

Data-driven operator learning for energy-efficient building control

Yuxin Bian, Yuanyuan Shi*

Department of Electrical and Computer Engineering, University of California San Diego, La Jolla, 92037, USA

ARTICLE INFO

Keywords:

Building energy systems
Ventilation control
Energy efficiency
Indoor air quality
Neural operator learning

ABSTRACT

Energy-efficient ventilation control plays a vital role in reducing building energy consumption while ensuring occupant health and comfort. While Computational Fluid Dynamics (CFD) simulations offer high-fidelity modeling of airflow for building HVAC design, their high computational cost makes them impractical for practical adoption in real-time building management system. In this work, we present a data-driven framework that combines the physical accuracy of CFD with the computational efficiency of machine learning to enable energy-efficient building ventilation control. Our method jointly optimizes airflow supply rates and vent angles to reduce energy use and adhere to air quality constraints. We train a neural operator transformer to learn the mapping from building control actions to airflow field distributions using high-resolution CFD data. This learned operator enables a gradient-based control framework capable of optimal decision-making. Experimental results demonstrate that our approach achieves substantial energy savings compared to maximum airflow rate control, rule-based control, and data-driven control based on regional average CO₂ predictions, while consistently maintaining safe indoor air quality. These results highlight the practicality and scalability of our method for enabling safe and energy-efficient building management.

1. Introduction

1.1. Background and motivation

Buildings account for nearly 40% of global energy consumption [1], with Heating, Ventilation, and Air Conditioning (HVAC) systems being among the primary contributors. As the demand for sustainable and intelligent energy solutions rises, ventilation systems have emerged as a critical area for optimization [2]. Proper ventilation not only reduces energy consumption but also plays a vital role in maintaining indoor air quality, directly impacting occupant health and overall comfort. In the wake of the COVID-19 pandemic, attention to building ventilation control has grown considerably [3]. Public health agencies and building operators increasingly recognize the need for adaptive ventilation systems that respond to occupancy and pollutant levels. However, the challenge lies in achieving this adaptivity without incurring excessive energy costs. A smarter, data-driven approach to ventilation is urgently needed—one that balances energy consumption and health in real-time.

Despite their potential, most buildings still rely on fixed or rule-based HVAC control strategies [4, 5, 6]. For example, beginning in spring 2020, Facilities Management at UC San Diego implemented a policy of maximum fresh-air intake with minimal or no recirculation during office hours [6], which led to a 2–2.5× increase in building energy consumption compared to nominal levels. While these ventilation strategies aim to improve indoor air quality, they often do so at the expense of significantly higher energy costs.

More advanced data-driven approaches, such as Model Predictive Control (MPC) [7, 8, 9, 10], or Reinforcement Learning(RL) [11, 12, 13, 14] have been explored to improve energy efficiency and air quality. However, most existing approaches predict only a single-point or spatially averaged CO₂ concentration to represent indoor air quality dynamics.

These approaches fail to capture the spatial and temporal dynamics of how control actions influence air quality, limiting their ability to make targeted adjustments. As a result, they can over-ventilate the entire space to compensate for unseen variations, leading to increased energy consumption.

To design effective and energy-efficient building management, it is critical to model and manage the *spatial-temporal* distribution of airflow and contaminants within indoor spaces. Inadequate spatial modeling not only leads to compromised indoor air quality but also causes significant energy waste, as over-ventilation or inefficient airflow distribution demands unnecessary HVAC power [15]. Computational Fluid Dynamics (CFD) provides this capability through partial differential equations (PDEs), offering high-resolution simulations of airflow behavior [15, 16, 17, 18]. However, solving PDEs using classical numerical techniques, such as finite element, finite volume, or spectral methods, is computationally prohibitive [19]. As a result, CFD modeling has been predominantly applied to HVAC system design and offline indoor air quality analysis, rather than real-time online control or operational deployment aimed at optimizing energy use in buildings. This gap between high-fidelity airflow modeling and real-time, energy-focused control remains a major bottleneck in advancing intelligent and sustainable building HVAC systems. There is a critical need for a framework that preserves the physical fidelity of CFD while enabling fast inference and optimization for real-time building energy management and ventilation control.

1.2. Literature review

Building HVAC control. Building control with indoor air quality considerations often presents a trade-off: enhancing air quality typically requires greater ventilation rates, which in turn increase energy consumption. Balancing these competing demands remains a central challenge in intelligent

*Corresponding author

ORCID(s): 0000-0002-6182-7664 (Y. Shi)

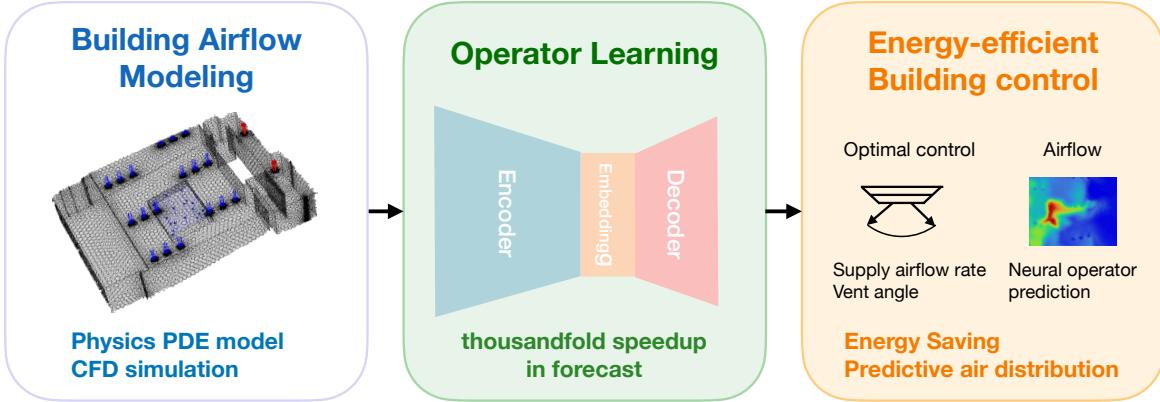


Figure 1: Schematic of our data-driven operator learning framework for energy-efficient ventilation control. Computational fluid dynamics (CFD) simulations are used to model complex 3D airflow and CO₂ spatiotemporal dynamics. A neural operator transformer is trained to learn the mapping between ventilation control actions and airflow field evolution. Leveraging high-fidelity simulation data, our approach enables real-time optimization of ventilation strategies to minimize energy consumption while maintaining indoor air quality.

HVAC control systems. To inform ventilation control decisions, indoor air quality dynamics are commonly approximated by modeling the temporal evolution of a single-point or spatially-averaged CO₂ concentration, either through simple data-driven methods or ordinary differential equation (ODE) models. For instance, Li et al. [20] modeled indoor CO₂ dynamics using ordinary differential equations (ODEs) and optimized the supply airflow rates through MPC. Li et al. [21] proposed to predict future CO₂ with a deep neural network, and optimized fan pressure and damper positions based on the average predictions. Shang et al. [14] modeled indoor PM2.5 concentrations with an ODE and employed reinforcement learning to control the air purifier exchange rate, improving energy efficiency while maintaining air quality. However, such ODE-based models neglect spatial variations, which either create poorly ventilated dead zones [3] that risk occupants' health and comfort, or lead to conservative control with high energy consumption [15].

To accurately model the airflow dynamics, partial differential equations (PDE)-based models including computational fluid dynamics (CFD) simulations should be adopted. Hosseinkoo et al. [3] modeled 2D pathogen concentration using convection-diffusion equations and optimized the velocity field through RL. However, directly optimizing the entire velocity field is impractical, as control is limited to the boundary airflow velocity at the supply air vents. Bian et al. [15] modeled indoor airflow and CO₂ dynamics using the Navier–Stokes and convection–diffusion equations, and optimized the ventilation control actions through PDE-constrained optimization. However, the proposed adjoint-based method is computationally expensive and was only demonstrated in a two-dimensional office room setting. Ning et al. [16] used CFD to study airflow fields and CO₂ concentrations within a bedroom to inform the design of supply outlet heights. However, due to the significant computational burden of CFD solvers, most existing work remains limited

to offline HVAC system design and analysis, rather than supporting online building control and optimization.

Operator learning. Recent advances in machine learning have demonstrated significant potential in accelerating the solution of PDEs, particularly for fluid dynamics problems. Neural operators, which learn mappings between functions defined on continuous domains, such as spatiotemporal processes and PDEs, have emerged as a powerful tool in this domain. Deep Operator Networks (DeepONets) [22] pioneered this direction by learning mappings between input functions and the solution operators of PDEs. Building on this, Fourier Neural Operators (FNO) [23] leverage the spectral domain to capture global dependencies in physical systems, demonstrating success in fast turbulent flow prediction and weather forecasting. Neural Operator transformer (GNOT) [24] further adapted these ideas to irregular geometries, particularly suitable for complex indoor environments. In the context of fluid dynamics, neural operator learning has demonstrated remarkable computational efficiency and accuracy trade-offs, enabling prediction times of milliseconds compared to hours for traditional CFD while maintaining acceptable accuracy. Recent studies have successfully applied these approaches to car pressure prediction [25], turbulent channel flow [26], and indoor airflow prediction [27].

1.3. Contributions and innovations

In this work, we propose a data-driven operator learning framework for energy-efficient building ventilation control, as illustrated in Figure 1. To the best of our knowledge, this is the first work that leverages operator learning to model the spatial-temporal room airflow dynamics for real-time building ventilation control. Our goal is to determine the optimal ventilation control actions including supply airflow rates and vent angles to minimize energy consumption while maintaining indoor air quality over the space. To achieve this, we develop a high-fidelity CFD simulation

dataset for building ventilation control, based on a real-world classroom environment. We then leverage operator learning to model the mapping from ventilation control actions to indoor airflow field distributions. The resulting surrogate model enables fast, gradient-based optimization of building control strategies, achieving significant energy savings while maintaining indoor air quality. Our key contributions are summarized as follows:

- Ensemble neural operator transformer: We propose to learn an ensemble neural operator transformer model to predict building airflow dynamics. This model achieves a 10.9% relative error and a remarkable speed-up of 250,000 times compared to CFD simulations, enabling accurate indoor air quality prediction with high computational efficiency.
- Ventilation control optimization: We optimize building ventilation control actions by integrating the learned neural operator model into an optimization-based control framework, that reduces 10%-50% energy consumption while maintaining air quality constraints compared against various building control methods.
- The performance has been verified in both a one-step building control scenario across three occupancy levels and a multi-step building control scenario with time-varying occupancy via simulation of a real-world classroom. We also develop and release a high-fidelity [building dataset](#), based on simulation of the classroom HVAC control. This dataset captures diverse ventilation actions and occupancy levels, providing a benchmark for training and evaluating machine learning models in building ventilation control applications.

2. Problem formulation

This work aims to develop a data-driven framework that integrates PDE-based models together with CFD simulations to capture the spatiotemporal dynamics of CO₂ in indoor environments and enable energy-efficient ventilation control. The CFD simulation is developed for a real-world classroom located in University of California, San Diego. The classroom measures approximately 19 × 13 × 3.5 meters and is equipped with a ceiling-mounted ventilation system. The ventilation system includes 2 outlet vents and 18 inlet vents grouped into six zones, each allowing independent control of airflow rate and supply airflow angle to optimize energy efficiency while maintaining high indoor air quality. Figure 2(a) illustrates the physical classroom and its corresponding CFD model representation, and Figure 2(b) presents an example of the CO₂ concentration and velocity field visualization.

2.1. Governing equations for CO₂ dynamics

In this section, we primarily introduce the governing equation for CO₂ transport, while the CFD simulation setup along with the dataset description is detailed in Section 3.

Let $\Omega \subset \mathbb{R}^3$ be the spatial domain of interests, and let $t \in \mathbb{R}^+$ denote time. We define $C(\mathbf{x}, t)$ as the CO₂ concentration field at spatial location $\mathbf{x} \in \Omega$ and time t , $m(t) = [m^r(t) \ m^a(t)] \in \mathbb{R}^{12}$ as the control actions (airflow rate $m^r \in \mathbb{R}^6$ and airflow angle $m^a \in \mathbb{R}^6$) for six groups of supply vents, and $n_p(t) \in \mathbb{R}$ as the occupancy level. The distribution of CO₂ in an indoor environment follows the advection-diffusion equation [15, 28], expressed as

$$\frac{\partial C(\mathbf{x}, t)}{\partial t} + \mathbf{u}(\mathbf{x}, t) \cdot \nabla C(\mathbf{x}, t) = D_{\text{eff}} \nabla^2 C(\mathbf{x}, t) + S(\mathbf{x}, t), \quad (1)$$

where

- $C(\mathbf{x}, t)$ is the CO₂ concentration (ppm) at location \mathbf{x} and time step t .
- $\mathbf{u}(\mathbf{x}, t)$ is the airflow velocity field obtained from CFD simulations, by solving the incompressible Navier-Stokes equations. The boundary conditions at the supply vents depend on the ventilation control $m(t)$, thereby allowing the control action to influence $\mathbf{u}(\mathbf{x}, t)$ throughout the domain. The numerical model for the airflow velocity field is presented in the Appendix.
- D_{eff} is the diffusion coefficient for CO₂ in the air.
- $S(\mathbf{x}, t)$ represents the CO₂ source term, which models occupant-generated CO₂ on the designated occupancy surface. This source depends on the occupancy level $n_p(t)$ and we assume that the exhaled air rate is 6 L/min per person following [29].

2.2. Data-driven modeling with neural operator

Solving PDEs numerically is computationally expensive, making traditional CFD models impractical for real-time building control applications. To address this limitation, we aim to learn a neural operator \mathcal{G}_θ that efficiently maps historical CO₂ concentrations and control actions to future CO₂ concentration distributions.

Formally, the neural operator learns the mapping:

$$\mathcal{G}_\theta : (C(\mathbf{x}, \tau)_{\tau \in [t-H, t]}, m, n_p) \mapsto C(\mathbf{x}, \tau)_{\tau \in (t, t+T]} \quad (2)$$

where $C(\mathbf{x}, \tau)_{\tau \in [t-H, t]}$ is the historical CO₂ fields over period $[t - H, t]$ and $C(\mathbf{x}, \tau)_{\tau \in (t, t+T]}$ is the predicted future CO₂ fields over the future interval $(t, t + T]$. Our forecasting approach incorporates historical CO₂ concentrations to account for temporal dependencies inherent to the system's physics (e.g., diffusion and advection dynamics). We further assume that the control m and occupancy n_p remain fixed over $(t, t + T]$. In many PDE-based fluid flow and CO₂ transport scenarios, transient dynamics evolve quickly over short time horizons—on the order of a few minutes. In practical building control, however, ventilation settings and occupancy often remain unchanged over these short intervals (e.g., a fixed HVAC setpoint during a 5-minute period [6]). Consequently, the modeling choice aligns with the typical operational strategy in real buildings.

In practice, we rarely have continuous access to the CO₂ distribution over space and time. Instead, we rely on

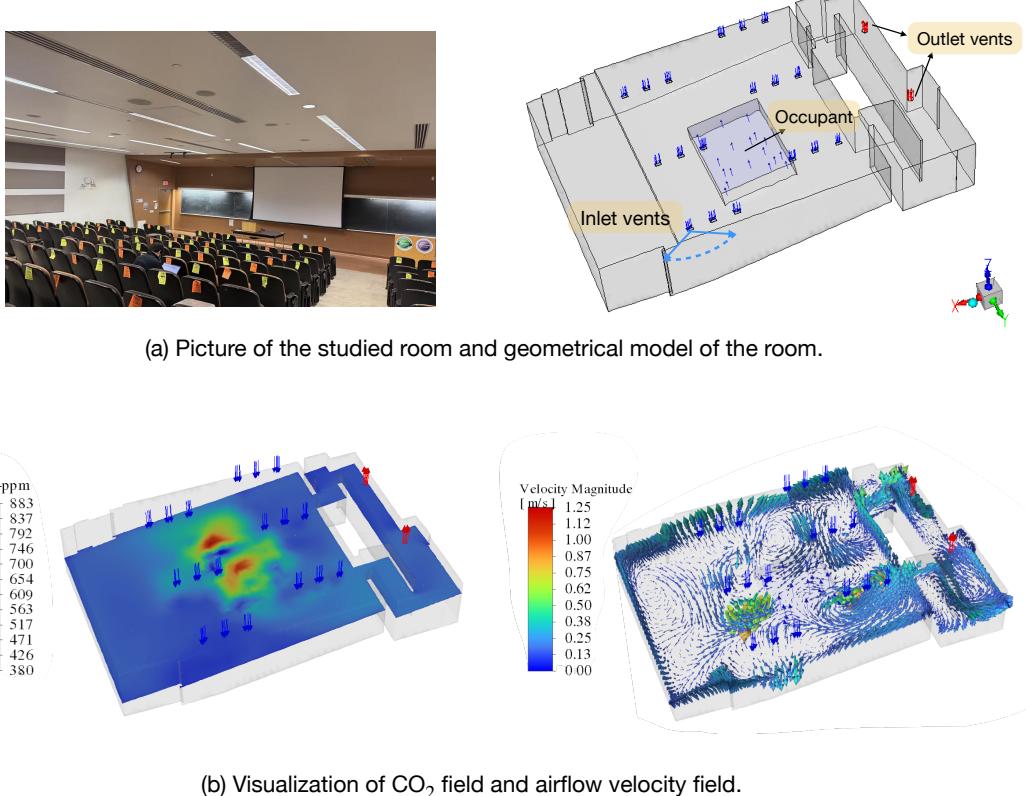


Figure 2: (a) The picture of the studied room and the geometry of the CFD model: a classroom with a ventilation system including 18 inlet vents and 2 outlet vents on the ceiling. (b) Visualization of the CFD simulation results of CO₂ concentration and airflow velocity fields - one example from the developed CFD dataset.

a discretized approximation of the underlying functions. For simplicity and consistency, we reuse the symbols t, H, T as discrete time indices. Let $\{x_i\}_{1 \leq i \leq N_x}$ be a spatial grid and $C_i^t \in \mathbb{R}$ be the CO₂ concentration at spatial grid point x_i and time step t . We collect the recent history of CO₂ fields over the last H time steps, along with the fixed parameters m and n_p . Define the discrete input set

$$\mathcal{A} = \{(x_i, C_i^{t-H:t})\}_{1 \leq i \leq N_x} \cup \{m\} \cup \{n_p\}, \quad (3)$$

although \mathcal{A} is finite-dimensional, it represents the sampled version of the underlying function $C(\mathbf{x}, \tau)$ over the history window $[t - H, t]$. From this discrete representation, the neural operator \mathcal{G}_θ produces predicted future concentrations $\hat{C}_{1 \leq i \leq N_x}^{t+1:t+T}$ (with short-hand notation \hat{C}), which approximate the continuous output function $\hat{C}(\mathbf{x}, \tau)$ for $\tau \in (t, t + T]$:

$$\hat{C}_{1 \leq i \leq N_x}^{t+1:t+T} = \mathcal{G}_\theta(\mathcal{A}). \quad (4)$$

2.3. Ventilation control optimization

The learned neural operator \mathcal{G}_θ can be integrated into the building ventilation control optimization problem. The resulting optimization problem is formulated in (5).

$$\min_{m=[m^r, m^a]} w_1 \|\hat{C} - C_{\text{target}}\|_2^2 + w_2 \|m - m^{(0)}\|_2^2 + w_3 \|m^r\|_1, \quad (5a)$$

$$\text{s.t. } \hat{C} \leftarrow \mathcal{G}_\theta(\{C_{1 \leq i \leq N_x}^{t-H:t}\} \cup \{m\} \cup \{n_p\}), \quad (5b)$$

$$\underline{m^r} \leq m^r \leq \overline{m^r}, \underline{m^a} \leq m^a \leq \overline{m^a}. \quad (5c)$$

In the objective function (5a), the first term quantifies the deviation between predicted CO₂ concentrations \hat{C} from Eqn (4) and the desired setpoint C_{target} over a prediction horizon T . This term is commonly used in indoor air quality ventilation control studies, as seen in prior works [20, 30]. The second term penalizes deviations of the optimized control actions from the previous control action $m^{(0)}$ that resulted in the past CO₂ concentration. This is important in building management, where large deviations in control actions are typically avoided to maintain system stability and ensure smooth operation [15, 31]. By encouraging control actions to remain close to $m^{(0)}$, we also reduce the risk of extrapolating beyond the model's learned domain, thereby increasing the reliability of the predictions. The third term represents the energy consumption, measured through the L1 norm of the ventilation rate [15]. The coefficients w_1, w_2, w_3 balance the relative importance of these objectives.

Constraint (5b) enforces that future CO₂ concentrations follow the dynamics predicted by the learned neural operator, while constraint (5c) ensures that both mechanical ventilation rates m^r and vent angles m^a remain within their physical limits.

Table 1

Data fields in the BEAR-CFD dataset pickle (.pkl) format.

Field	Description
HVAC surface	ndarray (N_{hvac} , 3), spatial coordinates of grid points on HVAC surfaces
CO_2 -HVAC	ndarray (N_{hvac}, \bar{T}), CO_2 concentration time series at HVAC surface
People surface	ndarray (N_{people} , 3), spatial coordinates of grid points on people surfaces
CO_2 -People	ndarray ($N_{\text{people}}, \bar{T}$), CO_2 concentration time series at people surface
steady case	int, Identifier for the initial steady-state condition used in the simulation
n_p	int, number of occupant
m_i^r	float, airflow rate (m^3/s) for i -th group of vents
m_i^a	float, angle ($^\circ$) for i -th group of vents

3. The BEAR-CFD data

In this section, we describe the BEAR-CFD dataset, which we will publish in conjunction with this work, along with its collection methodology.

3.1. BEAR-CFD data

The BEAR-CFD data is generated through a structured CFD simulation workflow designed to capture the spatiotemporal dynamics of CO_2 concentrations under varying ventilation and occupancy conditions, which will be available at <https://ucsdsmartbuilding.github.io/>.

The dataset was generated using ANSYS FLUENT 2023R2 and includes both steady-state and transient flow simulations. Specifically, it comprises 10 steady-state simulation cases, each converging to a distinct time-independent solution under unique boundary conditions, providing insights into equilibrium flow behavior. Additionally, the dataset contains 300 transient simulation cases, each initialized from a steady-state solution and spanning a 30-minute physical duration. The simulations use a time step size of 10 seconds, with results recorded every 30 seconds, resulting in 60 time steps per simulation. Transient simulations capture the temporal evolution of airflow dynamics, reflecting unsteady effects influenced by varying boundary conditions and flow interactions. We leverage this transient data to train our neural operators.

The dataset is distributed in both ANSYS FLUENT's native format (.cas and .dat files) and Python pickle (.pkl) format. The FLUENT files preserve the complete simulation environment and solution data, while the pickle format enables efficient programmatic access to the numerical results through standard Python libraries. The dataset includes spatiotemporal fields of CO_2 concentrations, airflow rates, vent angles, and occupancy, as detailed in Table 1.

3.2. CFD simulation setup

The dataset was collected through CFD simulations of a physical classroom environment at the University of California San Diego shown in Figure 2 (a).

3.2.1. Computational domain

The 3D meshed model of the indoor space fluid domain of the classroom with a simplified occupancy surface was generated with tetrahedral element using the ANSYS meshing tools. The computational domain represents a classroom, equipped with a ventilation system comprising 18 rectangular air inlets (each measuring $0.1349 \text{ m} \times 0.3048 \text{ m}$) organized into 6 groups, along with 2 outlets. The fluid domain was discretized into approximately 0.24 million tetrahedral elements to ensure accurate resolution of airflow and CO_2 transport dynamics for the CFD modeling and simulation. A transparent view of the 3D model, highlighting the mesh structure and key features, is shown in Fig. 2(a) on the right.

3.2.2. Governing equations

The simulations solved the incompressible Navier-Stokes equations for airflow dynamics, coupled with the species transport equations to track the distribution of CO_2 , O_2 , H_2O , and N_2 . Turbulence was modeled using the $k-\omega$ SST (Shear Stress Transport) model, which provides accurate predictions of near-wall airflow behavior and is well-suited for indoor environments [17]. The governing equations were discretized using a finite volume method, with second-order schemes for momentum and species transport.

3.2.3. Boundary conditions

The boundary conditions are defined as follows [18]:

- Inlet: A velocity-inlet condition is applied to the 18 rectangular inlet vents. The maximum airflow rate is set to 10 Air Changes per Hour (ACH) based on operational data from Pepper Canyon Hall, where the classroom is located at. The maximum supply airflow velocity is thus calculated to be $m^r = 3.24 \text{ m/s}$. The CO_2 concentration at the inlet is fixed at $C_{\text{fresh}} = 400 \text{ ppm}$, which is the CO_2 levels of fresh outdoor air.
- Outlet: A pressure-outlet condition is applied to the two exhaust vents.
- Occupancy: Human CO_2 exhalation is modeled as a mass-flow-inlet condition at occupant locations. Occupants are positioned within a designated rectangular region to reflect typical indoor spatial distributions, see Figure 2(a). The CO_2 concentration at these sources is set to $C_{\text{occupancy}} = 40,000 \text{ ppm}$, representing typical exhaled air, and the exhaled air rate is 6 L/min per person.
- Walls: We treat walls with the no-slip condition, meaning the airflow velocity at the wall is zero, which realistically represents friction and airflow resistance near solid surfaces.

3.3. Data collection

To generate the dataset, control parameters (supply airflow rates and angles) and occupancy levels were randomly sampled from uniform distributions. The six groups of inlet vents were modeled to operate independently, with the following parameters:

$$\begin{aligned} m_i^r &\sim U[\underline{m}^r, \bar{m}^r], m_i^a \sim U[45^\circ, 135^\circ], i \in [1, \dots, 6], \\ n_p &\sim U[10, 80], \end{aligned} \quad (6)$$

where m_i^r is the airflow rate for the i -th group of vents, bounded between $\underline{m}^r = 0.324\text{m/s}$ (10% of maximum) and $\bar{m}^r = 3.24\text{m/s}$; m_i^a is the airflow angle of the i -th group of vents, spanning 45° to 135° ; and n_p is the number of occupants in the classroom.

The simulations were conducted in two phases:

- Steady-state initialization: 10 steady-state simulations were performed with randomized parameters to generate equilibrium airflow and CO₂ concentration fields. These solutions served as physically consistent initial conditions for transient simulations.
- Transient simulation: 300 transient simulations were executed. For each simulation: a steady-state solution was randomly selected as the initial condition, control parameters $m_i^r, m_i^a, \forall i$ and occupancy n_p were randomly sampled from the uniform distributions, then the evolution of CO₂ concentrations was simulated and recorded at 30-second intervals over $\bar{T} = 60$ time steps (30 minutes).

We note that CO₂ concentrations were monitored at two critical planes:

- HVAC surface: A horizontal plane at 2.9-meter height near the ventilation inlets, capturing the supply and returning air quality.
- People surface: A horizontal plane at 1.6-meter height (average breathing height for standing adults), representing air quality at occupant exposure levels [32].

In our CFD simulation, the sitting surface is near the occupancy boundary, leading to potential inaccuracies from numerical artifacts and boundary effects. Thus, we focused on heights where airflow and CO₂ dispersion are more reliably captured.

4. Methodology

In this work, we propose a data-driven operator learning framework to model the indoor air quality and optimize ventilation control for energy efficiency. The core component is an ensemble neural operator transformer architecture, \mathcal{G}_θ , shown in Figure 3. In figure 3(a), the neural operator architecture enables fast building CFD simulations by processing multiple inputs: query points, supply airflow rates, airflow angles, and historical CO₂ concentrations. The

trained operator learning model uses these inputs to predict future CO₂, which are then utilized to determine optimal ventilation control parameters by solving the optimization problem defined in Equation (5). In the following subsections, we will first explain the concept of operator neural transformer, then detail describe the control algorithm.

4.1. Ensemble neural operator transformer

Existing operator learning approaches, while effective in many applications, often struggle with limited training data. To address this limitation, we enhance the General Neural Operator Transformer (GNOT) [24] by ensemble learning [33]. The network architecture of the proposed model together with the control is illustrated in Figure 3. In the following subsections, we will describe the input encoding and ensemble learning for GNOT.

4.1.1. Input encoding

The model takes the input mesh, historical CO₂ concentrations, and control parameters, including control actions and the number of occupants, as inputs. To accommodate these heterogeneous inputs, a general encoder, highlighted in green in Figure 3, is employed to transform them into the feature embedding $Y \in \mathbb{R}^{N \times n_e}$, where N denotes an arbitrary number of input elements and n_e is the embedding dimension. The model employs simple multilayer perceptrons (MLPs), denoted as f_{w1}, f_{w2}, f_{w3} , to map each type of input to its corresponding embedding.

- Input mesh: A MLP maps the mesh points to query embeddings $Y_{\text{mesh}} = (f_{w1}(\mathbf{x}_i))_{1 \leq i \leq N_x} \in \mathbb{R}^{N_x \times n_e}$.
- Past CO₂ concentrations: At time t , we have $\{\mathbf{x}_i, c_i\}_{1 \leq i \leq N_x}$, where $c_i = C_i^{t-H:t}$ represents the historical CO₂ levels at position \mathbf{x}_i . A MLP encodes both positions and concentrations to produce the feature $Y_C = (f_{w2}(\mathbf{x}_i, c_i))_{1 \leq i \leq N_x} \in \mathbb{R}^{N_x \times n_e}$.
- Control parameters: A MLP encodes control parameters $[m, n_p] \in \mathbb{R}^{1^3}$ into embeddings $Y_{\text{param}} = f_{w3}([m, n_p]) \in \mathbb{R}^{1 \cdot n_e}$.

4.1.2. Ensemble learning of GNOT

As shown in Figure 3(a), GNOT begins by encoding input features and updating them using a heterogeneous normalized cross-attention layer, followed by a normalized self-attention layer to refine representations. To effectively capture spatial heterogeneity, GNOT incorporates a geometric gating mechanism that leverages the query point coordinates to compute a weighted combination of expert feed-forward networks (FFNs). The model stacks N such attention blocks to produce the final predictions.

While GNOT is originally designed to predict only the mean of the prediction, we enhance its robustness by introducing an ensemble-based extension. Specifically, instead of training GNOT to minimize mean squared error or relative error, we modify it to predict a probability distribution for

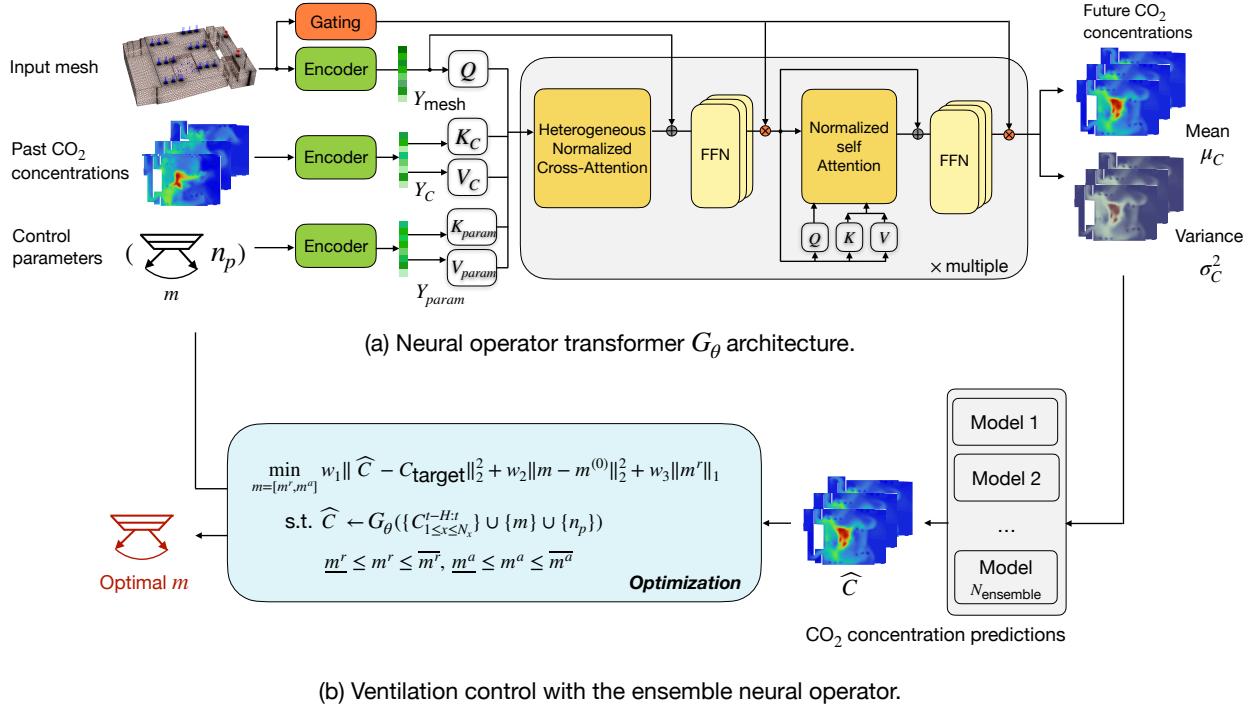


Figure 3: Overview of the proposed data-driven operator learning framework for energy-efficient ventilation control. The framework consists of two phases: (1) the learning phase, where neural operator transformers are trained to map past air field data and ventilation control parameters to future air field evolution, and (2) the control phase, where the trained ensemble neural operator is integrated into an optimization framework to solve the ventilation control problem. This approach enables real-time optimization of airflow supply rates and vent angles while maintaining air quality standards.

future CO₂ concentrations, characterized by the mean (μ_C) and variance (σ_C^2). Specifically, the model predicts:

$$\begin{aligned} \mu_C, \sigma_C^2 &= \mathcal{G}_\theta(\mathcal{A}), \\ \mu_C &= \mathbb{E}[\hat{C}_{1 \leq i \leq N_x}^{t+1:t+T}], \quad \sigma_C^2 = \text{Var}[\hat{C}_{1 \leq i \leq N_x}^{t+1:t+T}], \end{aligned} \quad (7)$$

where $[\mu_C]_i^t, [\sigma_C^2]_i^t$ represents the mean and variance at location i and time step t . The model is trained using the Negative Log-Likelihood (NLL) loss, defined as:

$$\mathcal{L} = \frac{1}{N_x T} \sum_{i=1}^{N_x} \sum_{k=1}^T \left(\frac{\log(2\pi[\sigma_C^2]_i^{t+k})}{2} + \frac{(C_i^{t+k} - [\mu_C]_i^{t+k})^2}{2[\sigma_C^2]_i^{t+k}} \right) \quad (8)$$

where $C_i^{t+k} \in \mathbb{R}$ is the true value at x_i and time $t+k$. By minimizing the NLL loss, the model learns to jointly optimize the mean and variance, effectively mitigating overfitting [33].

In addition, we train an ensemble of neural operator transformers, where the final prediction is obtained by averaging the outputs of multiple independently trained models. Let $\mu_C^{(n)}$ denote the mean predictions of the n -th model. We treat the ensemble as a uniformly-weighted mixture model and combine the predictions as follows:

$$\begin{aligned} \mu_C^{\text{ensemble}} &= \frac{1}{N_{\text{ensemble}}} \sum_{n=1}^{N_{\text{ensemble}}} \mu_C^{(n)}, \\ \mu_C^{(n)}, (\sigma_C^2)^{(n)} &= \mathcal{G}_{\theta_n}(\mathcal{A}), \end{aligned} \quad (9)$$

where N_{ensemble} represents the number of models in the ensemble, θ_n is the model parameters for n -th trained neural operator. This ensemble approach not only improves prediction accuracy but also enhances the reliability of ventilation control, as demonstrated in our experiments.

4.2. Control algorithm

With the learned ensemble neural operator transformer model, we are ready to solve the building control problem in (5). Recall that the objective function of the building control problem is defined as:

$$\mathcal{L}(m) = w_1 \|\hat{C} - C_{\text{target}}\|_2^2 + w_2 \|m - m^{(0)}\|_2^2 + w_3 T \|m^r\|_1 \quad (10)$$

where the first term $\|\hat{C} - C_{\text{target}}\|_2^2$ is defined as,

$$\|\hat{C} - C_{\text{target}}\|_2^2 = \frac{1}{N_x T} \sum_{k=1}^T \sum_{i=1}^{N_x} (\hat{C}_i^{t+k} - C_{\text{target},i}^{t+k})^2$$

and w_1, w_2, w_3 are weighting coefficients. \hat{C} represents the predicted CO₂ concentrations generated by the ensemble neural operator (9):

$$\hat{C} = \mu_C^{\text{ensemble}}. \quad (11)$$

To solve the building control problem in (5), we leverage the differentiability of the neural operator \mathcal{G}_θ and propose a gradient-based method to update the control actions for

	Model 1	Model 2	Model 3	Model 4	Model 5	Ensemble
l_2 error (Train)	6.35%	6.33%	6.33%	6.35%	6.33%	5.9%
l_2 error (Test)	12.09%	11.83%	11.82%	12.74%	13.01%	10.90%

Table 2

The l_2 error for five independently trained neural operator transformer models (Model 1 to Model 5) and their ensemble.

Algorithm 1 Algorithm for solving (5)

Require: Neural operator transformers \mathcal{G}_θ ,
1: Control inputs $C_{1 \leq i \leq N_x}^{t-H:t}, n_p$
Ensure: $m^{(0)}$ \triangleright initial control actions
2: **for** $ite = 0, 1, \dots$ MaxIt **do**
3: obtain future CO₂ predictions \hat{C} (9)(11)
4: evaluate loss $\mathcal{L}(m)$ (10)
5: compute the gradient $\nabla \mathcal{L}(m)$
6: update the control vector with step size η :

$$m^{(ite+1)} \leftarrow m^{(ite)} - \eta \nabla \mathcal{L}(m)$$

7: project the update m to satisfy the box constraints (5c)
8: **end for**
9: **Return** m

ventilation. To ensure that the control vector m remains within feasible bounds (5c), we use a projected gradient descent method. Specifically, after computing the gradient of the objective function with respect to m , the control vector is updated and then clipped to satisfy the predefined bounds. Our optimization procedure is summarized in Algorithm 1.

5. Numerical experiments

In this section, we evaluate the performance of the proposed framework through two experiments: (1) learning experiments to assess the accuracy of the ensemble neural operator, and (2) control experiments to evaluate the effectiveness of our data-driven ventilation control framework. The source code, input data, and trained models from all experiments will be available on GitHub¹.

5.1. Learning results

We train our neural operator transformer to predict CO₂ concentrations on the people surface. The number of query points is $N_x = 7462$, and we select $F = 12$ and $T = 6$, meaning that the model utilizes data from the past 12 time steps (equivalent to 6 minutes) to forecast CO₂ levels over the next 6 time steps (3 minutes). The dataset is divided into an 80% training set and a 20% testing set to evaluate model performance. We train $N_{\text{ensemble}} = 5$ independent models and compute the mean of their predictions as the final output. We use the AdamW optimizer with a cyclical learning rate schedule, and each model is trained for 200 epochs. All experiments are conducted on NVIDIA GeForce RTX 2080 Ti GPUs.

To evaluate model performance, we use the mean l_2 relative error as the primary metric. Let $C^{(d)}, \widehat{C}^{(d)} \in \mathbb{R}^{N_x \times T}$ denote the ground truth and the predicted *mean* future CO₂ concentrations for the d -th sample, respectively, and let D represent the total dataset size. The mean l_2 relative error is defined as:

$$l_2 = \frac{1}{D} \sum_{d=1}^D \frac{\|\widehat{C}^{(d)} - C^{(d)}\|_2}{\|C^{(d)}\|_2}. \quad (12)$$

Figure 4 shows the training convergence and test performance of the five neural operator transformers over 200 epochs. The left plot presents the NLL loss (8), while the right shows the l_2 relative error on both training and test sets. Table 2 lists the l_2 error metrics for the five trained models (Model 1 to Model 5) and the ensemble model, evaluated on both the training and testing datasets. The ensemble model consistently outperforms the individual models, indicating that combining the predictions of multiple models can lead to more robust and accurate results. This improvement in performance can be attributed to the ensemble's ability to mitigate the overfitting tendencies of individual models by averaging out their errors. Notably, the ensemble model achieves the lowest errors, with a training l_2 error of 5.9% and a testing l_2 error of 10.90%.

Figure 5 shows the ground truth CO₂ concentration fields with predictions from the ensemble model and Model 3, along with their relative errors for three test cases. Model 3 was selected as the best-performing single model based on its lowest prediction error. Overall, the neural operator framework shows strong capability in predicting complex spatial CO₂ distributions under varying control parameters. The ensemble model achieves lower relative errors and more accurately captures spatial patterns compared to Model 3. In Section 5.4, we further demonstrate its effectiveness in activating control actions and optimizing the air quality.

The computational efficiency of our neural operator transformer is a significant advantage over traditional CFD simulations. Specifically, the CFD simulation required 1,253.7 seconds to process transient flow across six time steps, while our neural operator transformer completed the same task in just 0.005 seconds. This represents a remarkable speed-up of approximately 250,000 times compared to the CFD approach. Such a dramatic reduction in computational time opens the potential for real-time control of complex ventilation systems, where rapid decision-making is crucial.

5.2. Ventilation control results

Now we evaluate control performance. We examine three scenarios distinguished by the number of occupants n_p ,

¹<https://github.com/alwaysbyx/BuildingControlCFD>

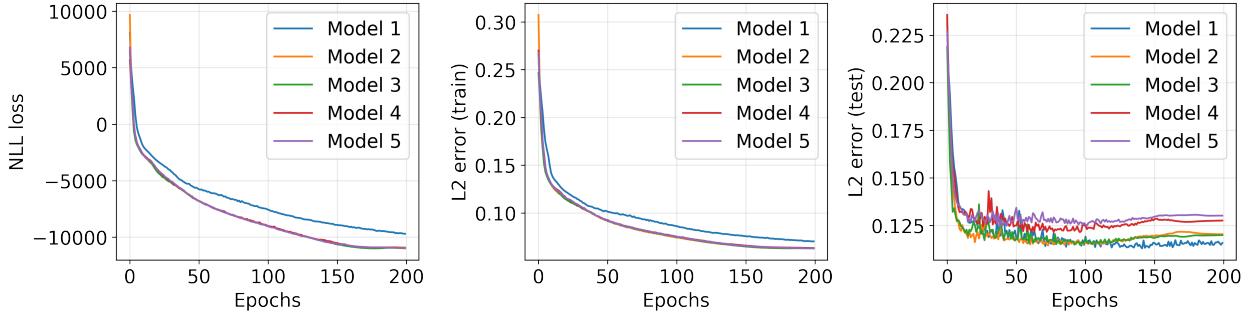


Figure 4: Training loss (NLL loss) during training (left) and the L_2 error for the training (middle) and test sets (right). The L_2 error (12) is computed based on the ground truth and the model's mean prediction output.

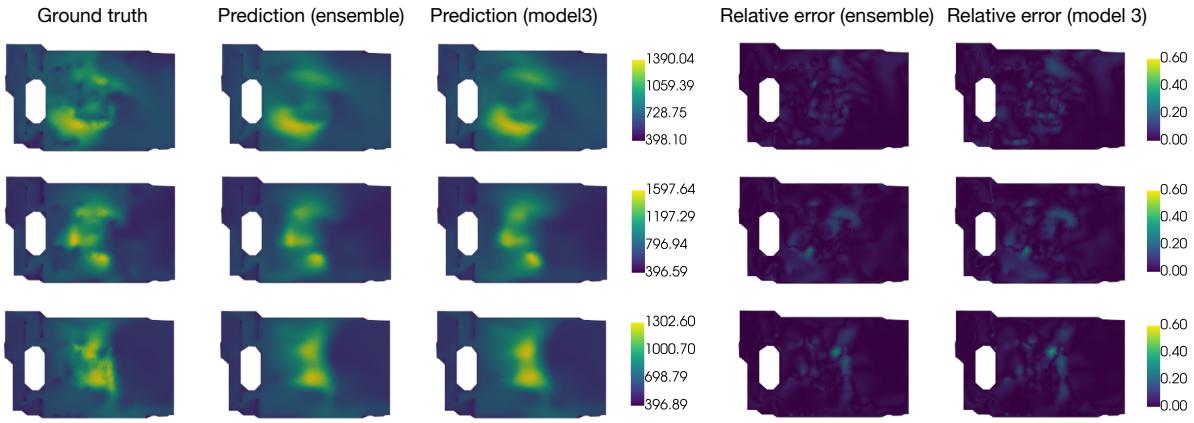


Figure 5: Operator Learning: Visualization of the ground truth, corresponding predictions from the ensemble model and Model 3, and the relative errors between the ground truth and predictions at the final time step.

corresponding to large $n_p = 75$, medium $n_p = 45$, and small values $n_p = 15$. According to ASHRAE Standard 62.1 [34], classrooms typically require 4–6 air changes per hour (ACH) to ensure adequate ventilation and indoor air quality. For all scenarios in this study, the initial control actions are initialized at 5 ACH (within the recommended range) for supply vents, with inlet angles fixed at a 90° downward orientation to align with conventional HVAC configurations. We evaluate the following control strategies:

- **Max Control:** The airflow rate is set to its maximum value, with the inlet angle fixed at 90° downward.
- **Baseline Control:** The control actions determined based on the ASHRAE standard as 5 ACH with 90° downward angle, serving as a baseline for comparison.
- **Rule-based Control:** The airflow rate is set proportionally to the occupancy level, calculated as the number of occupants divided by the maximum occupancy, with 90° downward angle.
- **Data-driven (Average ODE Model):** Following recent practice, we employ a neural network to predict the average CO₂ concentration at the center of the occupancy surface, rather than modeling the full spatial

distribution. Ventilation control actions are then optimized based on the predicted average concentration.

- **Ours:** Our model captures the spatiotemporal dynamics of CO₂ concentrations based on PDE models and CFD simulations. The control strategy is obtained by optimizing over this model to minimize energy consumption while maintaining acceptable air quality.

For our optimization solver, we choose $C_{\text{target}} = 400$ (ppm) and $w_1 = 1, w_2 = 0.4$ and $w_3 = 0.12$ to prioritize air quality improvement while balancing control stability and energy efficiency. We validate the three control strategies using CFD simulations to assess the resulting CO₂ concentrations. Performance is measured using three metrics: mean CO₂ (ppm), peak CO₂ (ppm), and CO₂ violation (%), defined as $\frac{\text{Peak CO}_2 - 1200}{1200}$. As suggested by [35], 1200 ppm is the maximum acceptable CO₂ level for human health. For each metric, we record both the temporal average over the simulation period and the value at the final time step. Energy consumption is modeled as the percentage of the maximum power required to operate the ventilation system at its maximum airflow rate, providing a normalized measure of energy usage relative to the system's peak capacity. This

Case	Control Strategy	Mean CO ₂ (ppm)		Peak CO ₂ (ppm)		CO ₂ > 1200ppm (%)		Energy Consumption (%)
		Average	Final Step	Average	Final Step	Average	Final Step	
1	Max Control	565.6	534.0	895.7	854.2	0.00	0.00	100.0
	Baseline Control	616.1	612.3	1203.0	1143.6	1.32	0.00	50.0
	Rule-based Control	567.4	533.5	898.1	807.9	0.00	0.00	93.8
	Data-driven(Average)	604.4	595.2	1021.6	971.6	0.00	0.00	83.3
	Ours	600.7	587.9	1060.2	1004.0	0.00	0.00	65.8
2	Max Control	546.7	512.2	800.0	741.3	0.00	0.00	100.0
	Baseline Control	599.9	595.9	1097.3	1108.4	0.00	0.00	50.0
	Rule-based Control	593.7	584.6	1057.6	1059.3	0.00	0.00	56.2
	Data-driven(Average)	584.7	562.1	912.7	914.6	0.00	0.00	83.0
	Ours	605.0	600.0	1044.1	881.4	0.00	0.00	43.8
3	Max Control	532.3	497.6	754.0	692.0	0.00	0.00	100.0
	Baseline Control	585.2	578.0	1034.2	1013.9	0.00	0.00	50.0
	Rule-based Control	650.7	712.7	1361.4	1623.1	15.67	35.26	18.8
	Data-driven(Average)	576.9	566.0	946.1	978.3	0.00	0.00	66.1
	Ours	589.4	585.4	984.4	949.8	0.00	0.00	55.2

Table 3

Comparison of CO₂ metrics (mean, peak, and violation percentages) and energy consumption across different control strategies. "Average" refers to the temporal mean over the simulation period, while "Final Step" refers to the CO₂ level at the end of the simulation. Our approach achieves significant ventilation energy savings compared to the other control approaches while maintaining acceptable CO₂ violation levels.

is expressed as:

$$E(m) = \frac{1}{6} \sum_{i=1}^6 m_i^r / \overline{m^r} \times 100\%, \quad (13)$$

where m_i^r represents the ventilation rate for the i -th vent, and $\overline{m^r} = 3.24$ (m/s) is the maximum ventilation rate.

The results are summarized in Table 3. The experimental results demonstrate the effectiveness of the proposed control strategy (Ours) in balancing air quality and energy consumption across all four cases.

Compared to the Max control strategy, our method achieves comparable air quality performance while reducing energy usage by 34–56%. Compared to the baseline control, our approach effectively lowers CO₂ levels and adheres to air quality standards without slight energy increase. Rule-based control suffers from CO₂ violations (e.g., 35.2% final-step violation in Case 3) due to its reliance on well-chosen static operation schedules. Unlike rule-based systems, which require manual fine-tuning of thresholds for different room layouts and occupancy levels, our method autonomously adapts to operation conditions. Thus, compared to rule-based control, our strategy reduces energy consumption by 12–28% in Cases 1–2 while maintaining safe CO₂ levels.

In addition, our method outperforms data-driven control(Average) by reducing energy consumption by at least 10% while maintaining comparable or better peak CO₂ levels. While this average-based method also employs the data-driven model and can optimize ventilation locally, it has two key limitations: (1) it primarily predict average CO₂ concentrations rather than spatial variations, and (2) it cannot effectively map control actions to real-time CO₂ distributions across different zones. Our framework addresses

this limitation by integrating the spatial-temporal CO₂ distribution into control decisions, achieving air quality standards and energy efficiency.

5.3. Multi-step ventilation control results

To further evaluate the effectiveness of the learned operator model in realistic control scenarios, we extend the single-step optimization strategy from Section 5.2 to a longer planning horizon. Specifically, we implement a multi-step ventilation planning procedure over a 30-minute horizon, divided into 10 planning steps. At each planning step, the system optimizes ventilation control for the next 6 time steps (i.e., 3 minutes), resulting in a total of 60 control steps. The first step $t = 0$ is initialized with historical CO₂ concentrations from Scenario 1 in Subsection 5.2, while subsequent steps use predicted CO₂ values obtained by recursively rolling out the learned operator model with previously optimized control actions.

Figure 6 illustrates the performance of the proposed control strategy in comparison to other control methods. The occupancy profile (top) demonstrates distinct variations in human presence. Our strategy dynamically adjusts ventilation in response to occupancy levels. At the beginning, when occupancy is high, the controller increases airflow to mitigate CO₂ accumulation. Although occupancy drops shortly after, the controller maintains a moderately high ventilation rate to clear the residual CO₂ from the earlier crowded period. In contrast, during later low-occupancy intervals, lower accumulated CO₂ allows the system to reduce airflow further, resulting in greater energy savings.

Over the full control period, the total energy consumption for the Max control, Baseline control, Rule-based control, Data-driven control(Average), and our proposed method is 100%, 50%, 63.1%, 55.3%, and 50.6%, respectively. In terms of CO₂ violations, the Max control strategy

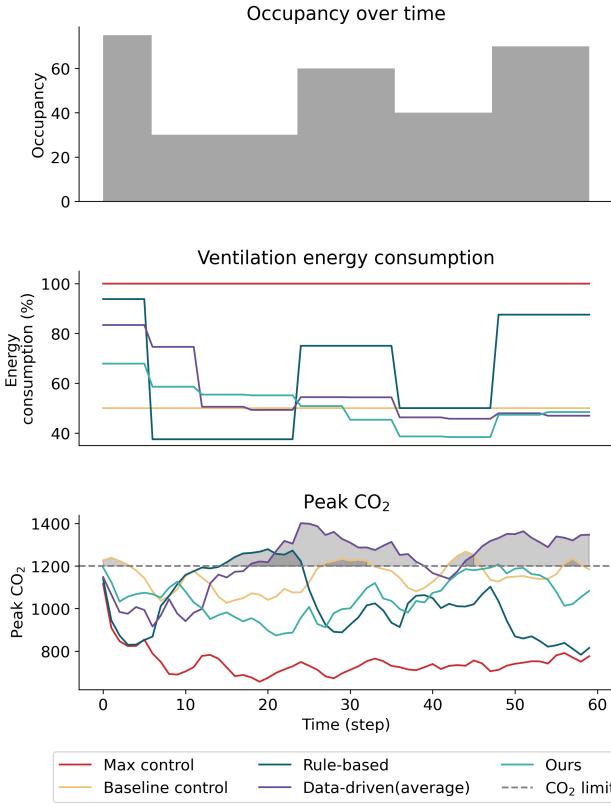


Figure 6: Comparison of occupancy profile, energy consumption, and peak CO₂ concentration over a 60-step period. The top panel illustrates dynamic occupancy changes, the middle panel shows corresponding energy consumption as a percentage of the maximum, and the bottom panel depicts CO₂ concentration levels with the threshold limit (1200 ppm) indicated by a dashed line.

incurs zero violations, while the baseline control results in 17 violations, the rule-based control results in 10, and the Data-driven(Average) control results in 37. Our proposed method incurs only one violation. The relatively poor performance of average-based control stems from its reliance on predicting only the average CO₂ concentration, without capturing the dynamic spatial shifts of peak concentrations over time. As a result, the control decisions fail to address localized high-risk areas, leading to both higher energy consumption and more CO₂ violations.

Notably, our approach achieves the lowest energy consumption (50.6%) among all strategies except the baseline, while dramatically improving air quality by reducing CO₂ violations to near zero.

5.4. Effectiveness of the ensemble model

In this subsection, we demonstrate the effectiveness of the ensemble model in improving air quality compared to individual models, serving as an ablation study. As shown in Section 5.1, the ensemble model achieves the lowest prediction error among all models. Building on these results,

we further illustrate that the ensemble model, which aggregates predictions from multiple independently trained neural operators, consistently outperforms individual models in improving air quality. This is achieved by effectively activating control actions to reduce CO₂ concentrations.

As recommended by the European REHVA guidelines [5], higher airflow rates generally lead to improved air quality. To focus on optimizing air quality, we set the coefficient for energy consumption $w_3 = 0$ in the building control problem (5). This allows the control strategy to prioritize air quality without considering energy cost. We optimize the control actions for a randomly selected case across all models, including both individual models and the ensemble model. The resulting control actions are illustrated in Figure 7. We observe that the ensemble model consistently reaches maximum airflow rates for all inlet vents, which align with the optimal action in this test case as the energy cost coefficient $w_3 = 0$. In contrast, the individual models exhibit variability in their control actions, with some models failing to achieve the maximum airflow rates across all vents.

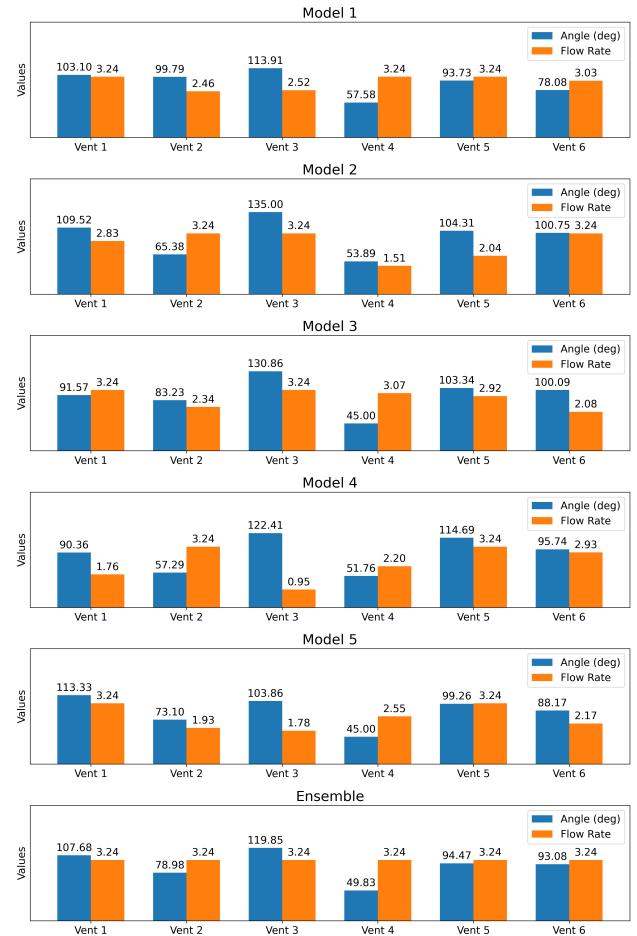


Figure 7: Control actions (flow rates and angles) for individual models (Model 1 to Model 5) and the ensemble model across six inlet vents for one case. The ensemble model consistently achieves maximum airflow rates for all inlet vents.

6. Conclusion and future work

In this work, we propose a novel operator learning framework for energy-efficient ventilation control. Our approach involves training an ensemble neural operator transformer to learn the mapping from past CO₂ fields and control actions to future CO₂ fields. The ensemble model demonstrates superior predictive performance compared to individual models, enabling more robust and reliable control actions. We further integrate the neural operator into a ventilation control optimization framework and employ a gradient-based method to optimize building control actions.

The ensemble neural operator achieves a 10.9% relative error while offering a thousandfold speed-up compared to traditional CFD simulations, enabling real-time building control. Using CFD simulations for validation, we demonstrate that our approach achieves substantial energy savings compared to maximum airflow control, rule-based control, and data-driven methods based on average CO₂ predictions. In addition, compared to baseline control, our method maintains similar energy consumption while significantly reducing CO₂ violations. Further more, we open-source the CFD data to facilitate further research in applying machine learning models to ventilation control in PDE-based settings.

Promising future research directions include extending the framework to larger and more complex building environments, such as multi-zone systems, to evaluate its scalability and generalization capabilities. Additionally, real-world experimentation with integrated sensing and actuation systems is planned to validate the framework's performance.

Acknowledgments

The authors gratefully acknowledge Professor Oliver Schmidt for insightful discussions on computational fluid dynamics (CFD) simulations. This work is supported under Department of Energy grant DE-SC0025495 and Schmidt Sciences AI2050 Early Career Fellowship.

References

- [1] K. J. Chua, S. K. Chou, W. Yang, J. Yan, Achieving better energy-efficient air conditioning—a review of technologies and strategies, *Applied Energy* 104 (2013) 87–104.
- [2] B. Chenari, J. D. Carrilho, M. G. Da Silva, Towards sustainable, energy-efficient and healthy ventilation strategies in buildings: A review, *Renewable and Sustainable Energy Reviews* 59 (2016) 1426–1447.
- [3] A. H. Hosseinkoo, S. Nabi, A. Hosoi, M. A. Dahleh, Data-driven control of covid-19 in buildings: a reinforcement-learning approach, *IEEE Transactions on Automation Science and Engineering* (2023).
- [4] S. Shi, S. Miyata, Y. Akashi, Event-driven model-based optimal demand-controlled ventilation for multizone vav systems: Enhancing energy efficiency and indoor environmental quality, *Applied Energy* 377 (2025) 124683.
- [5] REHVA, Covid 19 guidance, <https://www.rehva.eu/activities/covid-19-guidance>.
- [6] Y. Bian, X. Fu, B. Liu, R. Rachala, R. K. Gupta, Y. Shi, Bear-data: Analysis and applications of an open multizone building dataset, in: *Proceedings of the 10th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation*, 2023, pp. 240–243.
- [7] J. Drgoňa, J. Arroyo, I. C. Figueroa, D. Blum, K. Arendt, D. Kim, E. P. Ollé, J. Oravec, M. Wetter, D. L. Vrabie, et al., All you need to know about model predictive control for buildings, *Annual Reviews in Control* 50 (2020) 190–232.
- [8] Y. Chen, Y. Shi, B. Zhang, Optimal control via neural networks: A convex approach, in: *International Conference on Learning Representations*, 2018.
- [9] Y. Gao, S. Miyata, Y. Akashi, Energy saving and indoor temperature control for an office building using tube-based robust model predictive control, *Applied Energy* 341 (2023) 121106.
- [10] W. Su, Z. Ai, J. Liu, B. Yang, F. Wang, Maintaining an acceptable indoor air quality of spaces by intentional natural ventilation or intermittent mechanical ventilation with minimum energy use, *Applied Energy* 348 (2023) 121504.
- [11] Y. Du, H. Zandi, O. Kotevska, K. Kurte, J. Munk, K. Amasyali, E. McKee, F. Li, Intelligent multi-zone residential hvac control strategy based on deep reinforcement learning, *Applied Energy* 281 (2021) 116117.
- [12] T. Yang, L. Zhao, W. Li, J. Wu, A. Y. Zomaya, Towards healthy and cost-effective indoor environment management in smart homes: A deep reinforcement learning approach, *Applied Energy* 300 (2021) 117335.
- [13] H. Wang, X. Chen, N. Vital, E. Duffy, A. Razi, Energy optimization for hvac systems in multi-vav open offices: A deep reinforcement learning approach, *Applied Energy* 356 (2024) 122354.
- [14] W. Shang, J. Liu, C. Wang, J. Li, X. Dai, Developing smart air purifier control strategies for better iaq and energy efficiency using reinforcement learning, *Building and Environment* 242 (2023) 110556.
- [15] Y. Bian, X. Fu, R. K. Gupta, Y. Shi, Ventilation and temperature control for energy-efficient and healthy buildings: A differentiable pde approach, *Applied Energy* 372 (2024) 123477.
- [16] M. Ning, S. Mengjie, C. Mingyin, P. Dongmei, D. Shiming, Computational fluid dynamics (cfd) modelling of air flow field, mean age of air and co₂ distributions inside a bedroom with different heights of conditioned air supply outlet, *Applied Energy* 164 (2016) 906–915.
- [17] M. Abuhegazy, K. Talaat, O. Anderoglu, S. V. Poroseva, Numerical investigation of aerosol transport in a classroom with relevance to covid-19, *Physics of Fluids* 32 (10) (2020).
- [18] M. M. A. Mahmoud, P. Bahl, A. d. A. Aquino, C. MacIntyre, S. Bhattacharjee, D. Green, N. Cooper, C. Doolan, C. de Silva, A numerical framework for the analysis of indoor air quality in a classroom, *Journal of Building Engineering* (2024) 109659.
- [19] C. Michoski, M. Milosavljević, T. Oliver, D. R. Hatch, Solving differential equations using deep neural networks, *Neurocomputing* 399 (2020) 193–212.
- [20] B. Li, B. Wu, Y. Peng, W. Cai, Tube-based robust model predictive control of multi-zone demand-controlled ventilation systems for energy saving and indoor air quality, *Applied Energy* 307 (2022) 118297.
- [21] C. Li, C. Cui, M. Li, A proactive 2-stage indoor co₂-based demand-controlled ventilation method considering control performance and energy efficiency, *Applied Energy* 329 (2023) 120288.
- [22] L. Lu, P. Jin, G. Pang, Z. Zhang, G. E. Karniadakis, Learning non-linear operators via deepnet based on the universal approximation theorem of operators, *Nature machine intelligence* 3 (3) (2021) 218–229.
- [23] Z. Li, N. B. Kovachki, K. Azizzadenesheli, K. Bhattacharya, A. Stuart, A. Anandkumar, et al., Fourier neural operator for parametric partial differential equations, in: *International Conference on Learning Representations*.
- [24] Z. Hao, Z. Wang, H. Su, C. Ying, Y. Dong, S. Liu, Z. Cheng, J. Song, J. Zhu, Gnot: A general neural operator transformer for operator learning, in: *International Conference on Machine Learning*, PMLR, 2023, pp. 12556–12569.
- [25] Z. Li, N. Kovachki, C. Choy, B. Li, J. Kossaifi, S. Otta, M. A. Nabian, M. Stadler, C. Hundt, K. Azizzadenesheli, et al., Geometry-informed neural operator for large-scale 3d pdes, *Advances in Neural Information Processing Systems* 36 (2024).

- [26] Y. Wang, Z. Li, Z. Yuan, W. Peng, T. Liu, J. Wang, Prediction of turbulent channel flow using fourier neural operator-based machine-learning strategy, *Physical Review Fluids* 9 (8) (2024) 084604.
- [27] H. Gao, W. Qian, J. Dong, J. Liu, Rapid prediction of indoor airflow field using operator neural network with small dataset, *Building and Environment* 251 (2024) 111175.
- [28] A. Bulińska, Z. Buliński, A cfd analysis of different human breathing models and its influence on spatial distribution of indoor air parameters, *Computer Assisted Methods in Engineering and Science* 22 (3) (2017) 213–227.
- [29] Y. He, Y. Chu, H. Zang, J. Zhao, Y. Song, Experimental and cfd study of ventilation performance enhanced by roof window and mechanical ventilation system with different design strategies, *Building and Environment* 224 (2022) 109566.
- [30] X. Sha, Z. Ma, S. Sethuvenkatraman, W. Li, Online learning-enhanced data-driven model predictive control for optimizing hvac energy consumption, indoor air quality and thermal comfort, *Applied Energy* 383 (2025) 125341.
- [31] S. Zhang, Z. Ai, Z. Lin, Novel demand-controlled optimization of constant-air-volume mechanical ventilation for indoor air quality, durability and energy saving, *Applied Energy* 293 (2021) 116954.
- [32] ASHRAE Standard 55—thermal environmental conditions for human occupancy, Tech. rep., ASHRAE Inc., Atlanta, GA (1992).
- [33] B. Lakshminarayanan, A. Pritzel, C. Blundell, Simple and scalable predictive uncertainty estimation using deep ensembles, *Advances in neural information processing systems* 30 (2017).
- [34] ASHRAE Standard 62.1—ventilation for acceptable indoor air quality, Tech. rep., ASHRAE Inc., Atlanta, GA (2010).
- [35] W. Zhang, W. Wu, L. Norford, N. Li, A. Malkawi, Model predictive control of short-term winter natural ventilation in a smart building using machine learning algorithms, *Journal of Building Engineering* 73 (2023) 106602.

Appendix: Airflow dynamics modeling

Indoor air is modeled as an incompressible fluid consisting of four species, namely: oxygen (O_2), carbon dioxide (CO_2), water steam (H_2O) and nitrogen (N_2) [28]. The numerical model is based on the Navier-Stokes equations for incompressible flow, incorporating continuity, momentum, energy conservation, and turbulence model equations, along with species transport equations.

The continuity equation enforces the incompressibility of air,

$$\nabla \cdot \mathbf{u} = 0, \quad (14)$$

where \mathbf{u} refers to fluid velocity vector. The conservation equations for air constituents govern the transport and distribution of individual gas species within the airflow.

$$\frac{\partial Y_i}{\partial t} + \nabla \cdot (Y_i \mathbf{u}) = -\nabla \cdot \frac{1}{\rho} \mathbf{j}_i, \quad (15)$$

where i denotes three air constituents, namely, O_2 , CO_2 and H_2O , Y_i is the mass fraction of the i -th air constituent, and ρ is density of fluid. The mass flux of the i -th constituent can be calculated:

$$\mathbf{j}_i = -D_{eff} \nabla Y_i, \quad (16)$$

where D_{eff} is the effective diffusion coefficient which includes turbulence effects. The mass fraction of N_2 was calculated from the sum of mass fractions of all air species which should be equal to unity. CO_2 concentration can be converted from mass fraction of CO_2 with

$$C(\mathbf{x}, t) = Y_{CO_2}(\mathbf{x}, t) \cdot 10^6 \cdot \frac{\text{molecular weight of } CO_2}{\text{molecular weight of air}}, \quad (17)$$

where molecular weight of $CO_2 = 44.01\text{g/mol}$ and molecular weight of air is 28.97g/mol . The momentum equation describes the motion of air as an incompressible fluid, governed by the Navier-Stokes equations:

$$\frac{\partial(\rho \mathbf{u})}{\partial t} + \nabla \cdot (\rho \mathbf{u} \mathbf{u}) = -\nabla p + \rho \mathbf{g} + \nabla \cdot (\mu \nabla \mathbf{u}) - \nabla \cdot \tau_t, \quad (18)$$

where p is the pressure field, \mathbf{g} is a vector of gravitational acceleration, μ is a molecular dynamic viscosity and τ_t is a turbulence tensor. The energy conservation equation governs the transport of thermal energy within the airflow,

$$\frac{\partial(pe)}{\partial t} + \nabla \cdot (pe \mathbf{u}) = \nabla \cdot (k_{eff} \nabla T) - \nabla \left(\sum_i h_i j_i \right), \quad (19)$$

where e is a specific internal energy, k_{eff} is an effective heat conductivity, T is fluid temperature and h_i refers to a specific enthalpy of fluid. The turbulence model equations approximate the effects of small-scale turbulent eddies,

$$\tau_{t,ij} = \mu_t \left(\frac{\partial \mathbf{u}_i}{\partial x_j} + \frac{\partial \mathbf{u}_j}{\partial x_i} \right) - \frac{2}{3} \rho \kappa \delta_{ij}, \quad (20)$$

where μ_t is a turbulent viscosity, κ is a turbulent kinetic energy and δ_{ij} is Kronecker's delta.

The domain boundary ∂Z encompasses all surfaces, including walls, ventilation interfaces, and occupant boundaries. We further define Ω_{supply} as the inlet vent boundary, Ω_{occupant} as the occupant surface boundary, Ω_{return} as the outlet vent boundary. The airflow boundary conditions are then given by:

$$\mathbf{u}(\mathbf{x}, t) = m_i^r(t) \begin{bmatrix} \sin(m_i^a(t)) \\ 0 \\ -\cos(m_i^a(t)) \end{bmatrix}, \forall \mathbf{x} \in \Omega_{\text{supply},i}, \quad (21a)$$

$$\mathbf{u}(\mathbf{x}, t) = \frac{v_{\text{occupant}}}{A_{\text{occupant}}} \cdot N_{\text{occupant}} \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \forall \mathbf{x} \in \Omega_{\text{occupant}}, \quad (21b)$$

$$\mathbf{n} \cdot \nabla \mathbf{u} = 0, \forall z \in \mathcal{Z}_{\text{return}}, \quad (21c)$$

$$\mathbf{u}(\mathbf{x}, t) = 0, \forall \mathbf{x} \in \partial \Omega \setminus (\Omega_{\text{supply}} \cup \Omega_{\text{occupant}} \cup \Omega_{\text{return}}). \quad (21d)$$

Constraint (21a) specifies that the airflow velocity at i -th group of supply vents is determined by its corresponding airflow rate $m_i^r(t)$ and direction angle $m_i^a(t)$. Constraint (21b) relates the airflow rate to the number of occupants, where the exhaled air rate v_{occupant} is set to 6 L/min per person [29], and A_{occupant} denotes the occupant boundary area. Constraint (21c) sets the Neumann boundary conditions at the return vent and constraint (21d) applies Dirichlet conditions to all other boundaries by setting the airflow velocity as zero [15].