and pollutant dispersion. Stability is determined by buoyancy: warmer air rises (unstable), cooler air sinks (stable), while neutral conditions represent equilibrium between solar heating and mechanical turbulence. This study simulates three conditions: unstable (summer, 10:00–16:00)—strong solar radiation enhances vertical mixing; neutral (spring morning, 7:00–10:00)—balanced heating and mechanical turbulence; and stable (winter night, 22:00–05:00)—suppressed turbulence, trapping pollutants near the ground.

2.1.2. Reynolds' average Navier-Stokes (RANS) equation

For urban dispersion, LES is too costly for large-scale studies, so we use RANS with the standard *k-ε* model to balance efficiency and accuracy.

The continuity equation: For low Mach number flows, the fluid is considered incompressible, and its density variation is solely caused by temperature changes, where density is linearly related to temperature via the thermal expansion coefficient.

$$\frac{\partial \rho}{\partial t}+\frac{\partial}{\partial x_i}(\rho \overline{u}_i)=0, \tag{1}$$

Here $\overline{u}_i$ is average velocity in the direction x_i.

The momentum equation, which accounts for the buoyancy effect caused by density changes induced by temperature, is given by the following formula:

$$\frac{\partial}{\partial t}(\rho \overline{u}_i)+\frac{\partial}{\partial x_j}(\rho \overline{u}_i \overline{u}_j)=-\frac{\partial \overline{p}}{\partial x_i}+\frac{\partial}{\partial x_j}\left(\mu \frac{\partial \overline{u}_i}{\partial x_j}-\rho \overline{u'_i u'_j}\right)+(\rho-\rho_{ref})g_i, \tag{2}$$

Here g_i is the gravitational acceleration, ρ_{ref} is the reference density, and the Reynolds stress term is closed using the standard *k-ε* turbulence model.

Energy Equation: Energy transport is described by the following equation:

$$\rho c_p\left(\frac{\partial T}{\partial t}+\overline{u}_i \frac{\partial T}{\partial x_i}\right)=\frac{\partial}{\partial x_j}\left(\lambda \frac{\partial T}{\partial x_j}\right)+S_T, \tag{3}$$

Here, T denotes the average temperature, cp represents the specific heat capacity, λ stands for the thermal conductivity, and ST refers to the volumetric heat source.

Transport of Pollutants: The diffusion of pollutants is simulated using a scalar transport equation:

$$\frac{\partial}{\partial t}(\rho C)+\nabla\cdot(\rho\bar{u}C)=\nabla\cdot\left[\left(\rho D+\frac{\mu_t}{Sc_t}\right)\nabla C\right]+S_C, \quad (4)$$

Here, C represents the pollutant concentration, SC denotes the source term, D is the molecular diffusion coefficient, and Sct is the turbulent Schmidt number, with a default value of 0.7.

2.1.3. Flow Configuration, Calculation Domain, and Boundary Conditions

Field investigations of Shanghai's urban architecture (Puxi and Pudong) reveal that multiple rows of high-rise buildings often form street canyons, with large residential areas typically located nearby (Figure 1). Based on these observations, we developed simplified representative architectural models (Figure 2). The grid configuration is shown in Figure 3a (vertical *x-z* plane) and 3b (horizontal *x-y* plane). The simulated urban microenvironment comprises a street canyon and an adjacent medium-density residential area, with geometric parameters (summarized in Tables 1 and 2) selected based on standard urban aerodynamic considerations.

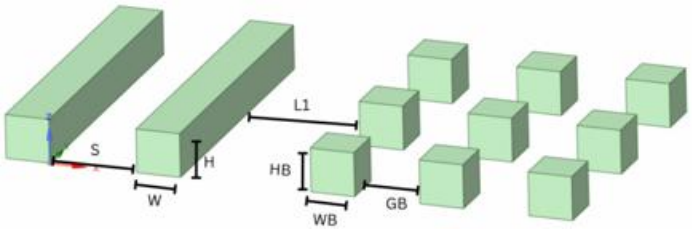


Figure 2. Flow Configuration Diagram.

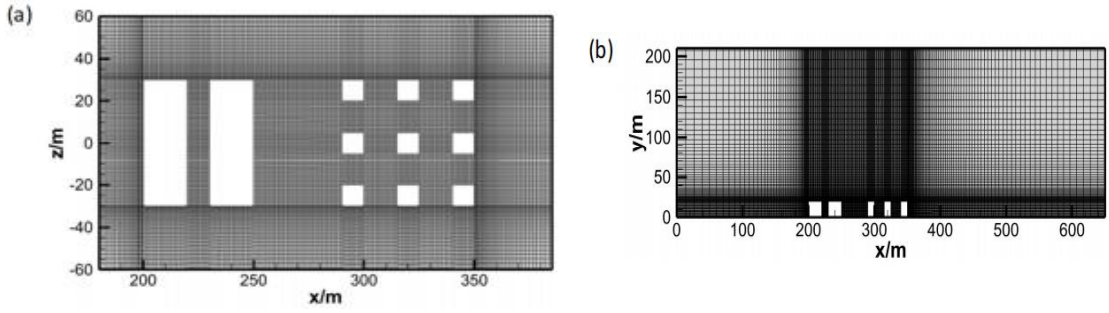


Figure 3. Computational grid: (a) vertical xoz plane, (b) horizontal xoy plane.

The street canyon has a height-to-width ratio H/S=2 and building width ratio H/W=1, with building height H=20 m (Table 1). The adjacent residential area (Table 2) features building height H_B=18 m (typical six-story), setback distance L_1=40 m (L_1/H=2), and building coverage λ=0.3–0.4. Boundary conditions simulate real urban atmospheric flow: inlet with a vertical wind profile following the urban logarithmic law; outlet as a pressure outlet with zero static pressure gradient; ground and building surfaces using standard wall functions; top boundary as a symmetric (sliding) wall placed sufficiently high; lateral boundaries with periodic conditions to represent an infinite urban surface and fully developed flow. The domain length ensures adequate inflow development and wake dissipation without reflux effects.

Table 1. Street canyon geometric parameters.

Parameter	Description	Value	Dimensionless ratio
H	Building height	20 m	-
S	Street width	10 m	H/S=2
W	Building width	20 m	H/W=1

Table 2. Residential area parameters

Parameter	Description	Value	Remarks
H_B	Building height	18 m	Typical 6-story residence
W_B	Building width	10 m	Typical building footprint
G_B	Building spacing	15 m	Ensure sufficient ventilation
L_1	Setback distance from street canyon	40 m	Derived from $L_1/H=2$
λ	Building coverage ratio	0.3–0.4	Medium-density urban development

2.1.4. Modeling of Vehicle Exhaust Emissions

Vehicle exhaust emissions are estimated using the unit road length method, with the emission rate calculated as the product of traffic volume and the emission factor per vehicle:

$$\text{Q (g/s)=Traffic Flow×Emission Factor per Vehicle.} \tag{5}$$

Traffic flow N (veh/s), obtained from monitoring data or reasonable assumptions, typically ranges from 1000–2000 veh/h/lane during peak hours and 200–800 veh/h/lane during off-peak hours. The emission factor E (g/km/veh) depends on vehicle type, fuel, and pollutant species; for a National V standard gasoline vehicle, NO_x emissions are 0.3–0.6 g/km, CO emissions are 1–2 g/km, $PM_{2.5}$ emissions are 0.005–0.01 g/km, and hydrocarbons (HC) emissions are 0.3–0.5 g/km. The average vehicle speed V (km/h) is used to convert distance-based emissions into time-based rates. The emission intensity q (g/s/m) per unit road length is then calculated accordingly. For example, when N=1000veh/h/lane (≈0.278 veh/s/lane), E=0.5g/km (NO_x), and V=40km/h, the resulting emission intensity q is obtained by:

$$q=\frac{NE}{V/3.6}=\frac{0.278\times0.5}{40/3.6}\approx0.0125 \text{ g/s/m}. \tag{6}$$

For a two-lane road with a total width of 7 meters, the total emission rate is:

$$Q_{total}=q\times\text{lane count}\approx0.0125\times2=0.025 \text{ g/s}. \tag{7}$$

2.2. Macro scale: Context-aware human-machine collaborative drone system

2.2.1. Overall System Architecture

The proposed context-aware human-machine collaborative drone system employs an edge computing architecture to overcome the response latency bottleneck of traditional "centralized ground station processing." The system is structured into three layers (Figure 4): (1) Perimeter Awareness Layer (onboard the drone)—responsible for real-time acquisition, preprocessing, and low-latency inference of multimodal environmental data; using sensor fusion algorithms, it synchronizes GPS-RTK positioning, IMU attitude, and gas concentration data to ensure each measurement has precise spatial coordinates. (2) Ground Inference Layer (local machine)—runs large language models, XGBoost prediction models, and an environmental database; it receives raw data from the edge device, generates pollution heatmaps, predicts concentration trends, and performs complex task inference. (3) Human-Computer Interaction Layer—supports users in issuing task commands via natural language or keyboard input, significantly lowering the operational barrier.

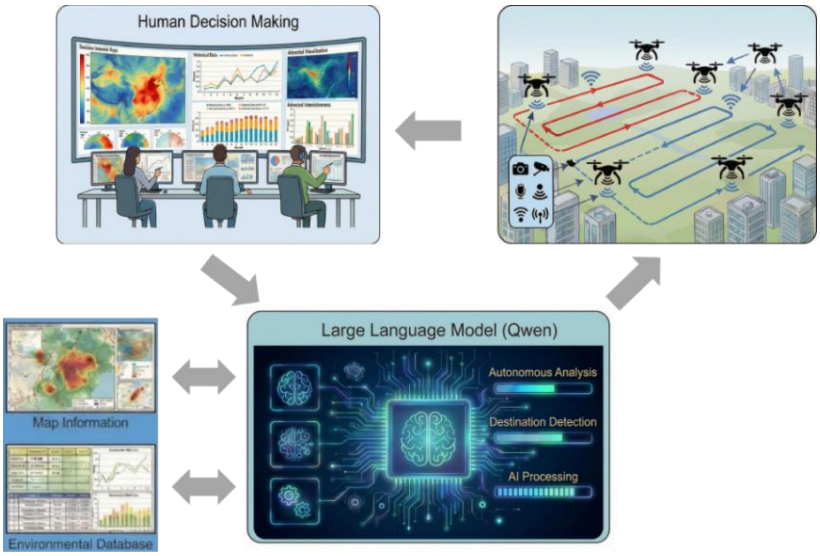


Figure 4. Human-Machine Collaborative Drone Framework with Situational Awareness.

2.2.2. Design of the drone hardware platform

Unmanned aerial vehicles (UAVs) serve as both the environmental perception platform and the intelligent execution agent within the situational awareness-based human-machine collaborative framework, following LLM commands. The hardware includes a flight controller for attitude control, a GPS-RTK module for precise positioning, and an ROS-based onboard computer for executing flight programs and MAVLink communication (Figure 5). Both control modes support LLM-guided pollution thermogram S-path scanning and fixed-point reconnaissance. Integrated gas sensors continuously monitor CO, NO_2, SO_2, $PM_{2.5}$, and PM_{10}, enabling real-time autonomous environmental monitoring and adaptive mission execution.

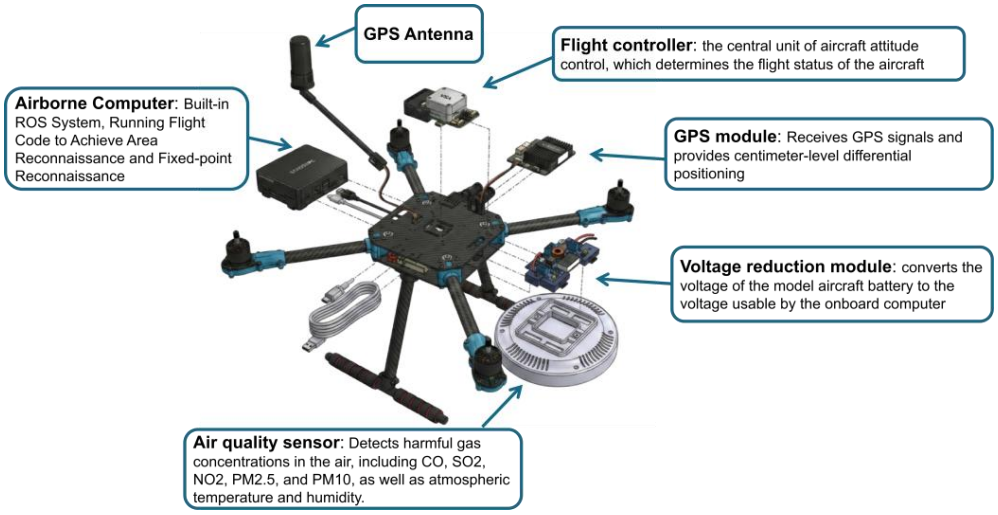


Figure 5. Schematic diagram of the drone

2.2.3. A pollutant concentration prediction model based on XGBoost

Pollutant concentrations are nonlinear and non-stationary, making historical data alone insufficient for high-precision prediction. This study constructs an input feature vector combining temporal features (hour, week, month, year) and historical hourly $PM_{2.5}$, PM_{10}, CO, SO_2, NO_2 records from Shanghai over the past decade. To validate XGBoost, three benchmark models—RF, LSTM, and SVR—were trained on the same dataset, with performance evaluated using RMSE, MAE, and R^2.

In Table 3, XGBoost outperformed RF, LSTM, and SVR across all metrics. For 1-hour CO concentration predictions, XGBoost achieved an RMSE of 2.3 μg/m³ and an R^2 of 0.92—an 18% reduction in RMSE and a 5% improvement in R^2 over LSTM. This is attributed to XGBoost's ability

to capture nonlinear feature interactions and its robustness to missing data and outliers. To further optimize performance, hyperparameter tuning was conducted over the following ranges: n_estimators (50–300), max_depth (3–10), learning_rate (0.01–0.3), γ (0–1), and λ (0–10). The final optimal combination was n_estimators=200, max_depth=6, learning_rate=0.1, γ=0.1, and λ=1.

Table 3. Performance comparison of different prediction models

Model	RMSE	MAE	R^2	Training time(s)
XGBoost	2.3	1.6	0.92	12.5
LSTM	2.8	2.0	0.87	45.2
Random Forest	3.1	2.2	0.85	9.7
SVR	3.7	2.6	0.79	25.3

2.2.4. A Task Control System Based on Natural Language

To enable intuitive human-machine collaboration, the framework uses a natural language interface allowing users to issue voice commands. A microphone array captures speech, an ASR model converts it to text, and the transcribed text is fed into GPT-4o integrated with a RAG module that combines map data, real-time sensor streams, and pre-calculated pollution heatmaps. When a user issues a command such as "Go to the location with the highest pollutant concentration," the LLM performs contextual reasoning: it queries the RAG database for the coordinates of the peak concentration in the latest heatmap, converts the user's intent into a structured flight plan, and transmits it to the flight controller via MAVLink for autonomous navigation and continuous multimodal data collection. A prompt template for this process is shown in Figure 6.

```
"""
You are a professional expert in analyzing unmanned aerial vehicle environmental
monitoring tasks. Please analyze user instructions based on the following
information and generate structured flight tasks.

[Environmental Context]
{rag_context}
Real time data
Current drone location: ({x}, {y}, {z})
Latest Pollution Heat Map: {heatmap_mase64}

[Drone Capability]
Supported flight modes: S-type cruise, fixed-point monitoring, return flight
Monitored pollutants: PM2.5, PM10, CO, NO2, SO2

[User Instructions]
{user_query}

[Output Requirements]
Please output strictly in JSON format, including the following fields:
-Task_type: Task type (cruise/monitor/return)
-Target: Target coordinates or area
-Duration: Monitoring duration (minutes)
-Pollutant: Types of pollutants monitored
-Explanation: Task Description
"""
```

Figure 6. Prompt Template.

Large language models achieve over 98% instruction parsing accuracy. For vague commands (e.g., "Go check that place"), the model proactively requests clarification; for unexecutable instructions (e.g., "Fly 100 km away"), it explains the limitations and offers alternatives. The integration of voice control, LLM inference, and automated heatmaps creates a closed-loop mechanism linking human intent to autonomous drone operations, enabling rapid, adaptive responses to dynamic pollution events and more efficient large-scale environmental monitoring.

2.3. The Influence of Atmospheric Stability on Temperature Distribution

Atmospheric stability controls pollutant dispersion: unstable conditions (daytime, summer) enhance vertical mixing, while stable conditions (nighttime, winter) suppress turbulence and trap pollutants near the ground. Figure 7 shows horizontal temperature fields. Under unstable conditions (Figure 7a), non-uniform isotherms reflect intense mixing, driving lateral exchange that transports pollutants from street canyons to residential areas with dilution. Under neutral conditions (Figure 7b), the uniform temperature field indicates mechanically driven dispersion following mean wind patterns. Under stable conditions (Figure 7c), a stratified thermal structure with cold air near the surface suppresses turbulence, allowing high-concentration plumes to reach residential areas with minimal dilution.

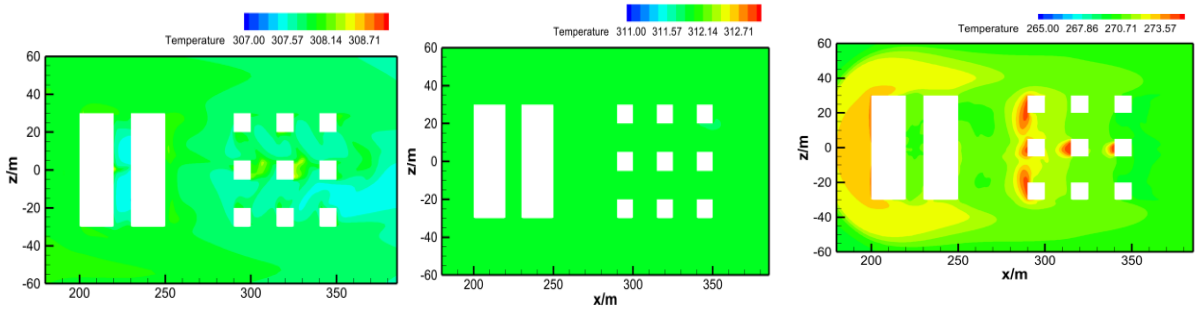


Figure 7. Horizontal (*x*–*z*) temperature contours: (a) Unstable, (b) Neutral, (c) Stable.

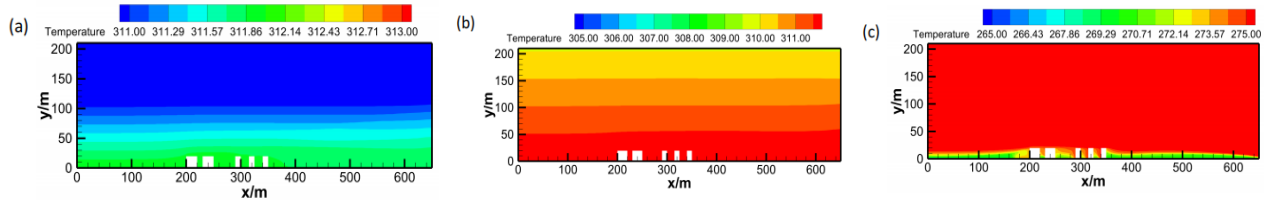


Figure 8. Vertical (*x*–*y*) temperature contours: (a) Unstable, (b) Neutral, (c) Stable.

The vertical temperature profile in Figure 8 is particularly critical. Under unstable conditions (Figure 8a), the near-neutral lapse rate supports intense vertical mixing and upward air currents, providing an effective emission mechanism for street canyons. A neutral atmosphere (Figure 8b) allows shear-driven turbulence to drive moderate vertical mixing. In stark contrast, under stable conditions (Figure 8c), the clearly defined inversion layer acts like a lid, severely suppressing vertical transport and causing pollutant accumulation within street canyons and the urban canopy (directly upwind of residential areas). Thus, the thermal structure reinforces the flow and turbulence patterns described earlier, collectively defining the diffusion environment.

2.4. The Impact of Atmospheric Stability on Pollution Concentration

The passive scalar concentration fields (Figure 9: horizontal; Figure 10: vertical) integrate flow, turbulence, and temperature, revealing exposure risks for downstream residential areas. Under unstable conditions (Figure 9a, 10a), intense turbulence and effective vertical mixing dilute the plume significantly by the time it reaches residential areas, resulting in broader dispersion but lower peak concentrations and good pedestrian-level air quality. Under neutral conditions (Figure 9b, 10b), a more concentrated plume is transported by coherent mean flow with moderate turbulence, forming a predictable transport pathway with intermediate concentrations and a well-mixed vertical distribution decreasing gradually with height. Under stable conditions (Figure 9c, 10c), suppressed turbulence and limited mixing keep the plume highly concentrated as it is advected toward residential areas, while an inversion layer sharply reduces concentrations above roof height, trapping pollutants within the urban canopy. Persistent recirculation zones between buildings further trap pollutants, exposing residents to prolonged high concentrations and posing the greatest adverse health risk.

In summary, atmospheric stability dominates pollutant transport from street canyons to residential areas. Unstable conditions enhance mixing and dilution, reducing peak concentrations but expanding

the spatial footprint. Neutral conditions produce predictable, mechanically driven dispersion. Stable conditions trap high-concentration plumes near the ground, creating prolonged, high-intensity exposure—the greatest health risk for local residents.

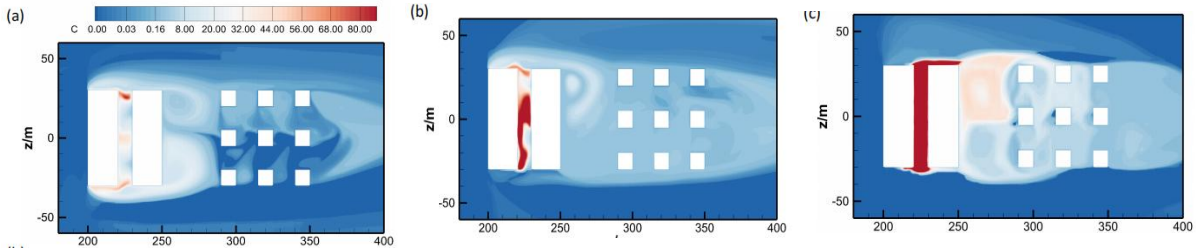


Figure 9. Horizontal pollutant contours: (a) Unstable, (b) Neutral, (c) Stable.

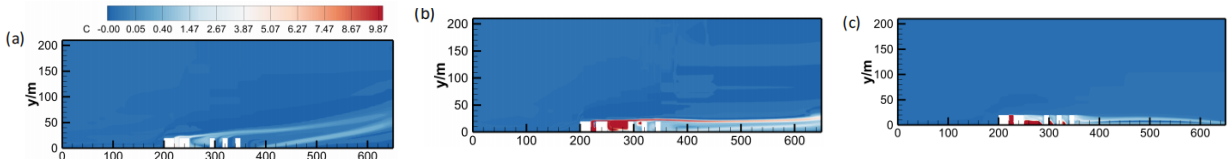


Figure 10. Vertical pollutant contours: (a) Unstable, (b) Neutral, (c) Stable.

2.5. Experimental Results of the Macro UAV Monitoring System

To evaluate pollution source localization, four smoke generators were randomly placed as simulated sources within a 10 m × 10 m indoor area. The drone followed an S-shaped path at 1 m/s, sampling at 1-m intervals (100 data points). Using cubic spline interpolation of CO data, the system generated a heatmap where the darkest regions matched the actual source locations (Figure 11). Positioning accuracy reached 0.34 m over three trials. The drone scanned the entire area within 10 minutes, producing a heatmap that clearly reveals pollutant distribution and enables rapid identification of key pollution zones.

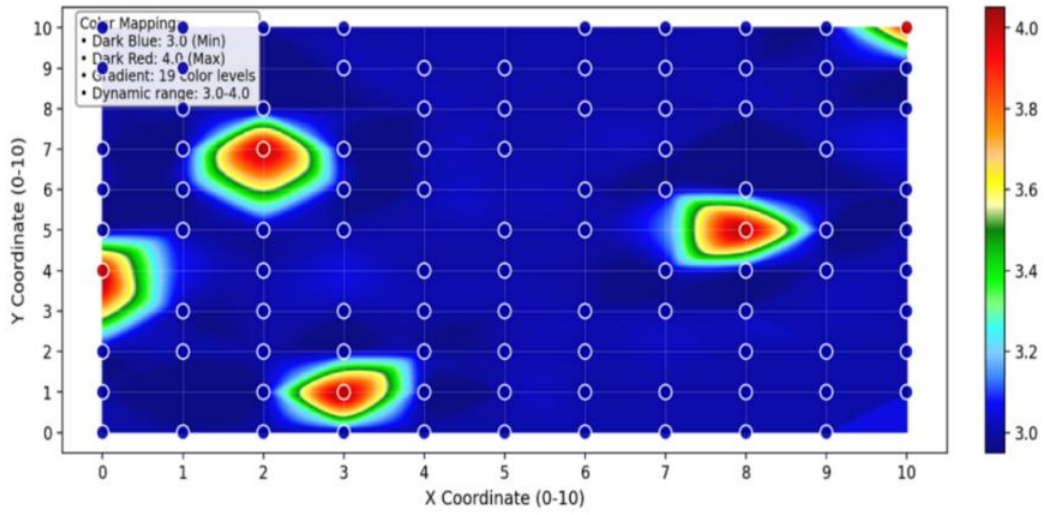


Figure 11. The heatmap of pollution in the offline experiment.

2.6. Performance Evaluation of the XGBoost Model

After data collection, the AI model analyzes and predicts pollutant trends. Figure 12 shows XGBoost performance for CO concentration prediction, incorporating multidimensional inputs (CO, NO_2, SO_2, PM2.5, PM10, temperature, humidity, time) with an 80:20 train–test split. In Figure 12, the learning curve (left) shows rapid decline and convergence of training and test RMSE. The upper-right comparison of actual vs. predicted CO distributions indicates strong agreement. The lower panel presents short-term CO predictions, closely matching recent historical observations.

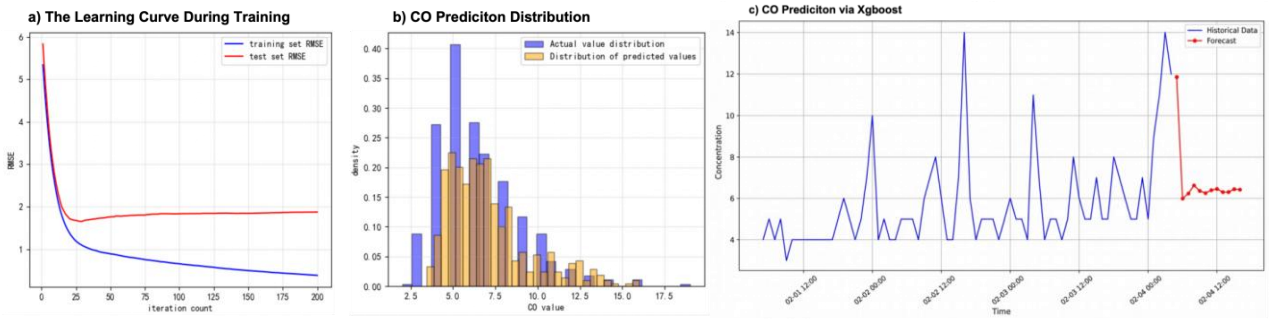


Figure 12. The learning curve: b) CO prediction distribution; c) CO prediction via Xgboost.

2.7. Natural Language Interaction Evaluation

After generating pollutant heatmaps and AI predictions, the RAG module integrates this information with geographic map data and feeds it into an LLM. Users then issue natural language instructions to specify new operational objectives. The LLM interprets user intent by reasoning through the integrated environmental and spatial context. In Figure 13, when a user commands "Go to the factory to detect pollutants," the LLM identifies the factory's precise coordinates from map data and converts them into executable flight commands for autonomous drone navigation. These results support the hypothesis that integrating LLM reasoning with physics-based modeling improves both macro-scale monitoring responsiveness and micro-scale prediction accuracy.



Figure 13. a) User voice input; b) Agentic drone flied to the location.

3. Experimental Verification and Application

3.1. Experimental Verification of Atmospheric Stability on Pollutant Diffusion

While Section 3 used CFD simulations to theoretically investigate how atmospheric stability affects pollutant dispersion, this section presents a physical experiment to validate those findings. A miniature experimental platform was built, comprising a compact chamber, 3D-printed building models, an ice-based cooling source, smoke tracers, and gas sensors (Figure 14). The platform aims to observe pollutant transport differences under stable and unstable conditions for qualitative comparison with the CFD simulations. An ESP32-S3 serves as the core data acquisition and transmission unit, with PM2.5, CO, and CO_2 sensors connected via I2C/UART for real-time sampling in the street canyon and downstream areas. The ESP32-S3 periodically reads concentration data, records timestamps, and transmits results to a host analysis program via serial or Wi-Fi, forming a structured sensor data stream for LLM input. The experimental setup consists of six semi-enclosed boxes (15×15×10 cm) with two hollow 50 cm structures forming a tall building model, creating a street canyon with distinct flow and retention characteristics. The hollow design reduces thermal conductivity, allowing sensors to detect air conditions rather than material temperatures. The downstream canyon section is designated as the residential area, aligning the experimental layout with the numerical simulation before.

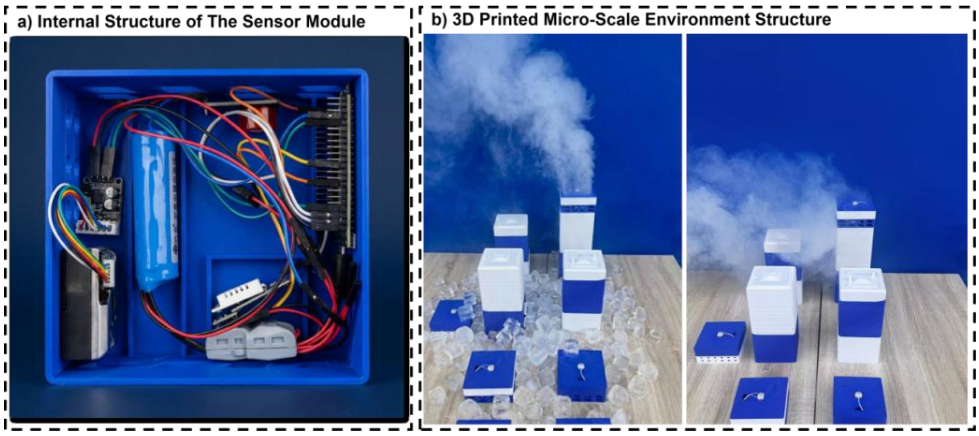


Figure 14. Gas detection module: (a) sensor internal structure, (b) 3D printed micro-scale housing.

Two atmospheric stability scenarios were created by adjusting ice block positions. With ice at the chamber bottom, a "cold bottom, warm top" inversion formed, weakening vertical mixing. Smoke diffused near the ground, accumulated along the canyon floor, and moved downstream with significant concentrations in residential areas—aligning with stable stratification simulations. Temperature sensors confirmed the inversion (coldest air at bottom, warmer above). With ice at the top (near building summit), a "cold top, warm bottom" distribution created weak thermal disturbances. Bottom warmer air rose slightly, generating weak but stable thermal updrafts. Under these conditions, smoke no longer accumulated along the ground but rose from the canyon floor and became significantly diluted near the building summit—closely matching simulation findings that unstable conditions promote vertical mixing and facilitate pollutant removal from the surface layer.

To compare horizontal transport under stable and unstable conditions, CO, CO_2, and $PM_{2.5}$ sensors were placed along the path between tall and short buildings. Under stable conditions (ice at bottom), smoke moved horizontally along the canyon due to ground-contact movement, reaching the downstream low-rise location earlier — aligning with simulations showing stable stratification inhibits vertical mixing. Under unstable conditions (ice at top), weak thermal disturbances caused smoke to rise and dilute in upper regions, weakening horizontal transport; CO sensors near low-rise buildings showed later, lower peak readings (Figure 15). Thus, unstable conditions reduce near-ground transport, causing smoke to "disperse upward first, then spread outward," resulting in slower horizontal progression.

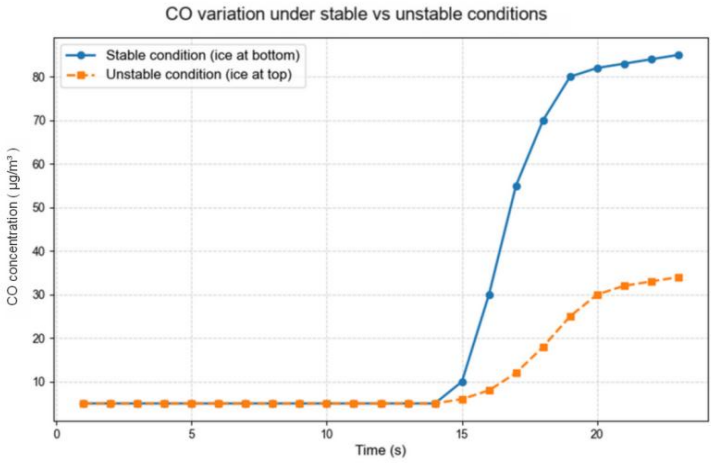


Figure 15. Comparison of ground-level CO diffusion patterns under stable vs. unstable conditions

The results show that under stable conditions, near-ground pollutants readily reach downstream areas, while unstable conditions induce vertical uplift, reducing horizontal transport and delaying downstream sensor responses. This aligns with observed smoke patterns and validates the simulation mechanism that stability controls horizontal transport efficiency. Despite measurement errors, the

experiment clearly demonstrates that stable conditions cause ground-level pollutant accumulation. These observations are highly consistent with numerical simulations in terms of physical mechanisms, providing a reliable experimental basis for the subsequent LLM-based reasoning system and confirming that even simplified setups can reconstruct the key stability effects in urban micro-scale aerodynamics.

3.2. A mobile application based on large language models and fluid dynamics simulations

Theoretical simulations and experiments have revealed that atmospheric stability is the key factor determining urban pollutant transport and exposure levels. In street canyon–residential area microenvironments, stability influences dispersion pathways, vertical mixing, and retention times. Stable conditions suppress turbulence, trapping pollutants near surfaces and creating localized pollution "traps." Unstable conditions promote thermal convection, enhancing mixing and rapid dilution. As such microenvironments proliferate, conducting case-by-case data collection, AI training, and deployment is unsustainable. Moreover, during atmospheric transitions (e.g., nighttime cooling or morning heating), AI models fail to capture abrupt changes in turbulence, leading to prediction delays or errors. To address this, this project proposes a forecasting framework integrating large language models (LLMs) with fluid dynamics simulations, and has developed a corresponding mobile application (Figure 16). The system allows users to ask natural language questions (e.g., "How will air quality around Lujiazui change in the next two hours?"). Multimodal prompt engineering—the core innovation—grounds LLM responses in first-principles physical theories (atmospheric stability patterns from simulations), geographic data, and real-time metrics (stability, temperature, CO, $PM_{2.5}$). The output module then analyzes and predicts pollutant dispersion within the microenvironment.

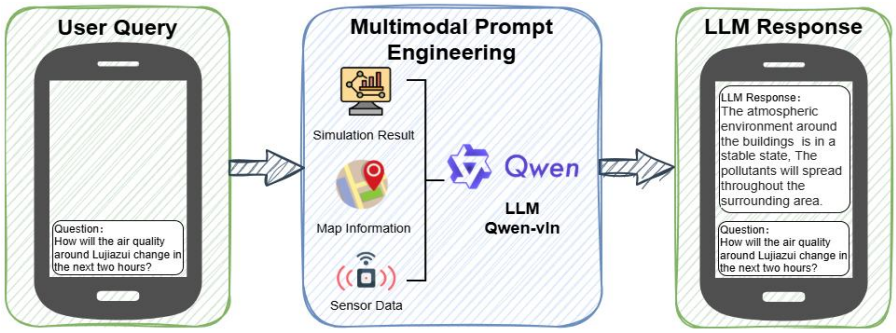


Figure 16. Large Language Model Workflow.

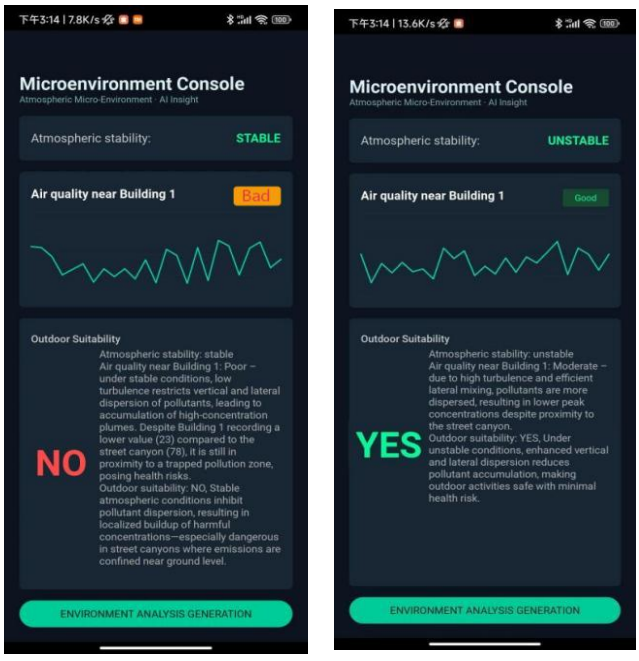


Figure 17. Analysis of the APP under Different Atmospheric Stability Conditions.

In the mobile prototype, this project uses Python + LangChain with the Qwen-plus model. The backend creates a conversational LLM interface and uses prompt templates to generate multimodal system prompts integrating three information types: (1) CFD-derived atmospheric stability and dispersion patterns (env_insight); (2) real-time or experimental CO sensor data (sensor_data); and (3) base64-encoded map images of street networks, canyons, and residential areas. Atmospheric stability is initially preset (e.g., "unstable") and can later be auto-determined by temperature sensors. CO data is formatted as structured text (e.g., "street canyon concentration / near-building concentration"). Using LangChain utilities, the system combines user natural language queries (e.g., "What is the environmental condition near building 4?") with multi-source inputs. Qwen-plus then generates explanatory responses on pollution dispersion and exposure risks under physical and sensor constraints. This shows that open-source frameworks can build a full prototype pipeline integrating CFD, sensor data, map imagery, and LLM inference, laying the foundation for lightweight mobile environmental intelligence tools.

After physical validation, the LLM-based micro-environment analysis system was tested. Figure 17 shows the mobile app's conclusions for the same query (building 4) under unstable vs. stable conditions. Under unstable conditions (Figure 17a), the system noted high mixing and low CO near building 4, concluding "effective dilution" and "low exposure risk"—matching experimental observations of rapid vertical dispersion. Under stable conditions (Figure 17b), the system identified weak mixing and retention, concluding "poor diffusion efficiency" and "long-term downstream risk"—consistent with near-ground haze accumulation. These examples demonstrate that the system integrates physics, sensor data, and map information into a unified reasoning framework, producing interpretable, context-aware environmental analysis.

3.3. Experimental Verification of the Intelligent UAV Monitoring System

To evaluate the context-aware human-machine collaborative drone system, this study established a two-tier framework combining computer simulations and field experiments, assessing motion algorithm validation, source localization accuracy, prediction precision, and natural language reliability—all repeated three times. For motion validation, the Gazebo simulation system established a drone flight model and terrain (Figure 18); through continuous testing, takeoff trajectory matched preset parameters each time. Field experiments were conducted in a closed 10 m × 10 m outdoor facility (25 ± 1°C, 50 ± 5% RH, wind <0.5 m/s). Four smoke generators (S1–S4) simulated pollution sources with CO emission rate of 0.01 g/s. The drone used flight parameters: cruise altitude 10 m, speed 1 m/s, sampling interval 1 m, sensor frequency 1 Hz. Data processing used a laptop (Intel i7-12700H, 16 GB RAM) with ChatGPT-4 and a 10-s timeout. Thermal maps showed CO locations matching actual smoke patches. System errors stem from three sources: (1) sensor errors controlled within 5% after calibration; (2) positioning errors—GPS-RTK provides cm-level accuracy, with attitude-induced shifts limited to <3% using an active stabilizer; (3) wind field effects—when wind exceeds 5 m/s, errors exceed 15%, triggering automatic return.

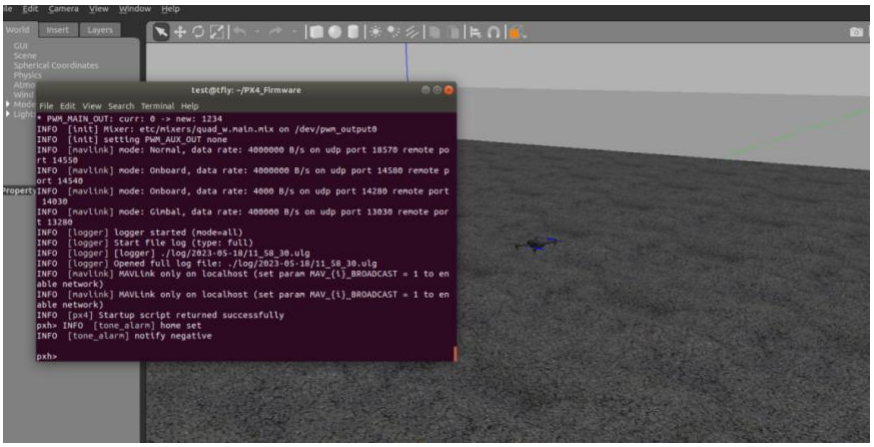


Figure 18. Gazebo Simulation Image.

4. Discussion

At the micro scale, CFD simulations and control experiments confirm that atmospheric stability dominates pollutant dispersion in "street canyon–residential area" systems. The inversion layer under stable stratification suppresses vertical mixing, causing persistent near-surface accumulation—CO concentrations in stable canyons exceed 80 μg/m³, compared to 5–35 μg/m³ under unstable conditions, aligning with real-world data [19]. Unlike previous studies focused on isolated canyons, this work elucidates transport pathways to residential areas and highlights long-term health risks under stable conditions. At the macro scale, a context-aware drone system overcomes fixed station limitations: natural language-driven navigation achieves response speeds ten times faster than manual inspections, XGBoost outperforms LSTM and Random Forest [12], and integrated LLM/RAG technologies lower operational barriers [15]. However, micro-scale ice-block experiments cannot fully replicate real atmospheric radiation (future LES needed), and macro-scale tests were limited to a controlled 10 m × 10 m area—field validation in real industrial parks under dynamic conditions is required.

From a cross-scale integration perspective, this study is the first to successfully couple microscopic physical mechanisms with macroscopic intelligent monitoring. Diffusion patterns from microscopic CFD simulations provide a physical foundation for macroscopic early warning systems, while large-scale drone data enable boundary condition refinements for microscopic models. This bidirectional feedback breaks traditional macro-micro separation, establishing a "mechanism-data"-driven environmental monitoring paradigm. Future work will enhance nested coupling between macroscopic meteorological models and microscopic CFD, enabling full-scale pollution prediction from urban to building-cluster scales and supporting millisecond-level responses via lightweight edge models. Beyond technical performance, the framework promotes environmental information equity by converting professional data into understandable health recommendations via a mobile app, empowering residents in underserved communities. The synergy of micro-CFD and macro-drone data is a unique feature: drone-collected high-resolution gridded data can serve as CFD boundary conditions, while CFD physical insights guide adaptive drone missions (e.g., increased patrol frequency during stable nocturnal periods)—a bidirectional feedback rarely addressed in previous work.

5. Conclusion and Future work

This study develops an LLM-enhanced multi-scale framework bridging micro-scale physical mechanisms and macro-scale intelligent sensing to address urban environmental monitoring's scale mismatch. At the micro scale, CFD simulations and experiments reveal that stable stratification creates an inversion lid suppressing vertical mixing, causing ground-level CO accumulation (>80 μg/m³ in street canyons) and transport to residential areas, while unstable conditions enhance convection (5–35 μg/m³)—enabling an LLM-CFD system for interpretable real-time assessments. At the macro scale, a context-aware drone framework achieves natural language-driven navigation, 0.34 m source localization, 2.3 μg/m³ CO prediction RMSE, and <3-minute response, with drone data feeding back to calibrate micro-scale models. This proves that integrating LLMs with physics-informed modeling and autonomous sensing establishes a "mechanism-data" paradigm, breaking traditional micro-macro isolation. Practical applications include autonomous industrial park inspections, real-time pollution warnings for sensitive urban areas, and carbon accounting for zero-carbon demonstration parks (e.g., Shanghai Minhang).

The framework will evolve toward full-scale, intelligent, adaptive monitoring. Technically, it will integrate satellite, drone, and ground data into city-level digital twins with edge-based LLMs for millisecond response and multi-agent coverage of square-kilometer areas. Socially, it will become an urban ecological brain, translating data into public health guidance to promote environmental equity and build low-carbon, harmonious cities.

References

[1] Leelőssy, Ádám, Ferenc Molnar Jr, Ferenc Izsák, Ágnes Havasi, István Lagzi, and Róbert Mészáros. "Dispersion modeling of air pollutants in the atmosphere: a review." *Central European Journal of Geosciences* 6, no. 3 (2014): 257-278.

[2] Blocken, Bert. "Computational Fluid Dynamics for urban physics: Importance, scales, possibilities, limitations and ten tips and tricks towards accurate and reliable simulations." *Building and environment* 91 (2015): 219-245.

[3] Sanchez, Beatriz, Jose-Luis Santiago, Alberto Martilli, Magdalena Palacios, and Frank Kirchner. "CFD modeling of reactive pollutant dispersion in simplified urban configurations with different chemical mechanisms." *Atmospheric Chemistry* and *Physics* 16, no. 18 (2016): 12143-12157.

[4] Wang, Qun, Jian Hang, Yifan Fan, and Yuguo Li. "Urban plume characteristics under various wind speed, heat flux, and stratification conditions." *Atmospheric Environment* 239 (2020): 117774.

[5] Huang, Yuan-Dong, Ren-Wei Hou, Ze-Yu Liu, Ye Song, Peng-Yi Cui, and Chang-Nyung Kim. "Effects of wind direction on the airflow and pollutant dispersion inside a long street canyon." *Aerosol and air quality research* 19, no. 5 (2019): 1152-1171.

[6] An, K., J. C. H. Fung, and S. H. L. Yim. "Sensitivity of inflow boundary conditions on downstream wind and turbulence profiles through building obstacles using a CFD approach." *Journal of Wind Engineering and Industrial Aerodynamics* 115 (2013): 137-149.

[7] Liu, Fan, Hua Qian, Zhiwen Luo, and Xiaohong Zheng. "The impact of indoor thermal stratification on the dispersion of human speech droplets." *Indoor Air* 31, no. 2 (2021): 369-382.

[8] Masoumi-Verki, Shahin, Payam Gholamalipour, Fariborz Haghighat, and Ursula Eicker. "Embedded LES of thermal stratification effects on the airflow and concentration fields around an isolated high-rise building: Spectral and POD analyses." *Building and Environment* 206 (2021): 108388.

[9] Bazdidi-Tehrani, Farzad, Payam Gholamalipour, Mohsen Kiamansouri, and Mohammad Jadidi. "Large eddy simulation of thermal stratification effect on convective and turbulent diffusion fluxes concerning gaseous pollutant dispersion around a high-rise model building." *Journal of Building Performance Simulation* 12, no. 1 (2019): 97-116.

[10] Huang, Yuan-Dong, Nuo Xu, Su-Qi Ren, Li-Bing Qian, and Peng-Yi Cui. "Numerical investigation of the thermal effect on flow and dispersion of rooftop stack emissions with wind tunnel experimental validations." *Environmental Science and Pollution Research* 28, no. 9 (2021): 11618-11636.

[11] Liu, Xiaoping, Mei Wu, Ziying An, and Tao Chen. "Study on the combined effect of wind and buoyancy on cross-unit contamination around a high-rise residential building." *Sustainable Cities and Society* 82 (2022): 103860.

[12] Cui, Peng-Yi, Zhuo Li, and Wen-Quan Tao. "Investigation of Re-independence of turbulent flow and pollutant dispersion in urban street canyon using numerical wind tunnel (NWT) models." *International Journal of Heat and Mass Transfer* 79 (2014): 176-188.

[13] Wang, Yong, Ke Zhong, Ningbo Zhang, and Yanming Kang. "Numerical analysis of solar radiation effects on flow patterns in street canyons." *Engineering Applications of Computational Fluid Mechanics* 8, no. 2 (2014): 252-262.

[14] Ioannidis, Giannis, Paul Tremper, Chaofan Li, Till Riedel, Nikolaos Rapkos, Christos Boikos, and Leonidas Ntziachristos. "Evaluating the spatial coverage of air quality monitoring stations using computational fluid dynamics." Atmosphere 16, no. 3 (2025): 326.

[15] Villa, Tommaso Francesco, Felipe Gonzalez, Branka Miljievic, Zoran D. Ristovski, and Lidia Morawska. "An overview of small unmanned aerial vehicles for air quality measurements: Present applications and future prospectives." *Sensors* 16, no. 7 (2016): 1072.

[16] Javan, Farzaneh Dadrass, Farhad Samadzadegan, and Ahmad Toosi. "Air pollution observation—bridging spaceborne to unmanned airborne remote sensing: a systematic review and meta-analysis." *Air Quality, Atmosphere & Health* 18, no. 8 (2025): 2481-2549.

[17] Zhang Y, Wang L, Chen H, et al. VOCs emission source identification in chemical industrial parks using UAV-based PID sensor and machine learning[J]. Journal of Hazardous Materials, 2024, 472: 134567. https://doi.org/10.1016/j.jhazmat.2024.134567

[18] Li, X., Zhang, Q., Zhang, Y. et al., Attribution of PM2.5 Exposure in Beijing-Tianjin-Hebei Region to Emissions: Implication to Control Strategies, Science Bulletin, 2017, 62 (13): 957-964.

[19] Jeanjean, Antoine PR, Riccardo Buccolieri, James Eddy, Paul S. Monks, and Roland J. Leigh. "Air quality affected by trees in real street canyons: The case of Marylebone neighbourhood in central London." *Urban Forestry & Urban* Greening 22 (2017): 41-53.