



NVAPF: An adaptive particle filter algorithm for CO₂-based natural ventilation rate estimation with high temporal resolution and stability

Sen Miao ^{*} , Marta Gangolells, Blanca Tejedor

Department of Project and Construction Engineering, Group of Construction Research and Innovation (GRIC), Universitat Politècnica de Catalunya, C/ Colom, 11, Ed. TR5, 08222 Terrassa (Barcelona), Spain



ARTICLE INFO

Keywords:

Natural ventilation rate
CO₂ tracer gas
Bayesian filter
Adaptive particle filter
Educational buildings

ABSTRACT

CO₂-based ventilation rate estimation techniques have been widely used in relevant studies in naturally ventilated educational buildings. Such techniques are non-invasive, low-cost, simple, accurate, and do not interfere with the activities of indoor occupants. However, the estimation is significantly affected by the CO₂ measurement noise, the uncertainties associated with CO₂ generation rate, and complex natural ventilation dynamics. Existing techniques were found to have limited capabilities to deal with these aspects. Therefore, this research proposed an adaptive particle filter algorithm for CO₂-based natural ventilation rate estimation and validated it through a case study in an educational building. Compared with the existing transient mass balance model and the extended Kalman filter technique, the estimation stability has been improved by nearly 6 times and 3 times, respectively. More importantly, the proposed algorithm is significantly more robust to abrupt changes in indoor CO₂ and can effectively avoid large drifts in the estimated ventilation rate due to sudden window opening and sudden changes in room occupancy, demonstrating great practical applicability for real-time estimation with a high temporal resolution of 1 minute. To help relevant users with practical applications, the study also analyzed the algorithm parameter settings and the impact of simplification strategies commonly used in relevant studies, such as the use of a fixed outdoor CO₂ concentration, an averaged CO₂ generation rate, and an assumed constant room occupancy. Finally, considering that applying the proposed algorithm requires programming skills, an open, user-friendly software has been developed for relevant users for a convenient implementation.

1. Introduction

Indoor air quality is one of the most important aspects of indoor environmental quality in educational buildings. It has become a major concern for the public, government, and scientific community, since the outbreak of the COVID-19 pandemic [1,2]. Relevant building standards such as EN16798-1 [3] and ASHRAE 62.1 [4] specify the minimum ventilation rate requirements for maintaining good indoor air quality in classrooms, with an air exchange rate of 4 times/hour or a fresh air supply rate of 10 L/s per person. Since the COVID pandemic, the Harvard T.H. Chan School of Public Health has recommended increasing this threshold to 6 times/hour or 15 L/s per person, in order to prevent massive infection of students and to ensure a safe learning environment for them [5,6].

Natural ventilation is often the primary means of maintaining good indoor air quality, as most schools are not equipped with mechanical ventilation systems [7]. Unlike mechanical ventilation systems, which can operate at a controlled and stable ventilation rate, natural ventilation is determined by both buoyancy and wind effects, and is influenced by relevant building characteristics such as facade orientation and window geometry [8,9]. Therefore, the actual natural ventilation rates often need to be estimated through field measurements.

The occupant-released CO₂ tracer gas approach is a technique of natural ventilation rate estimation [10,11]. This approach uses CO₂ as a tracer gas, estimating the airflow rate based on the transient mass balance equation according to the change in the indoor CO₂ concentration over a period of time [12]. It has well-recognized advantages: non-invasive, low-cost, simple but accurate. More importantly, this

List of abbreviations: APF, Adaptive particle filter; C_{in}, Indoor CO₂ concentration; C_{out}, Outdoor CO₂ concentration; EKF, Extended Kalman filter; G, CO₂ generation rate; HVAC, Heating, ventilation, and air conditioning; IOT, Internet of Things; MPE, Mean percentage error; N, Number of occupants; Q_{vent}, Ventilation rate; V, Room volume; TMBM, Transient mass balance model.

* Corresponding author.

E-mail address: sen.miao@upc.edu (S. Miao).

method does not interfere with the activities of the indoor occupants compared to other methods such as direct measurement using air flow meters or using other tracer gases such as SF6 [10,11]. As a result, it has become the most important method for estimating natural ventilation rates in educational buildings and has been widely applied in relevant field studies in schools [7,13–17].

However, there are certain limitations to this approach. On one hand, the estimation of the ventilation rate can be significantly affected by the precision of the CO₂ measurement device. The measurement error of the CO₂ sensor is typically ± 50 ppm. Such measurement noise can significantly affect the reliability of the calculation results if it is not properly addressed [10,18,19]. A commonly used approach to reduce the measurement noise is to use average values of CO₂ observations over a period of time like 10 to 30 min [15–17]. However, this is certainly not a proper solution when a high temporal resolution (e.g., 1 min) or a real-time estimation of the ventilation rate is required. On the other hand, the occupant-released CO₂ tracer gas technique takes into account the amount of generated CO₂ in the room. The CO₂ generation is determined by the individual CO₂ generation rate, which has uncertainties due to individual differences and activities of occupants. The transient mass balance method uses a fixed CO₂ generation rate in the equation, which does not account for the uncertainties associated with CO₂ generation.

Bayesian filter (recursive Bayesian estimation) is a very useful technique to address the above limitations [20]. It can effectively deal with the measurement noise of CO₂ to achieve reliable estimation of ventilation rates while considering the uncertainty of CO₂ generation. As important members of the Bayesian filter family, the Kalman filter and extended Kalman filter have been applied to ventilation rate estimation [18,21]. However, the Kalman filter is based on the linear system, which is more suitable for steady-state estimation (stable, constant ventilation rates) rather than dynamic natural ventilation rates. The extended Kalman filter approximates a nonlinear system based on the Kalman filter, but it is very sensitive to abrupt changes in the system, such as an abrupt increase in ventilation rate due to window openings or sudden changes in occupancy. In fact, as another important member of the Bayesian filter family, the particle filter is a technique specifically designed for nonlinear systems with well-recognized robustness. This technique is more suitable for estimating the natural ventilation rate, especially the complex dynamics related to window operation and occupancy in reality. However, to the best of the author's knowledge, till now, there is no research that has explored the development of such particle filters for natural ventilation rate estimation.

Therefore, this research proposes an adaptive particle filter algorithm for estimating ventilation rates in naturally ventilated buildings, which has better practical applicability than the existing techniques. To help relevant users in their practical application, this research also tested several common simplification strategies used in relevant studies, such as using a fixed outdoor CO₂ concentration, an averaged CO₂ generation rate, and an assumed constant room occupancy. Finally, considering that applying the proposed algorithm requires certain programming skills, an open, user-friendly software has been developed to help relevant researchers and engineers for a convenient implementation. The results of this research certainly contribute to both scientific knowledge and practical applications on this research topic.

Following this introduction, Section 2 describes the methodology, Section 3 analyzes and discusses the results, and Section 4 summarizes the conclusions and recommendations.

2. Methodology

This section introduces the methodology of this research. Section 2.1 briefly introduces the occupant-released CO₂ tracer gas approach for ventilation rate estimation. Section 2.2 elaborates on the existing ventilation rate estimation techniques, including the transient mass balance model and the extended Kalman filter technique. Section 2.3

proposes a particle filter algorithm for ventilation rate estimation. Section 2.4 describes the case study conducted to validate the proposed algorithm.

2.1. Occupant-released CO₂ tracer gas approach

The tracer gas method is widely used to estimate ventilation rates. Unlike the traditional tracer gas methods, the occupant-released CO₂ tracer gas approach uses CO₂ generated by occupants. As mentioned earlier, this approach avoids the external gas injection and extensive equipment setups that interfere with occupants' activities [10,11]. The airflow rate in a given space can be estimated based on the changes in indoor CO₂ concentrations over time (Fig. 1).

The CO₂ tracer gas approach assumes that CO₂ is uniformly distributed in the space. Ventilation promotes the mixing of indoor air. The CO₂ measurement sensor should be placed within the breathing zone of the occupants, at a height of 0.75–1.80 m above the ground, and avoid direct contact with CO₂ sources and ventilated fresh airflow [12,22]. Related experimental research indicated that for spaces with a floor area of less than 100 m², one sensor placed at the room center is sufficient. While multi-point measurement with three or more sensors is recommended for rooms larger than 200 m² [23].

2.2. Existing CO₂-based ventilation rate estimation techniques

This section introduces existing CO₂-based natural ventilation rate estimation techniques, including the conventional method of transient mass balance model (Section 2.2.1) and the extended Kalman filter (Section 2.2.2).

2.2.1. Transient mass balance model

In a given room of a volume V , the indoor CO₂ concentration depends on indoor CO₂ generation, ventilation rate, outdoor CO₂ concentration, satisfying the mass balance equation [10]:

$$\frac{dC_{in}}{dt} \cdot V = (C_{out} - C_{in}) \cdot Q_{vent} + N \cdot G \quad (1)$$

With a small time interval ($dt=\Delta t=1$ min), the transient mass balance model (TMBM) can be expressed as:

$$C_{in,t} = C_{in,t-1} + \left(\frac{N_t \cdot G}{V} - \frac{Q_{vent,t}}{V} (C_{in,t-1} - C_{out,t}) \right) \cdot \Delta t \quad (2)$$

Where $C_{in,t}$ denotes indoor CO₂ concentration at time step t (ppm), $C_{in,t-1}$ is the indoor CO₂ concentration at previous time step $t-1$ (ppm), C_{out} is the outdoor CO₂ concentration (ppm), V is the room volume (m³), N_t is the number of occupants in the room at time step t , G is the individual CO₂ generation rate (m³/min/person), and Q_{vent} is the ventilation

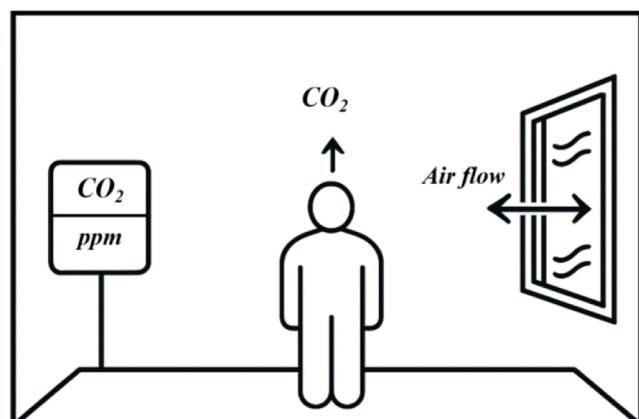


Fig. 1. Schematic diagram of the occupant-released CO₂ tracer gas approach.

rate at time step t (m^3/min).

When the indoor and outdoor CO_2 concentrations, the number of occupants, and the individual CO_2 generation rate at each time step are known, the corresponding ventilation rate can be calculated by a numeric solver.

2.2.2. Extended kalman filter

The extended Kalman filter (EKF) is a nonlinear extension of the well-known Kalman filter [24–26]. Compared to the conventional mass balance equation, this technique can address noise in CO_2 measurements while accounting for uncertainties associated with ventilation rate and CO_2 generation rate. It works on a defined state-space model, estimates (predicts) the state of the system at the next time step, and incorporates information from new measurements to correct the state estimate.

To be more specific, for a given room of a volume V , the indoor CO_2 concentration follows the mass balance equation with a small time interval ($\Delta t=1 \text{ min}$). The system follows the dynamics:

$$C_{in,t+1} = C_{in,t} + \frac{N_t \cdot G_t}{V} - \frac{Q_{vent,t}}{V} (C_{in,t} - C_{out,t}) \quad (3)$$

$$Q_{vent,t+1} = Q_{vent,t} + w_{Qvent} \quad (4)$$

$$G_{t+1} = G_t + w_G \quad (5)$$

The difference with the traditional mass balance model is that the ventilation rate Q_{vent} and the CO_2 generation rate G are added with the process noises w_{Qvent} and w_G to characterize the uncertainties associated with their changes over time.

Hence, the state vector of the system at time step t can be expressed as:

$$X_t = \begin{bmatrix} C_{in,t} \\ Q_{vent,t} \\ G_t \end{bmatrix} \quad (6)$$

Where $C_{in,t}$ is the indoor CO_2 observation, while $Q_{vent,t}$ and G_t are hidden states (latent variables) to be estimated.

Hence, the state of the system at the next time step $t+1$ can be expressed as:

$$X_{t+1} = f(X_t) + w_t, \quad w_t \sim N(0, Q) \quad (7)$$

Where $f(X_t)$ is the state transition function:

$$f(X_t) = \begin{bmatrix} C_{in,t} + \frac{N_t G_t}{V} - \frac{Q_{vent,t}}{V} (C_{in,t} - C_{out,t}) \\ Q_{vent,t} \\ G_t \end{bmatrix} \quad (8)$$

Q is the process noise covariance matrix:

$$Q = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \sigma_{Qvent}^2 & 0 \\ 0 & 0 & \sigma_G^2 \end{bmatrix} \quad (9)$$

Since the system is nonlinear, EKF uses a first-order Taylor expansion to approximate state transition function around the previous state estimate \hat{X}_t . Hence:

$$X_{t+1} \approx f(\hat{X}_t) + F_t(X_t - \hat{X}_t) + w_t \quad (10)$$

Where F_t is the Jacobian matrix:

$$F_t = \begin{bmatrix} 1 - \frac{Q_{vent,t}}{V} & -\frac{(C_{in,t} - C_{out,t})}{V} & \frac{N_t}{V} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (11)$$

The predicted state of the system is:

$$\hat{X}_{t+1|t} = f(\hat{X}_t) \quad (12)$$

The predicted covariance matrix is:

$$P_{t+1|t} = F_t P_t F_t^T + Q_t \quad (13)$$

Where P_t is the state covariance matrix characterizing the uncertainties between the states of CO_2 , Q_{vent} and G , while Q_t is the process noise covariance.

EKF updates the predicted state based on actual observations. The observation model of the system can be expressed as:

$$Z_t = H X_t + v_t, \quad v_t \sim N(0, R) \quad (14)$$

Where Z_t is the observation of the system, H is the observation matrix, and v_t is the measurement noise.

Since CO_2 is the only observed state of the system, the observation matrix $H = [1, 0, 0]$, $R = \sigma_{sensor}^2$, which is the variance of CO_2 sensor measurement error σ_{sensor} . Hence the system observation is:

$$Z_t = C_{observation,t} = C_{in,t} + v_t \quad (15)$$

The difference between the observed and predicted CO_2 concentration is called innovation in EKF, which can be expressed as:

$$Y_t = C_{observation,t} - C_{predicted,t} \quad (16)$$

$$C_{predicted,t} = C_{in,t|t-1} \quad (17)$$

The innovation covariance is:

$$S_t = H P_{t|t-1} H^T + R \quad (18)$$

Like a standard Kalman filter, EKF uses Kalman Gain to calibrate the estimate through observations, which can be expressed as:

$$K_t = P_{t|t-1} H^T S_t^{-1} \quad (19)$$

Hence, the state estimate can be updated:

$$X_{t|t} = X_{t|t-1} + K_t Y_t \quad (20)$$

The covariance matrix is also updated:

$$P_{t|t} = (I - K_t H) P_{t|t-1} (I - K_t H)^T + K_t R K_t^T \quad (21)$$

Where I is the identity matrix.

The above briefly describes the basic principle of EKF. For more detailed instructions on this technique, Labbe [26] is a very useful reference. Duarte et al. [18] first constructed and implemented an EKF for ventilation rate estimation in MATLAB, while the EKF built in this research is generally the same.

It is worth mentioning that to initialize the EKF, it is necessary to assume the initial values for the system states (CO_2 , Q_{vent} , G). In practice, the assumed initial values only affect the first few estimates. After several rounds of iterations, the EKF can automatically calibrate and track the state of the system.

2.3. Proposal of an adaptive particle filter algorithm

Both particle filter and extended Kalman filter (EKF) are Bayesian filters for state estimation of dynamic systems, but they are fundamentally different in the way of handling uncertainty and non-linearity [20, 27]. The EKF is essentially an extension of the linear Kalman filter to approximate nonlinear systems, which limits its performance in practice. In a linear system, the ventilation rate is assumed to be constant. In reality, the natural ventilation rate fluctuates and can change abruptly due to the window operation. The particle filter is naturally designed for nonlinear systems, with better flexibility, reliability and robustness, and is more suitable for natural ventilation rate estimation.

To be more specific, in a room of volume V , the indoor CO_2 concentration still follows the mass balance equation with a small time

interval ($\Delta t=1$ min). The system follows the dynamics:

$$C_{in,t} = C_{in,t-1} + \frac{N_t \cdot G_t}{V} - \frac{Q_{vent,t}}{V} (C_{in,t-1} - C_{out,t}) + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma_{sensor}^2) \quad (22)$$

Where ε_t represents the measurement noise.

The particle filter uses a set of weighted particles N_p (e.g., $N_p=1000$, 1000 particles) characterizing the probability distribution of the unknown state of ventilation rate Q_{vent} and CO₂ generation rate G :

$$Q_{vent,t} = Q_{vent,t-1} + w_{Q_{vent}}, \quad w_{Q_{vent}} \sim N(0, \sigma_{Q_{vent}}^2) \quad (23)$$

$$G_t = G_{t-1} + w_G, \quad w_G \sim N(0, \sigma_G^2) \quad (24)$$

Where $w_{Q_{vent}}$ and w_G are the process noises.

In reality, the CO₂ generation rate remains relatively stable and does not suddenly increase or decrease exponentially. However, the ventilation rate can change drastically due to window operation. Therefore, a more appropriate approach is to set the process noise of the ventilation rate $w_{Q_{vent}}$ in an adaptive manner as follows:

$$\sigma_{Q_{vent,adaptive}}^2 = \sigma_{Q_{vent}}^2 \cdot \left(1 + \frac{(Q_{vent,t+1} - Q_{vent,t})}{Q_{vent,t}} \right) \quad (25)$$

This means that the process noise is scaled based on a defined base value, according to the relative change in the predicted ventilation rate Q_{vent} in adjacent time steps. This form of process noise is more flexible and ideal for the natural ventilation scenario with high dynamics and uncertainty. When the ventilation rate is relatively stable, its relative change is low, and the process noise remains near the defined value. If the ventilation rate suddenly increases or decreases, the relative change becomes significantly larger, and the process noise can be temporarily scaled to help the filter effectively track the changes.

Hence, ventilation rate Q_{vent} follows:

$$Q_{vent,t} = Q_{vent,t-1} + w_{Q_{vent,adaptive}}, \quad w_{Q_{vent,adaptive}} \sim N(0, \sigma_{Q_{vent,adaptive}}^2) \quad (26)$$

Such a particle filter that uses adaptive process noise settings is also called an adaptive particle filter (APF).

Then, the APF computes the weight of each particle using the likelihood of observing the predicted state. For each particle i , the predicted state of CO₂ is:

$$C_{in,t}^i = C_{in,t-1} + \frac{N_t G^i}{V} - \frac{Q_{vent}^i}{V} (C_{in,t-1} - C_{out,t}) \quad (27)$$

The likelihood of observing the measured $C_{in,t}$ given the predicted $C_{in,t}^i$ can be calculated through a log-likelihood function:

$$\log P(C_{in,t} | Q_{vent}^i, G^i) = -\frac{(C_{in,t} - C_{in,t}^i)^2}{2\sigma_{sensor}^2} - \log(\sqrt{2\pi\sigma_{sensor}^2}) \quad (28)$$

The particle weight is updated as:

$$w_t^i = \frac{\exp(\log P(C_{in,t} | Q_{vent}^i, G^i))}{\sum_{j=1}^{N_p} \exp(\log P(C_{in,t} | Q_{vent}^j, G^j))} \quad (29)$$

Hence, the final estimate of ventilation rate \hat{Q}_{vent} at time step t is computed as the weighted mean of the particle set:

$$\hat{Q}_{vent,t} = \sum_{i=1}^{N_p} w_t^i Q_{vent}^i \quad (30)$$

The CO₂ generation rate at time step t can also be estimated if necessary by:

$$\hat{G}_t = \sum_{i=1}^{N_p} w_t^i G^i \quad (31)$$

In this process, the weights of most particles gradually become

negligible over time. This causes a few high-weight particles to dominate in estimates, and the filter does not effectively approximate the true distribution. Therefore, the APF is designed to incorporate stratified resampling at each time step, in order to prevent particle degeneracy and maintain particle diversity.

Simply, the stratified resampling generates stratified sample positions:

$$u_i = \frac{i-1+U_i}{N_p}, U_i \sim U(0, 1) \quad (32)$$

Then it calculates the cumulative sum of particle weights C_j , and iterates through u_i to find the smallest where $u_i < C_j$:

$$C_j = \sum_{k=1}^j w_t^k \quad (33)$$

The corresponding particle index j then is then assigned to the new particle i , and all weights are then reset to be uniform:

$$Q_{vent,t}^i = Q_{vent,t}^j \quad (34)$$

$$G_t^i = G_t^j \quad (35)$$

$$w_t^i = \frac{1}{N_p}, \forall i \quad (36)$$

Furthermore, as the particle filter needs to be initialized like the EKF, a data-driven initialization strategy can be designed for the APF. Instead of randomly assuming an initial ventilation rate $Q_{vent,0}$, the average ventilation rate $\bar{Q}_{vent,m}$ calculated from the first m CO₂ observations is used as the initial value, and the APF initializes particles around this initial ventilation rate:

$$Q_{vent,0}^i \sim N(\bar{Q}_{vent,m}, \sigma_{Q_{vent,adaptive}}^2) \quad (37)$$

This data-driven initialization strategy is more flexible and allows the filter to track the system state more quickly.

The above briefly describes the fundamental principle of the proposed APF. [Appendix A](#) presents a flowchart of the algorithm.

2.4. Validation of the proposed adaptive particle filter with case study

A case study was conducted to validate the performance of the proposed APF and to compare it with the existing techniques. Subsequently, a sensitivity analysis was performed to investigate and analyze the impact of the algorithm parameter settings and simplification strategies commonly used in relevant field studies, including the use of a fixed outdoor CO₂ concentration, an averaged CO₂ generation rate, and a constant room occupancy.

2.4.1. Case study description

The case study was conducted between November and December 2024, in two classrooms of the TR1 building of the Terrassa campus of the Universitat Politècnica de Catalunya (UPC). These two classrooms are completely naturally ventilated, while their photos and architectural characteristics are shown in [Fig. 2](#). In the case study, a total of 6 field tests were conducted during the class held in these two rooms, with 3 tests in each room and each test lasting approximately 3 hours. The occupants of the rooms were mainly college students with an average age of 22. Notably, as the class was in the engineering school, the gender ratio of the students was disproportionate, with approximately 80 % of male students and 20 % of female students.

In the field measurements, two Comet U3430 sensors were used to measure indoor and outdoor CO₂ concentrations, respectively. Following the specifications of ASTM D6245–18 [12], one sensor was placed in the center of the classroom at a height of 1.1 m and 2 m away from disturbances, measuring the indoor CO₂ concentration. The other

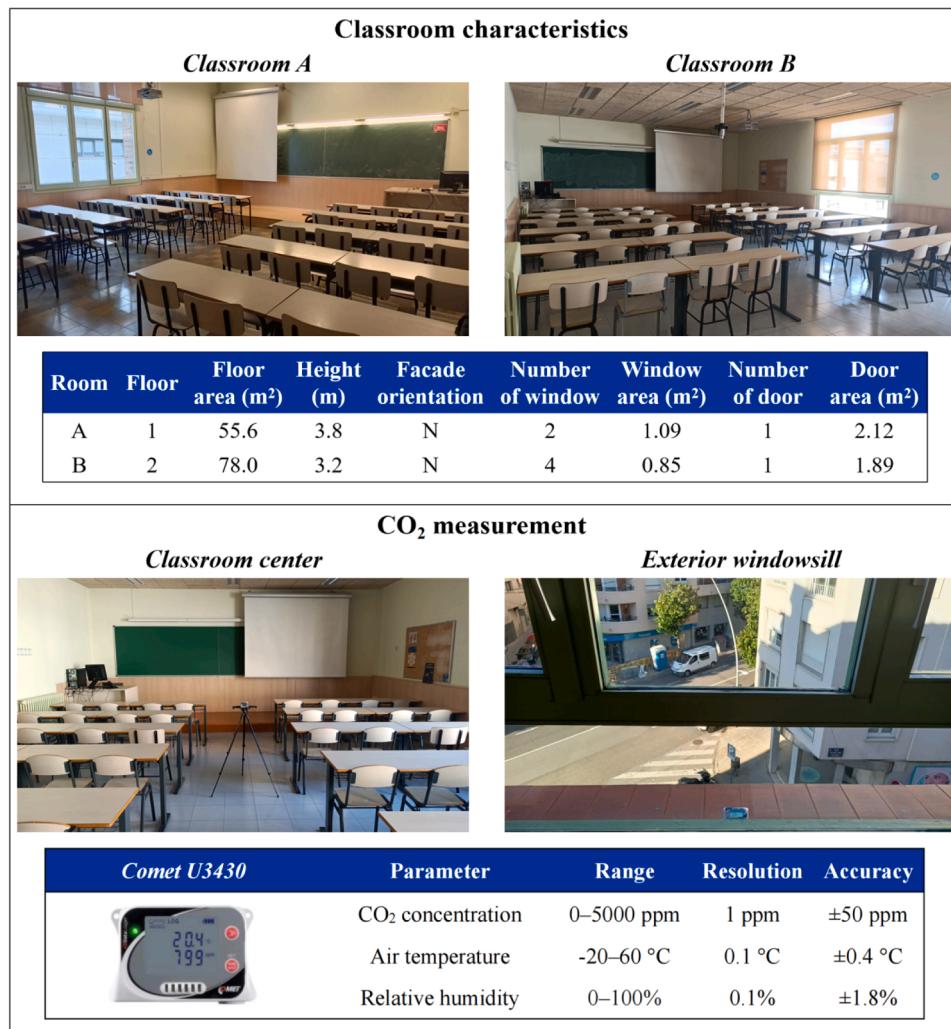


Fig. 2. Classroom characteristics and CO₂ measurement.

was placed on the exterior windowsill, measuring the CO₂ concentration of the outdoor airflow. The placement of the sensors and technical specifications are shown in Fig. 2. The sensors were turned on 15 min before each measurement to stabilize the readings and the data were recorded at 1-minute intervals.

During the measurement, the researchers stayed in the classroom to record the changes in room occupancy and the opening/closing of windows, also at 1-minute intervals. The door was remained closed during the tests, and only opened when people entered or left the room.

Hence, a total of 6 time series datasets were obtained, each lasting 200 min, containing indoor and outdoor CO₂ observations, the number of occupants, and the window opening areas.

2.4.2. Comparison of ventilation rate estimation techniques

Based on the data collected from the case study, the ventilation rates were estimated and compared using the transient mass balance model (TMBM), the extended Kalman filter (EKF), and the proposed adaptive particle filter (APF). TMBM uses the collected data directly without any processing. The EKF and APF take into account the uncertainties of the CO₂ generation rate and ventilation rate, and the measurement noise of CO₂.

For CO₂ generation rate G, referring to Persily and Jonge [28], the generation rate for male students in a sedentary state (1.2 met) is 0.288 L/min/person, and for female students is 0.228 L/min/person. The teacher's CO₂ generation rate is very close to that of the student (0.276 L/min/person and 0.216 L/min/person), therefore, the student's CO₂

generation rates were used. The reference CO₂ generation rates were directly used in the transient mass balance model, while for EKF and APF, the uncertainties of the CO₂ generation rate were considered. According to Persily and Jonge [28], the generation rate has a deviation of about 15 % between activity levels. Hence, the process noise variance σ_G^2 for the generation rate G is defined as 15 % of G.

In reality, the CO₂ generation rate G remains relatively stable and does not suddenly increase or decrease exponentially. However, the ventilation rate Q_{vent} can change drastically due to window operation. Therefore, the process noise $\sigma_{Q_{vent}}^2$ needs to be properly defined for the ventilation rate Q_{vent} to characterize the dynamic ventilation rate over time. The preliminary test found that a process noise $\sigma_{Q_{vent}}^2$ of 10 is an appropriate value, while a sensitivity analysis was also performed for examining and discussing the setting of this parameter.

Regarding the measurement noise of CO₂, the CO₂ sensor used in this study has a measurement error of ±50 ppm, which is similar to other commercial CO₂ measurement sensors available on the market. The error of the measurement instrument is typically considered to follow the "2 sigma" principle, which means that the σ_{sensor} that characterizes CO₂ measurement noise should be defined as 25.

The EKF and the APF used the same parameter settings relating to the ventilation rate, the CO₂ measurement noise, and the CO₂ generation rate. The EKF was also designed with a data-driven initialization like APF, with the same setting of the initial time steps m of 10. However, APF has an additional parameter of the particle number N_p , which was set to 1000. The setting of the initialization time steps m and the particle

number N_p is also to be examined and discussed later through a sensitivity analysis. Hence, Table 1 summarizes the parameter setting for EKF and APF.

The estimated ventilation rates were compared using total variation (TV). This indicator quantifies the overall oscillation (variation) of the curve that reflects the stability of the estimation, which can be calculated as follows:

$$TV = \sum_{i=1}^{n-1} |Q_{vent,i+1} - Q_{vent,i}| \quad (38)$$

Where $Q_{vent,i}$ is the estimated ventilation rate at time step i , $Q_{vent,i+1}$ is the estimated ventilation rate at time step $i + 1$, and n is the total number of the time step.

2.4.3. Sensitivity analysis

In addition to the comparison with existing ventilation rate estimation techniques, a sensitivity analysis based on the proposed APF was also conducted to explore the impact of the algorithm parameter settings and commonly used simplification strategies in practice, in order to better help relevant users in their practical applications.

Unlike CO₂ measurement noise and CO₂ generation rate uncertainties that can be defined based on available specifications and references, the settings of several parameters needs to be further examined and discussed through sensitivity analysis, including the ventilation rate process noise variance $\sigma_{Q_{vent}}^2$, particle number N_p , and initialization time step m . Table 2 summarizes the parameter values tested in the sensitivity analysis.

In practice, some simplification strategies are also commonly used, which also needs to be analyzed. These strategies include:

- (1) Using a fixed outdoor CO₂ concentration instead of real-time observations. Hence, according to NOAA [29], the atmospheric CO₂ concentration was around 420 ppm, which was used for the sensitivity analysis.
- (2) Using an averaged CO₂ generation rate for all occupants without distinguishing between genders. Hence, an averaged CO₂ generation rate of 0.258 L/min/person was used for the sensitivity analysis.
- (3) Using a constant room occupancy provided by the teachers (i.e. the number of students attending the class) instead of a real-time occupancy record. Hence, the occupancy of each measured class was assumed to be a constant value during the class period for the sensitivity analysis.

For the above sensitivity analysis, the mean percentage error (MPE) was used to measure the error in the estimated ventilation rates:

$$MPE = \frac{1}{n} \sum_{i=1}^n \left(\frac{Q_{vent,new,i} - Q_{vent,i}}{Q_{vent,i}} \times 100 \right) \quad (39)$$

Where $Q_{vent,i}$ is the estimated ventilation rate at time step i , $Q_{vent,new,i}$ is the estimated ventilation rate based on the simplification strategy at time step i , and n is the total time step.

In this study, the calculation and analysis were performed on the

Table 1
Parameter setting for extend Kalman filter and adaptive particle filter.

Parameter	Extended Kalman filter	Particle filter
CO ₂ measurement noise - σ_{sensor}	25	25
CO ₂ generation process noise variance - σ_G^2	15 %•G	15 %•G
Ventilation rate process noise variance - $\sigma_{Q_{vent}}^2$	10	10
Initialization time step - m	10	10
Particle number - N_p	/	1000

Table 2
Parameter values for sensitivity analysis.

Parameter	Value
Ventilation rate process noise variance - $\sigma_{Q_{vent}}^2$	1, 5, 10, 20, 30
Number of particles - N_p	1, 10, 100, 1000, 5000
Initialization time step - m	1, 5, 10, 20, 30

Google Colab platform using Python 3.10. The Numpy, Pandas, and SciPy packages were used for data processing and calculation, the Matplotlib package was used for results visualization, and the PyQt5 and pyinstaller packages were used for user interface and application development.

3. Results and discussion

This section analyzes and discusses the obtained results, including comparison of the estimated natural ventilation rate (Section 3.1), sensitivity analysis (Section 3.2), and further discussion on the practical implications and limitations (Section 3.3).

3.1. Comparison of the estimated ventilation rate

Fig. 3 shows the time series of indoor and outdoor CO₂ observations, the number of occupants, and the window opening area of the 6 measurements, each with a 200-minute time step. Appendix B provides supplementary reference on indoor and outdoor temperatures and outdoor reference wind speed during the measurement period. In general, indoor CO₂ concentrations ranged from 471 ppm to 1641 ppm, depending on occupancy and window operation. Outdoor CO₂ concentrations ranged from 385 to 459 ppm, with an average concentration of 420 ppm. This value is consistent with the atmospheric CO₂ concentration reported by NOAA [29]. In general, the number of occupants in classroom A was higher than that in classroom B. The occupancy of the classrooms is affected by the teaching schedule, sometimes with short breaks in the middle of the class (such as cases A-1, A-2, and B-1), and sometimes with some students leaving the class earlier (such as cases B-1 and A-3). The window operation in each measurement was different, and the openable window area of classroom B is nearly 1 time larger than that of classroom A. As seen, these case study measurements properly characterized and reflect the complex window operations and occupancy in real natural ventilation environments, which can effectively evaluate the capabilities of ventilation rate estimation techniques.

Based on the collected observations, the ventilation rate was estimated using the transient mass balance model (TMBM), the extended Kalman filter (EKF), and the proposed adaptive particle filter (APF), as shown in Fig. 4. Notably, in order to better observe the impact of window operation and occupancy changes on the estimation, the window opening period and the time of occupancy change are highlighted using green and orange background colors, respectively.

As can be seen in Fig. 4, the ventilation rate estimated by the TMBM has very large variations due to the CO₂ measurement noise. EKF handles most of the noise well, but the estimated ventilation rate is still associated with obvious oscillations. In contrast, the ventilation rate estimated by APF is significantly more stable and smooth. On average, the total variation of the ventilation rate estimated by the TMBM reaches 1315. The value is 552 for EKF and only 208 for APF. This means that the stability of APF estimation is over 6 times better than that of the conventional transient mass balance method and nearly 3 times higher than that of EKF.

More importantly, the TMBM and the EKF were found to be very sensitive to "instantaneous changes" in CO₂ observations. When the classroom windows were suddenly opened, the indoor CO₂ concentration dropped abruptly, causing the estimated ventilation rates to drift to a very high value and then gradually stabilize back. When the number of occupants in the room changed, the amount of CO₂ generation also

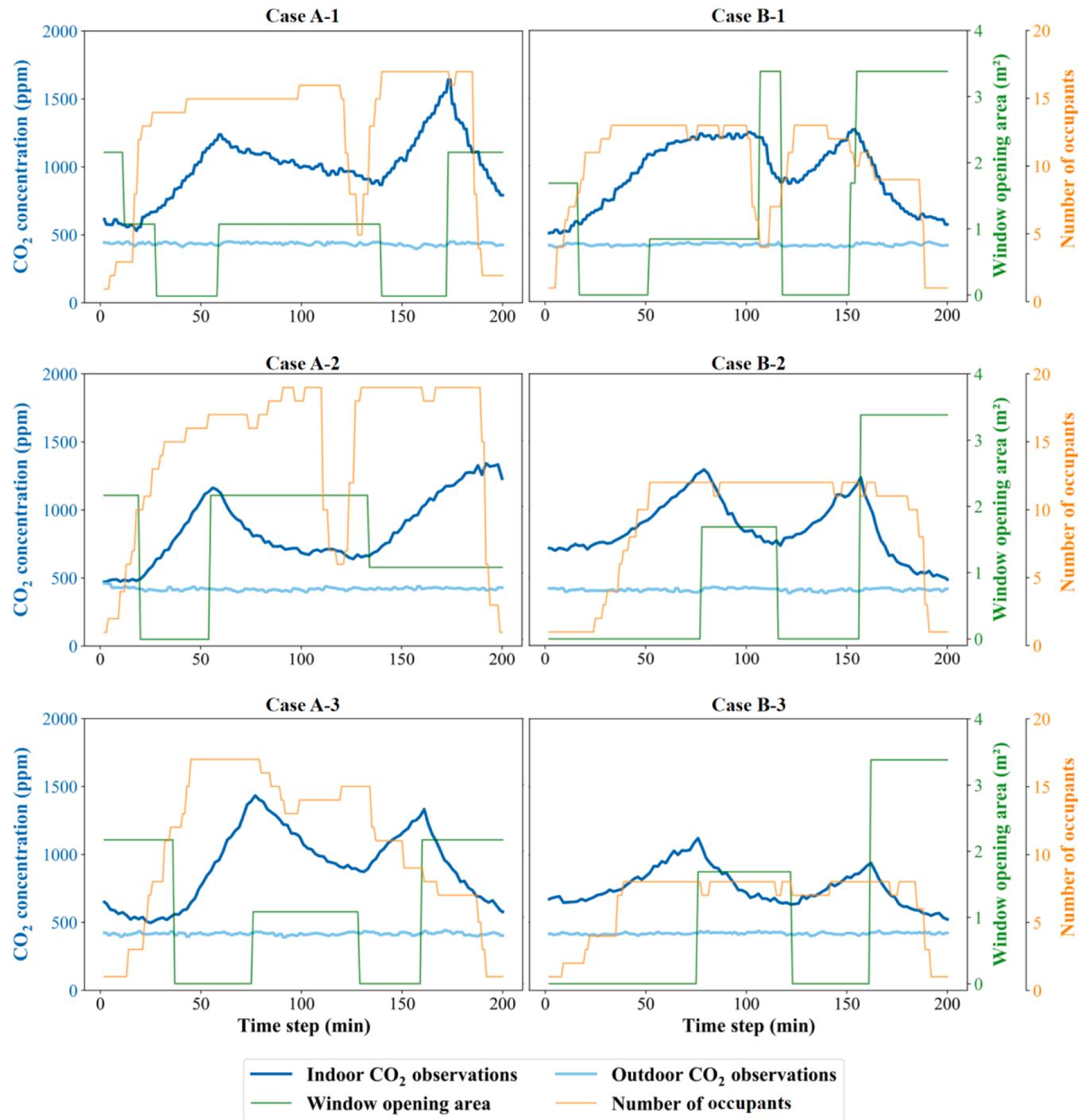


Fig. 3. Overview of measurement observations.

changed between adjacent time steps in the calculation, causing a sudden rise or drop in the estimated ventilation rates. These observed cases are certainly unreasonable in reality because the natural ventilation rate depends on buoyancy and wind effects and is not related to occupancy. In contrast, APF shows recognized stability to such sudden changes without drifts or abrupt rises and drops in the estimated ventilation rate.

In addition, it is worth mentioning that when the windows are closed, the indoor CO₂ concentration progressively rises, and the TMBM and EKF often generate an estimate of the ventilation rate close to 0. This is also unreasonable since there is always a small infiltration in the space even when the windows and doors are all closed. In contrast, APF estimated a small and stable ventilation rate in these periods, which is

certainly more consistent with reality.

The above results showed that the proposed APF demonstrated better stability, robustness, and practical applicability for the natural ventilation rate estimation.

3.2. Sensitivity analysis

This section analyzes and discusses the results of the sensitivity analysis based on the proposed APF, in terms of the algorithm parameter settings (Section 3.2.1) and common simplification strategies used in practice (Section 3.2.2).

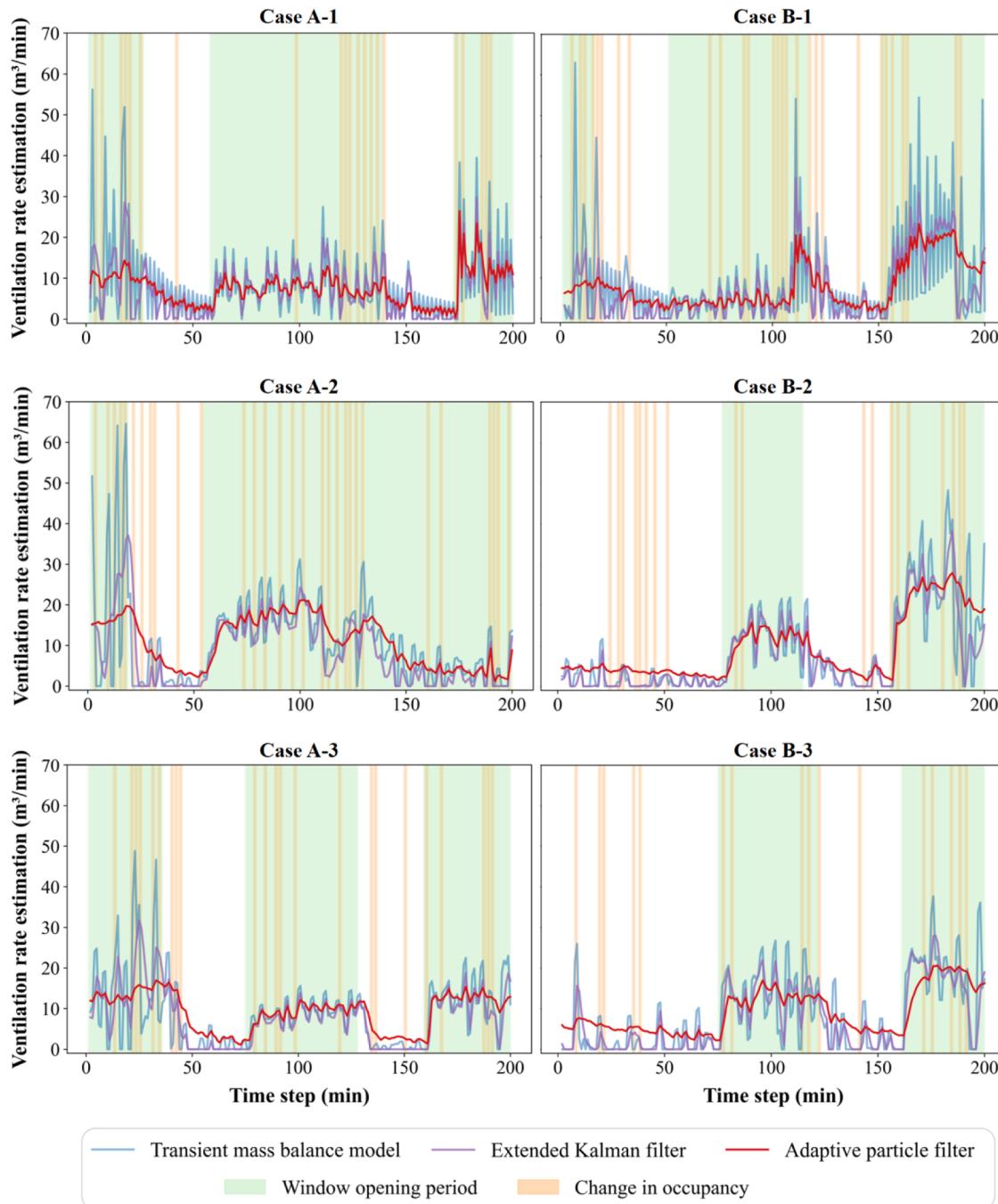


Fig. 4. Comparison of estimated ventilation rates.

3.2.1. Parameter settings

As mentioned in Section 2.4.3, unlike CO₂ measurement noise and CO₂ generation rate uncertainties that can be defined based on available specifications and references, the settings of several parameters for APF needs to be further examined and discussed, including the base process noise variance of ventilation rate σ_{Qvent}^2 , number of particles N_p , and initialization time steps m .

3.2.1.1. Base process noise variance. The setting of base process noise variance σ_{Qvent}^2 determines the allowed changes of ventilation rate per time interval (1 min in this study). A small value of base process noise variance may suppress potential variation, thus limiting the reaction to rapid changes in the ventilation rate. In contrast, a large value provides more flexibility, but may increase the variation of estimations to certain

extents. Since the natural ventilation is very dynamic and uncertain, there is no clear reference available for the setting of the base process noise variance and thus it needs to be determined empirically. Hence, a total of 5 different values were tested (Table 2), including: 1, 5, 10 (original), 20, and 30. Fig. 5 presents the estimated ventilation rates based on these values. Notably, the ventilation rates directly calculated by the original transient mass balance model (TMBM) are also visualized for a more intuitive comparison.

As seen, when the base process noise variance was set to 1, the estimated ventilation rates were very smooth, while the total variation declined to only 74 (the original APF was 208). However, the estimations in this case were substantially suppressed and could not properly reflect the dynamics of the ventilation rate. Compared with the original APF, the estimated ventilation rates were 11.57 % lower, as indicated by

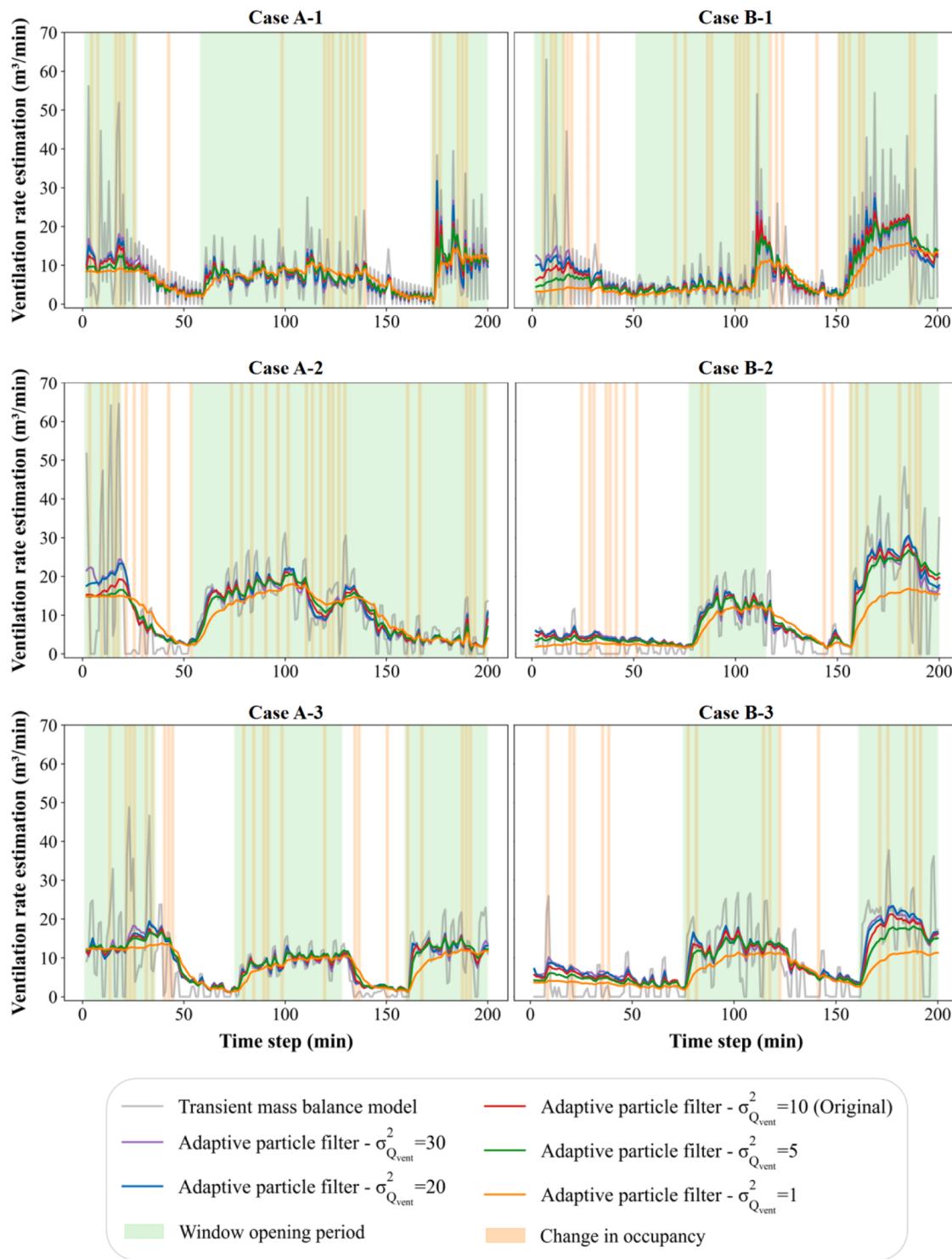


Fig. 5. Ventilation rate estimation with different parameter setting of base process noise variance.

the mean percentage error (MPE). The situation was significantly improved when the base process noise variance was set to 5. The MPE decreased to -2.43% , though the total variation slightly increased to 154. However, it can be observed that the ventilation rates were still suppressed when the window was suddenly opened, as demonstrated in cases A-1, B-1, B-2, and B-3. For the scenarios with a base process noise variance of 10 (original), 20, and 30, the estimated ventilation rates were consistent with each other and reacted quickly to sudden changes. Compared to the original APF, their estimation deviation (MPE) was only 2.35% and 3.09% , respectively. Larger base process noise variance

also led to greater fluctuations, as the total variation increased to 279 and 323, respectively. Nevertheless, these values are much lower than those of the original TMBM (1315) and the EKF (552).

Additionally, an interesting finding is that the setting of the base process noise variance can slightly affect the initialization of the estimation, especially when the initial state involves complex occupancy changes with window opening (e.g., cases A-1, B-1, and A-2). However, since the Bayesian filter can automatically calibrate the estimates based on newly observed information. The estimated ventilation rates became consistent after 10 steps of iterations.

3.2.1.2. Number of particles. The main feature of the particle filter is that it employs a large number of particles to explore and estimate system states, which can be defined by the number of particles N_p . A small value can significantly limit the power of the filter and the reliability of the estimation. Therefore, a relatively large value is preferable, though it leads to a higher computational demand. For comparison, a total of 5 different values were tested (Table 2), including: 1, 10, 100, 1000 (original), and 10,000. Fig. 6 shows the estimated ventilation rates based on these settings.

As can be seen, the APF was completely ineffective when the number of particles was set to 1. The estimation deviation (MPE) was nearly 249 % compared to the original APF (1000 particles). When using 10

particles, the situation was much better but the estimation drifted frequently, as shown in all cases. In contrast, estimations were very consistent when using 100, 1000, and 10,000 particles. Compared to the original APF (1000 particles), their estimation deviation (MPE) was only 0.5 %, while the total variation was almost the same.

Although APF is more computationally intensive than EKF, it is still very efficient. The calculations in this study were performed on the Google Colab platform with the default configuration of an Intel Xeon CPU and 13 GB of RAM. For the dataset with 200 time steps used in this study, it was found that the EKF can complete the computation within 0.2 s, whereas the original APF (with 1000 particles) used around 2 s. However, even with 10,000 particles, the APF can complete the

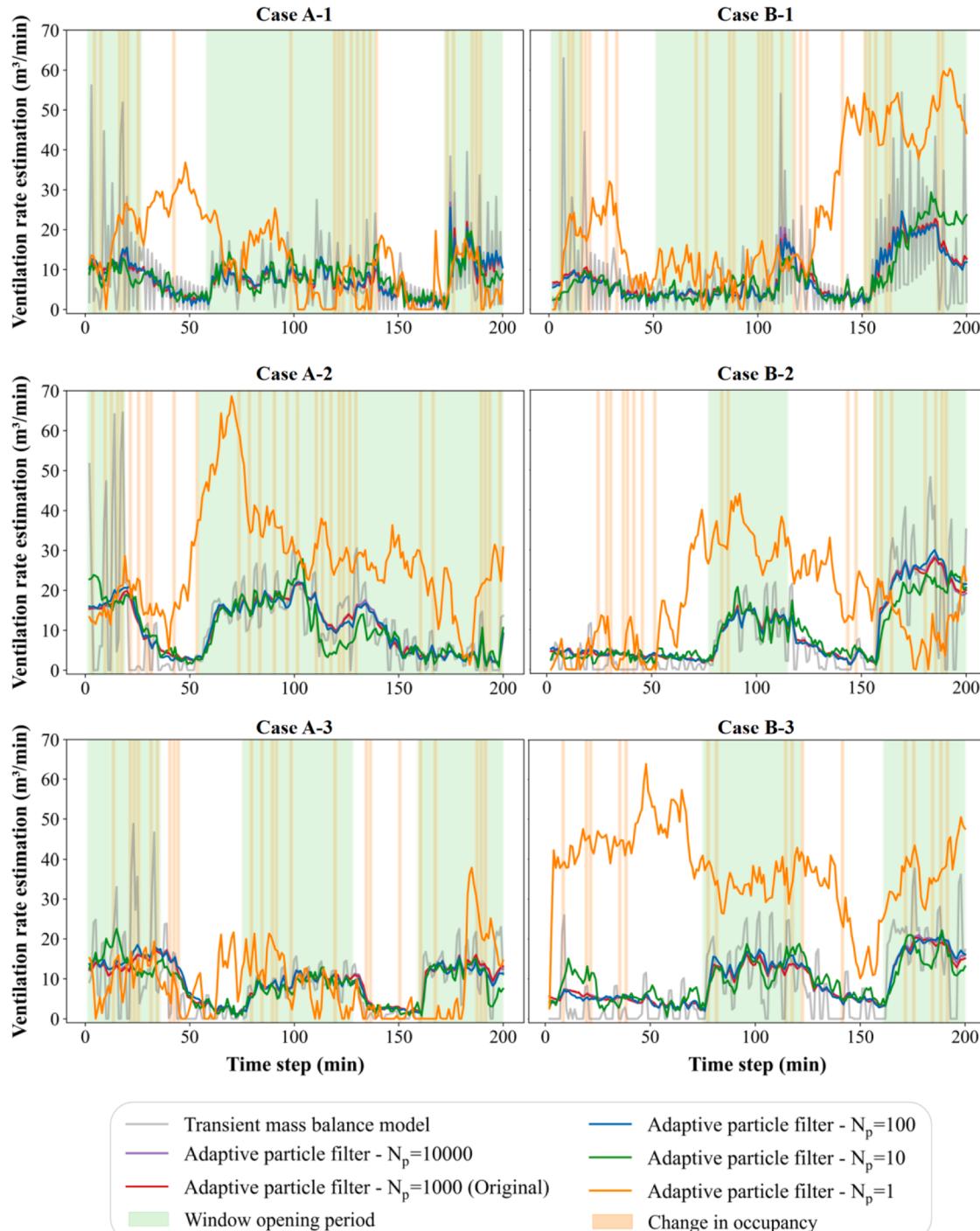


Fig. 6. Ventilation rate estimation with different parameter setting of particle numbers.

computation within 5 s. This computational demand is certainly acceptable for practical applications.

3.2.1.3. Initialization time steps. As introduced in Section 2.3, the APF was designed with a data-driven initialization, which initializes the estimation using the mean value of the first several ventilation rates directly calculated based on the original transient mass balance model (TMBM). This parameter can be set by the initialization time steps m . Similarly, a total of 5 different values were tested (Table 2), including 1, 5, 10, 20, and 30. Fig. 7 presents the estimated ventilation rates based on these defined values.

It can be seen that the setting of initialization time steps only

determines the first several estimations, while the subsequent estimations became consistent. When the classroom windows were initially closed, the initial ventilation rates were generally the same regardless of the initialization time steps, as shown in cases B-2 and B-3. However, when there were complex occupancy changes with window openings, the ventilation rate directly calculated by the TMBM varied greatly. Thus, the calculated mean ventilation rates differed considerably, which in turn affected the initialization of the APF (e.g., cases A-1 and A-2). However, as mentioned before, the Bayesian filter can automatically update and calibrate the estimation with new observations. Consequently, the estimations gradually became consistent after iterations regardless of their initial values. Therefore, setting the initialization time

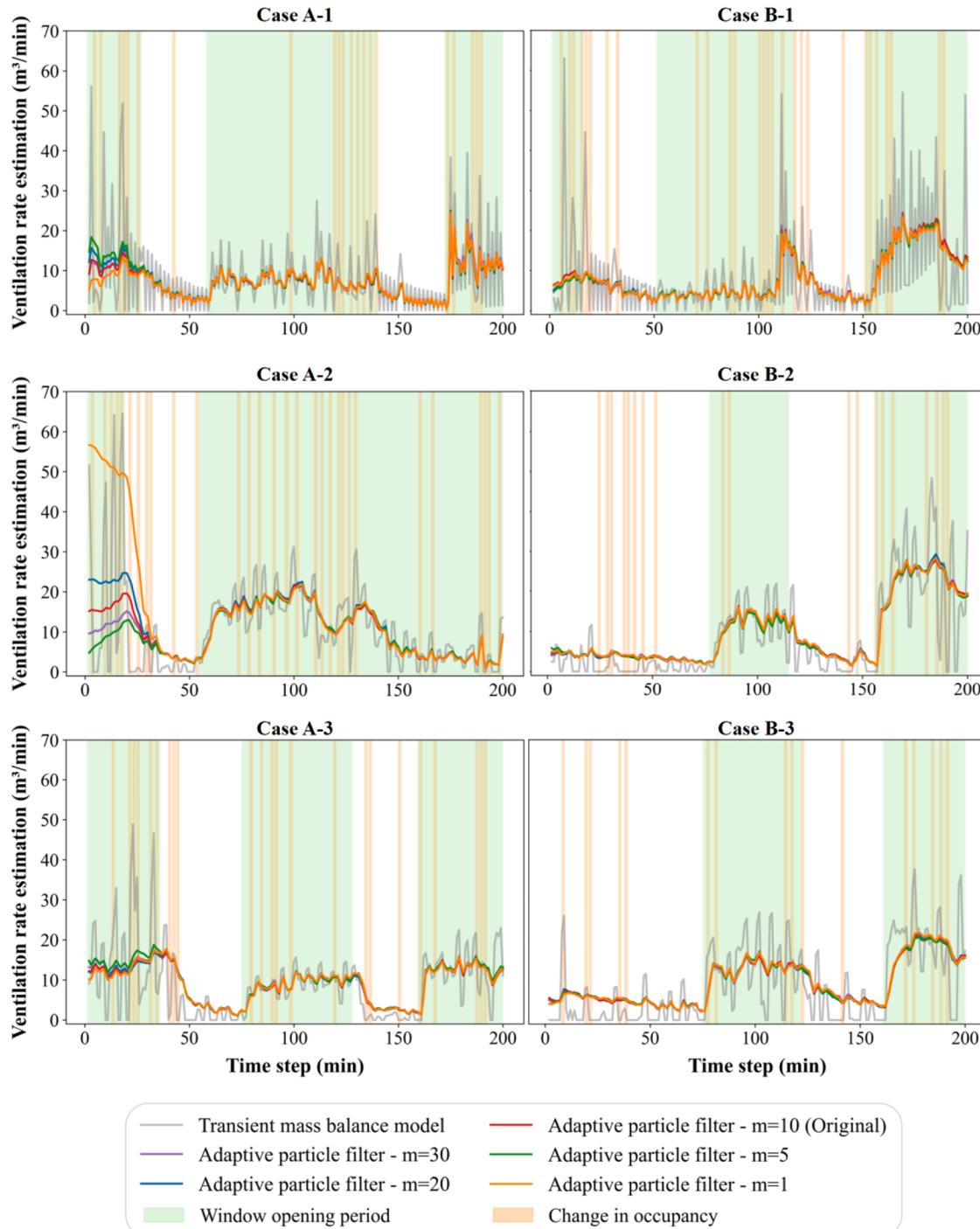


Fig. 7. Ventilation rate estimation with different parameter setting of initialization time steps.

steps to 5 or 10 is considered sufficient from a practical point of view.

3.2.2. Simplification strategies

In practice, a few simplification strategies are often adopted to reduce the complexity of data collection during field investigations. These strategies include the use of a fixed outdoor CO₂ concentration, an averaged CO₂ generation rate, and a constant room occupancy. In order to help relevant users apply the proposed APF in practice, the impact of these simplification strategies was further analyzed. The APF followed the original settings described in Section 2.4.2 (Table 1).

3.2.2.4. Fixed outdoor CO₂ concentration. The field studies usually assume a fixed outdoor CO₂ concentration when estimating ventilation rates, rather than using real-time outdoor CO₂ measurements.

Therefore, according to NOAA [29], a constant outdoor CO₂ concentration of 420 ppm was used for the sensitivity analysis. As mentioned earlier, the average value of the measured outdoor CO₂ is also 420 ppm, but with a fluctuation between 385 and 459 ppm.

Fig. 8 shows the ventilation rates estimated using the assumed outdoor CO₂ and the measured outdoor CO₂. As seen, as the real outdoor CO₂ fluctuates, small overestimation and underestimation occur in some cases after using a fixed outdoor CO₂ concentration. However, the estimation had a very small overall deviation, less than ±1.5 % on average. This suggests that using a fixed CO₂ concentration does not have a serious impact on the estimated ventilation rate, but it should be noted that the assumed outdoor CO₂ value should be close to the measured value.

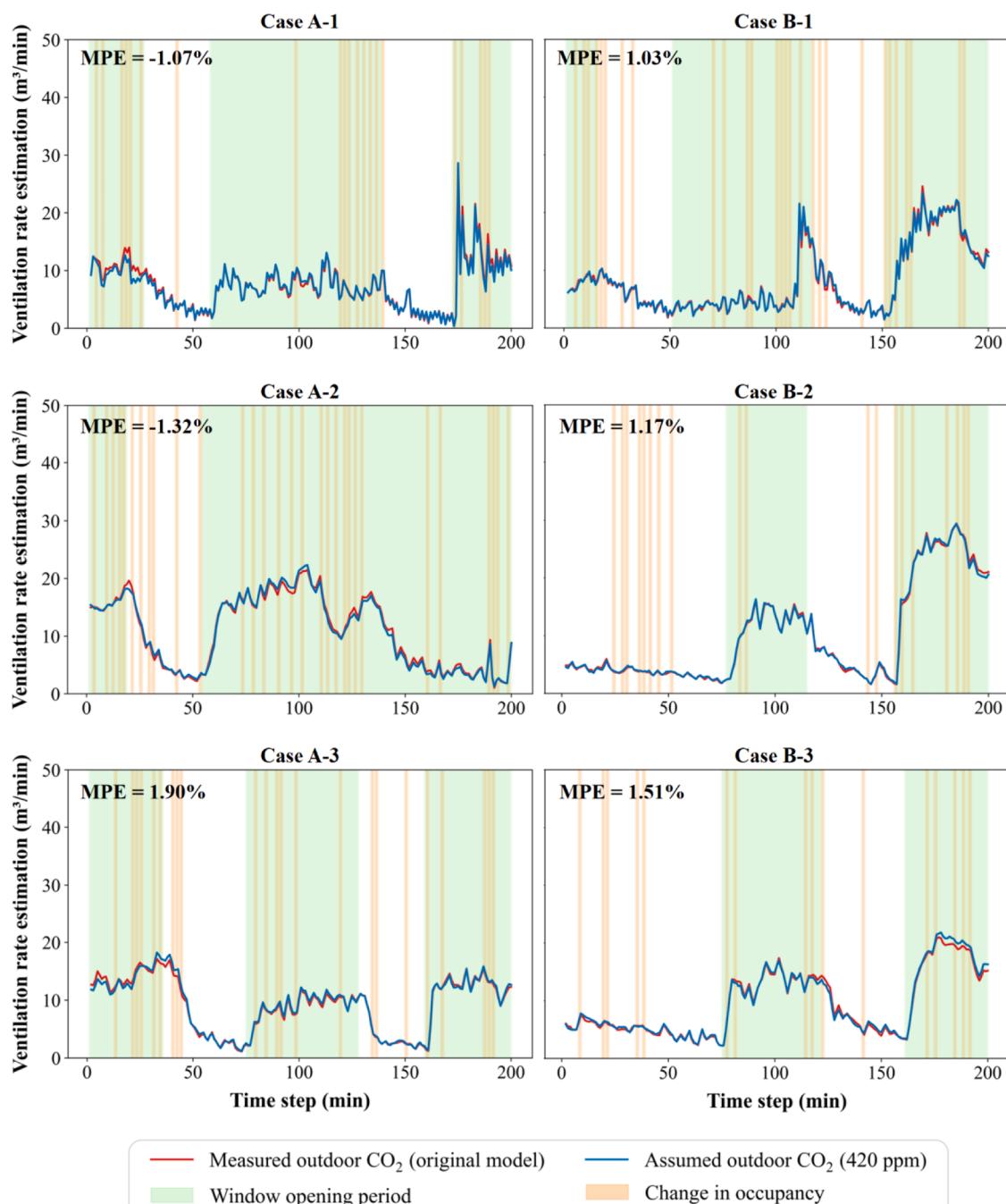


Fig. 8. Estimated ventilation rates by assumed and measured outdoor CO₂.

3.2.2.5. Averaged CO₂ generation rate. Relevant field studies often assume an average CO₂ generation rate for all occupants regardless of gender. In this study, the original ventilation rate was estimated using CO₂ generation rates of 0.288 L/min/person for male students and 0.228 L/min/person for female students. Hence, an average value of 0.258 L/min/person was used for the sensitivity analysis, and Fig. 9 shows the results.

As mentioned in Section 2.3, APF can handle the uncertainty of the CO₂ generation rate and estimate it together with the ventilation rate as hidden states (latent variables). It was found that the estimated CO₂ generation rate for male students ranged from 0.233 to 0.356 L/min/person, with an average of 0.303 L/min/person. For female students, it ranged from 0.206 to 0.260 L/min/person, with an average of 0.231 L/min/person. As mentioned in the case study description, the ratio

between male and female students is about 80 %–20 %, so the real average generation rate should be around 0.283 L/min/person. However, when the assumed average generation rate of 0.258 L/min/person was used in the estimation, the estimated generation rate ranged from 0.213 to 0.328 L/min/person, with an average of 0.270 L/min/person, which is lower than the actual CO₂ generation rate. Therefore, as shown in Fig. 9, when an average CO₂ generation rate was used, the ventilation rate was slightly underestimated by 2.98 % to 5.60 %, within 4 % on average.

However, judging from the overall results, the impact of using the average generation rate on the estimation is not very serious. This is because the CO₂ sensors have limited measurement precision and it is difficult to effectively capture small differences in the CO₂ generation rates. Assuming a room of 250 m³, a person with a CO₂ generation rate

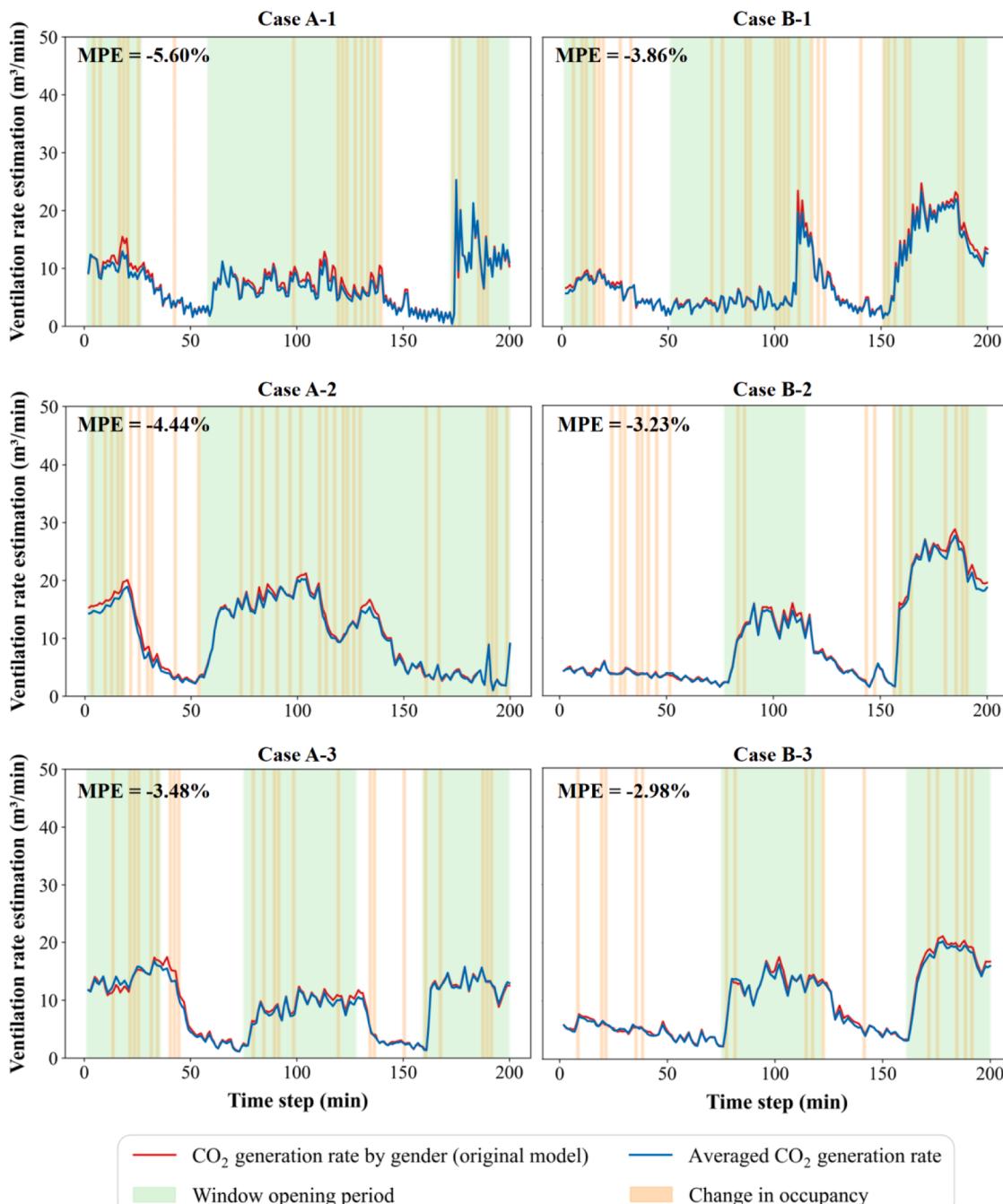


Fig. 9. Estimated ventilation rates by original and averaged CO₂ generation rate.

of 0.258 L/min could only result in an increase in indoor CO₂ of around 1 ppm/min. This is far below the CO₂ measurement error of ± 50 ppm. Therefore, the CO₂ measurement noise still accounts for the main impact on the estimation results. However, given the small error it may cause, a practical strategy is to use the averaged CO₂ generation rate based on the ratio of male to female occupants in the room.

3.2.2.6. Constant room occupancy. The field studies often use a constant occupancy based on the number of students in the class provided by the teacher, rather than using real-time occupancy observations. Hence, a sensitivity analysis was performed assuming a constant number of students during the class period for each measurement. Fig. 10 shows the comparison between the assumed occupancy and the real occupancy. In general, the A-1, A-2, A-3, and B-1 cases had large variations during the class, so the assumed occupancy was much higher than the actual value. In contrast, the assumed occupancy for B-2 and B-3 cases is very close to

the actual value. Notably, the assumed occupancy did not take into account the fact that students gradually entered and left the classroom.

For simplicity of analysis, the occupants were considered as a whole without distinguishing between genders when adjusting the occupancy, so the ventilation rate for both real and assumed occupancy was calculated based on the average CO₂ generation rate. Fig. 11 shows the estimated ventilation rates. It can be seen that the A-1, A-2, A-3, and B-1 cases had a large overestimation from 9.70 % to 17.73 %, generally over 10 %. In contrast, B-2 and B-3 had a very small underestimation of about 3 %. This is because, unlike the other cases, these two cases had a very small change in occupancy, so the assumed occupancy was very close to the real value. Accordingly, the underestimations were mainly from the periods when students entered and left the classroom.

The results suggest that, in practice, small and short term changes in occupancy (e.g., one or two students going to the lavatory for a short time) may not have a serious impact on the estimated results. However,

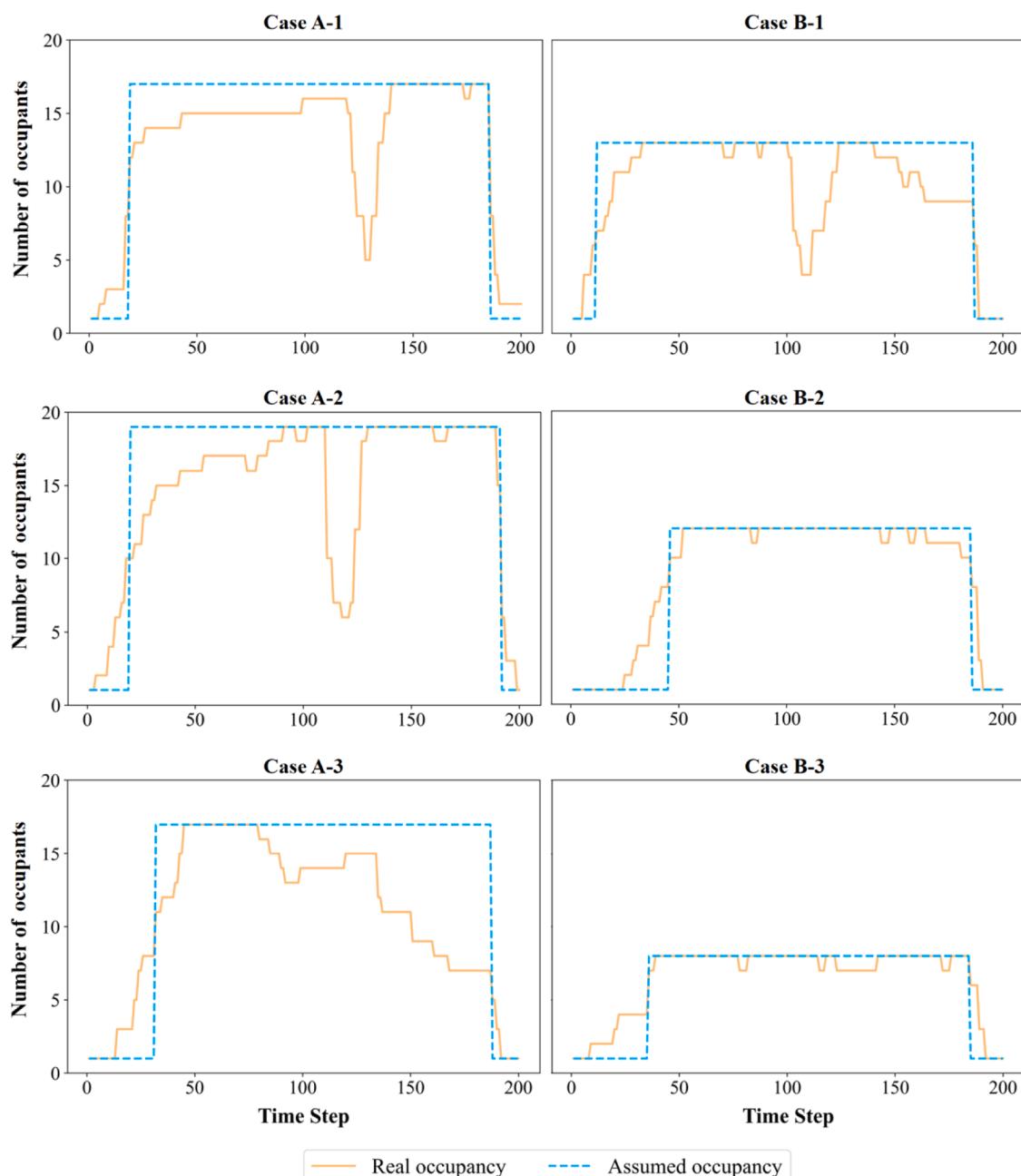


Fig. 10. Comparison between real and assumed occupancy.

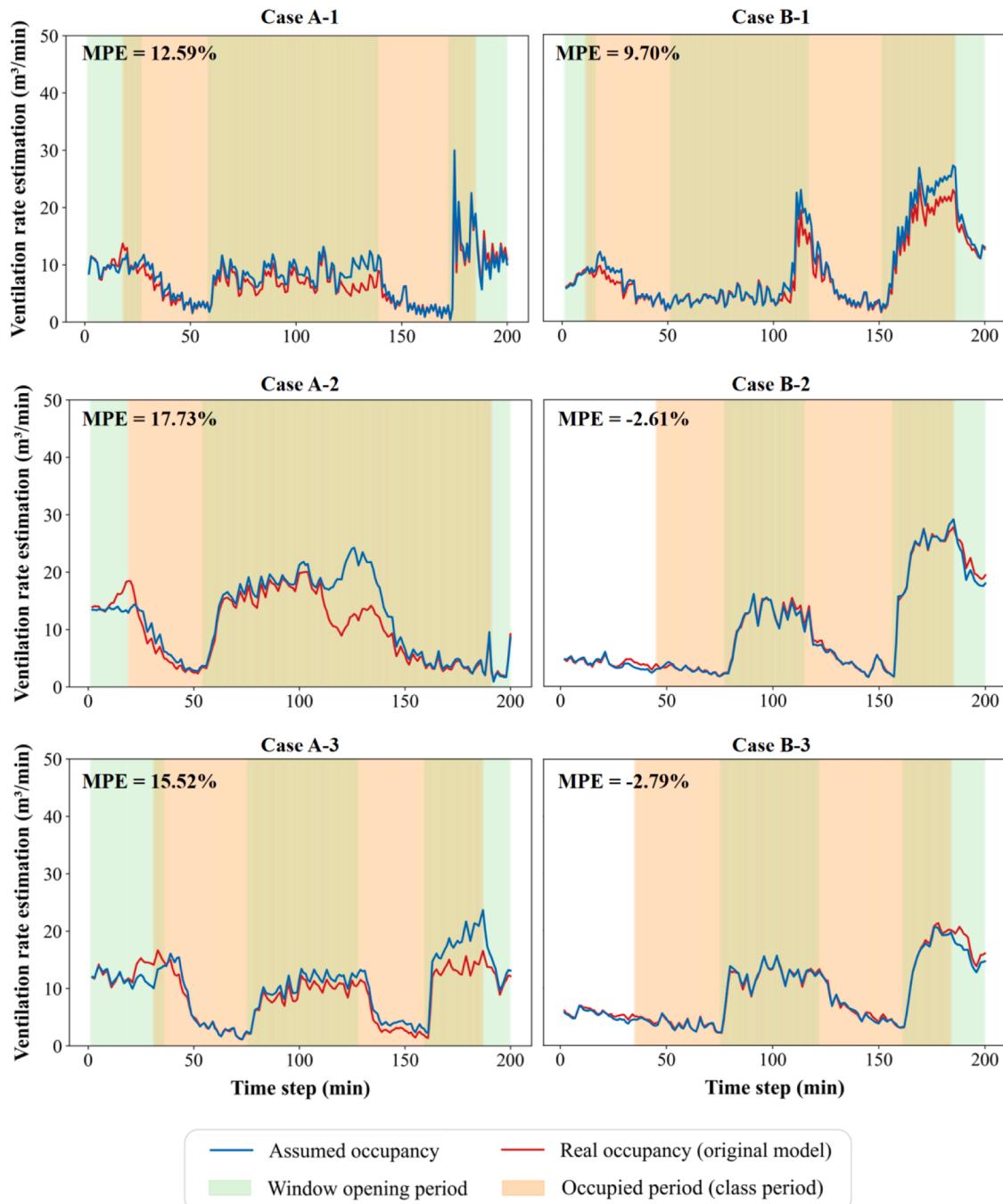


Fig. 11. Estimated ventilation rate by real and assumed occupancy.

it should be very cautious about the situations where there may be large changes (e.g., half of the students leaving in the middle of the class). In reality, teachers may be focused on teaching and may not record and report these changes in detail, which can lead to significant estimation errors due to missing information.

3.3. Practical implications and limitations

It is necessary to further discuss the practical implications of the developed APF algorithm to help future research. Firstly, considering that the execution of the developed algorithm requires certain programming skills, an open, user-friendly application software was developed to help relevant researchers and engineers apply the proposed algorithm in practice. The software has been compiled into an

executable (.exe) file and encapsulates all the required Python packages (Table 3). Users can run this application directly on a 64-bit Windows system without any additional requirements. The software is available on the GitHub platform: <https://github.com/Misaeon/CO2-Based-Natural-Ventilation-Rate-Estimation-Tool>

Table 3
Python packages, versions and functions for the developed software.

Package	Version	Function
Numpy	1.26.4	Processing & computation
Pandas	2.2.2	Processing & computation
SciPy	1.31.1	Processing & computation
Matplotlib	3.8.4	Results visualization
PyQt5	12.13.0	User interface

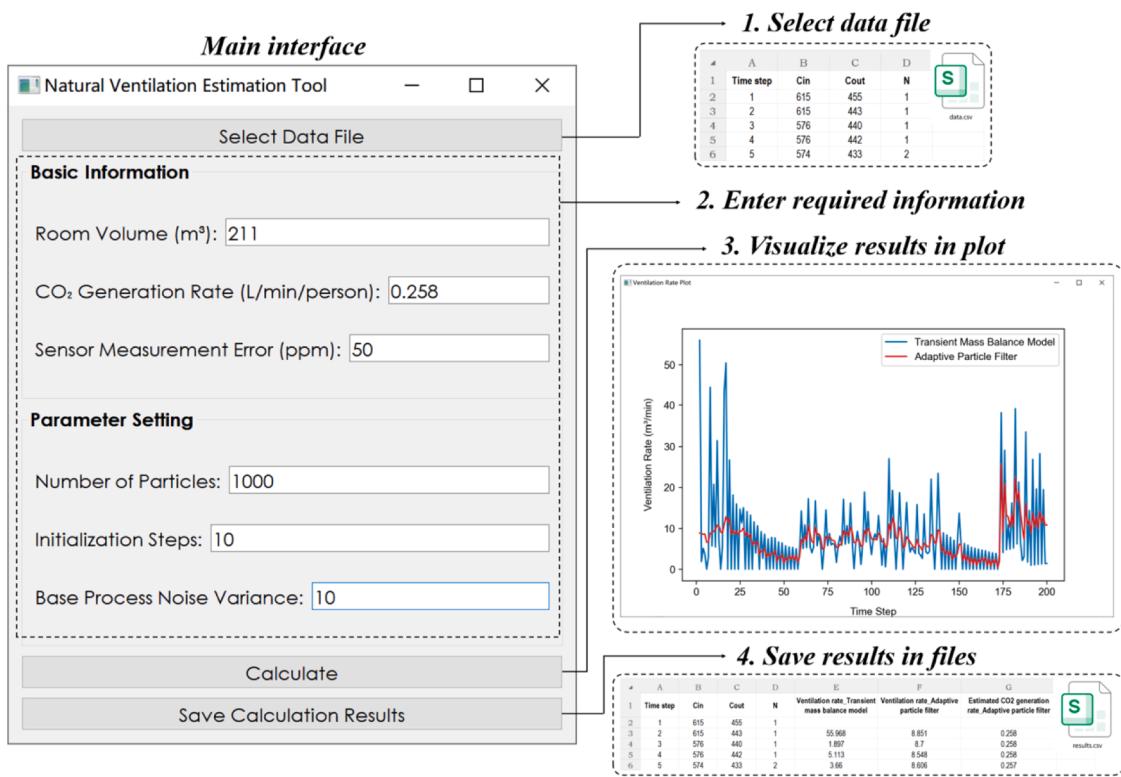


Fig. 12. Execution process of the developed software.

Fig. 12 shows the execution process of the developed software. First, the user needs to save the data collected from the field measurement into a .csv file, which should include the columns of the measurement time step (Time step), the corresponding indoor CO₂ observation (C_{in}), the outdoor CO₂ observation (C_{out}), and the number of occupants (N). The .csv file can be loaded by clicking the "Select Data File" button on the user interface. Then, the user needs to enter the required information for ventilation rate estimation on the user interface, including the volume of the room, the CO₂ generation rate, the measurement error of the CO₂ sensor, as well as the parameter setting regarding the number of particles, the initialization steps, and base process noise variance. Next, when the "Calculate" button is clicked, the software calculates the ventilation rate based on both the conventional transient mass balance model and the proposed APF, and displays the estimated ventilation rates in a line graph. Finally, by clicking the "Save Calculation Results" button, the user can save the calculation results locally on the computer. Notably, in addition to the "Ventilation rate_Transient mass balance model" calculated by the mass balance model and the "Ventilation rate_Adaptive particle filter" estimated by the APF, the results also include the estimated CO₂ generation rate by the APF in the column of "Estimated CO₂ generation rate_Adaptive particle filter".

In theory, the temporal resolution of the estimation by the proposed APF algorithm depends entirely on the interval of the measurement data. This study used observations with 1-minute intervals. If relevant users can collect data at smaller intervals (e.g., 30 or 10 s), the APF can achieve an even higher temporal resolution that better reflects the ventilation dynamics. Moreover, the proposed APF algorithm also has great practical value to be integrated into smart building management systems. Nowadays, Internet of Things (IOT) sensors can measure real-time CO₂ concentrations and identify real-time indoor occupants through deep vision [30,31]. Since the proposed APF algorithm has better reliability and robustness than existing techniques in naturally ventilated indoor environments, it is very suitable to be integrated with IOT monitoring systems for real-time ventilation rate estimation, contributing to future smart indoor environment management.

Certainly, the limitation of this study also needs to be discussed to help future research. The natural ventilation environment is often associated with high complexity and uncertainty, which brings challenges to the field measurements. This study was conducted in real classroom environments without measuring the ground-truth ventilation rate using balometers, because it would cause serious disturbances to the teaching activities. Therefore, future research is recommended to conduct experiments in controlled laboratories in order to measure the ground-truth of the natural ventilation rates to benchmark the APF estimates. This can provide valuable references for the setting of the ventilation process noise and contribute to the improvement of the fundamental state-space model for the Bayesian filter (both EKF and APF). In addition, this study used a single-point (room center) CO₂ measurement, which is based on the assumption that CO₂ is homogeneously mixed under natural ventilation. This method is effective for spaces with a floor area of less than 100m² [23]. However, whenever possible, relevant field studies are recommended to use the average value of multiple-point measurements to further improve the accuracy of CO₂ observations.

4. Conclusions and recommendations

This research proposes an adaptive particle filter (APF) algorithm for CO₂-based natural ventilation rate estimation. The performance of the proposed algorithm is validated through a case study, and the effects of commonly used simplification strategies in practical applications are

analyzed through sensitivity analysis.

The results show that the proposed APF algorithm improves the estimation stability by nearly 6 times and 3 times, compared to the existing transient mass balance method and extended Kalman filter, respectively. More importantly, the proposed algorithm is significantly more robust to abrupt changes in indoor CO₂ and does not cause large drifts in the estimated ventilation rate due to window openings or sudden changes in occupancy. In addition, unlike existing techniques, APF effectively estimates stable, small infiltrations. The proposed APF algorithm demonstrated great capacity for natural ventilation rate estimation with high temporal resolution and stability, with recognized practical applicability for naturally ventilated environments.

The sensitivity analysis results suggest that the parameters of the APF algorithm should be set to reasonable values. It is recommended that the number of particles be more than 100, while the base process noise variance of the ventilation rate can be 10 to 20. Smaller particle sizes and base process noise would limit the power of the APF and affect the reliability of the estimation. For initialization time steps, a value of 5 or 10 steps is considered sufficient for practical applications. For simplification strategies commonly used in field investigations, adopting a fixed outdoor CO₂ value may cause a very small estimation error of less than ±2 %. Using an averaged CO₂ generation rate can lead to an overestimation of less than 5 %, so it is recommended to calculate the average CO₂ generation rate based on the actual gender ratio in practice. When assuming a constant room occupancy, unrecorded large changes in the room occupancy may lead to an overestimation of more than 15 %. Therefore, it is always recommended to record real-time occupancy observations for ventilation rate estimation.

Finally, considering that the implementation of the proposed APF requires programming skills, an open, user-friendly software has been developed to help relevant users apply it more conveniently in practice. Future research should include independent airflow measurements to benchmark the APF estimates, further explore the possibility of further improving the proposed algorithm, as well as the integration of the proposed algorithm with the IOT monitoring systems for real-time natural ventilation rate estimation.

CRediT authorship contribution statement

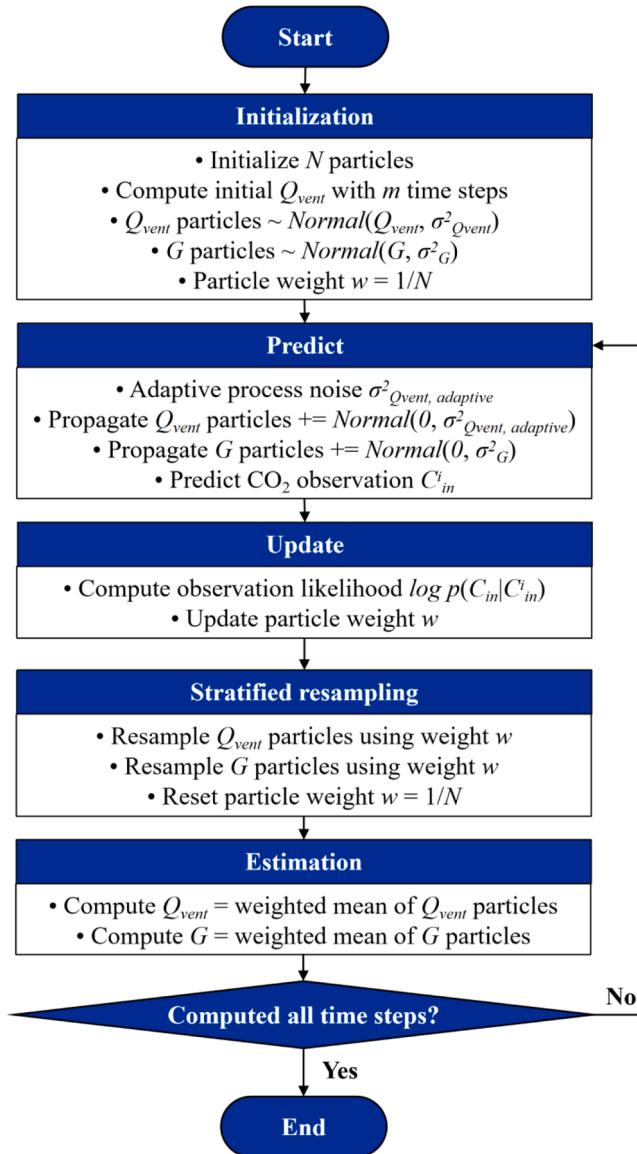
Sen Miao: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Marta Gangolells:** Writing – review & editing, Supervision, Project administration, Conceptualization. **Blanca Tejedor:** Writing – review & editing, Supervision, Conceptualization.

Declaration of competing interest

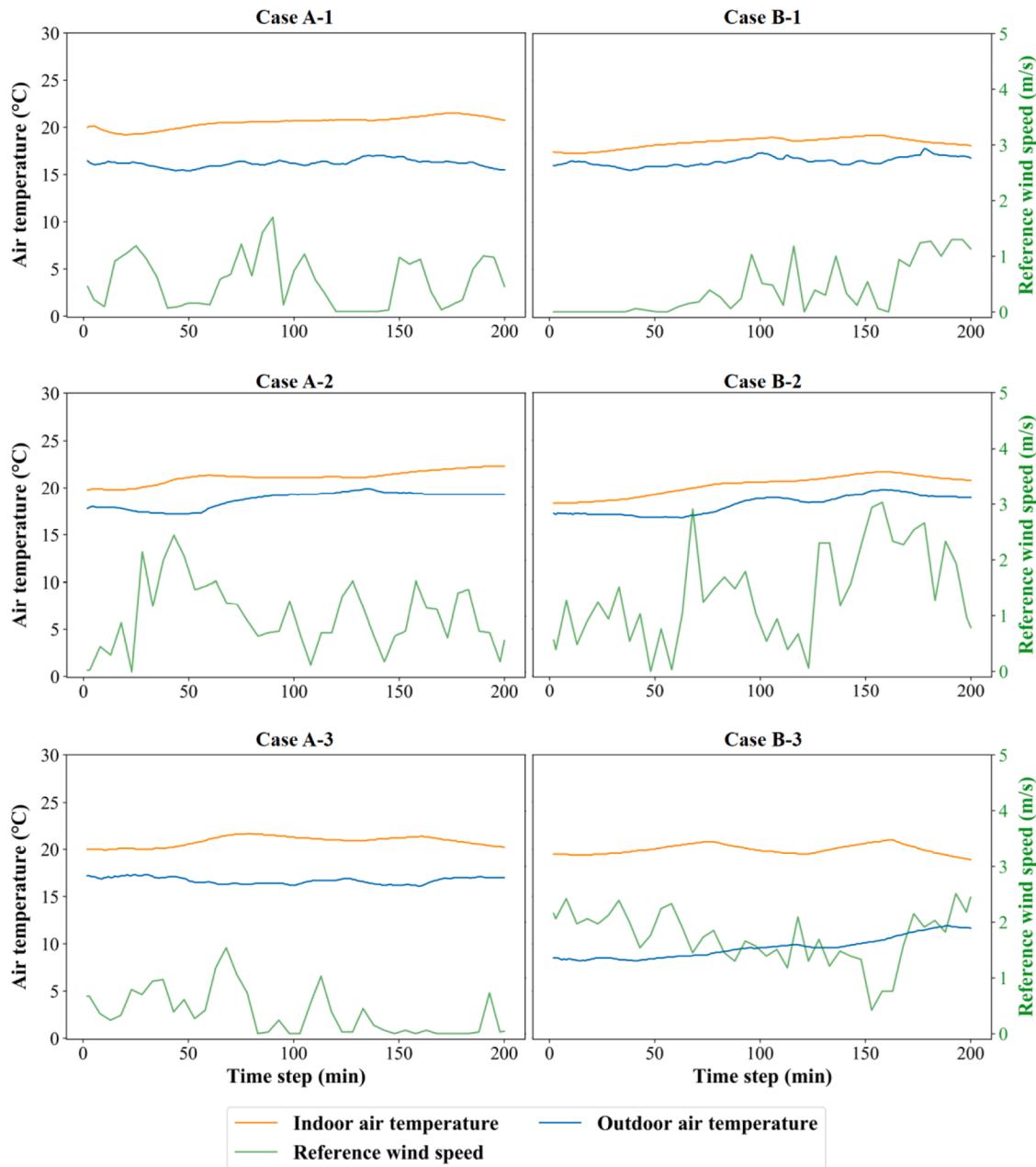
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research is part of the R&D project IAQ4EDU, reference no. PID2020-117366RB-I00, funded by MCIN/AEI/10.13039/501100011033. This work was supported by the Catalan Agency AGAUR under their research group support program (2021 SGR 00341). The author Sen Miao is funded by the China Scholarship Council (CSC) as a full-time PhD student, reference no. 202208390065.

Appendix A. Flow chart of the proposed adaptive particle filter


Appendix B. Indoor/outdoor air temperature and reference wind speed



*Indoor and outdoor air temperatures were measured at 1-minute intervals using Comet U3430 sensors in the case study. Outdoor reference wind speed was measured at 5-minute intervals by a Bresser 5-in-1 weather station located inside the university campus, while the observations were then interpolated to 1-minute intervals.

Data availability

Data will be made available on request.

References

- [1] M.T. Miranda, P. Romero, V. Valero-Amaro, J.I. Arranz, I. Montero, Ventilation conditions and their influence on thermal comfort in examination classrooms in times of COVID-19. A case study in a Spanish area with Mediterranean climate, Int J Hyg Env. Health 240 (2022) 113910, <https://doi.org/10.1016/j.ijheh.2021.113910>.
- [2] S. Miao, M. Gangolells, B. Tejedor, A comprehensive assessment of indoor air quality and thermal comfort in educational buildings in the Mediterranean climate, Indoor Air (2023) 6649829, <https://doi.org/10.1155/2023/6649829>, 2023.
- [3] CEN, EN 16798-1:2019, Part 1: Indoor environmental Input Parameters For Design and Assessment of Energy Performance of Buildings Addressing Indoor Air quality, Thermal environment, Lighting and Acoustics, European Committee for Standardization (CEN), Brussels, 2019.
- [4] ASHRAE. 2022. ASHRAE 62.1: ventilation and acceptable indoor air quality. The American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE), Atlanta.
- [5] E. Jones, A. Young, K. Clevenger, P. Salimifard, E. Wu, M.L. Luna, M. Lahvis, J. Lang, M. Bliss, P. Azimi, J. Cedeno-Laurent, C. Wilson, J. Allen, Healthy schools:

- Risk Reduction Strategies For Reopening Schools, Harvard TH Chan School of Public Health Healthy Buildings program, 2020.
- [6] IDAEA, Ventilation guide for indoor spaces. <https://www.idaea.csic.es/wp-content/uploads/2021/12/Ventilation-guide-of-indoor-environments.pdf>, 2021. >Accessed on 10 January 2025.
- [7] A.J. Aguilar, D.P. Ruiz, M.D. Martínez-Aires, M.L de la Hoz Torres, Indoor environment in educational buildings: assessing natural ventilation, in: D. Bienvenido-Huertas, J. Durán-Alvarez (Eds.), Building Engineering Facing the Challenges of the 21st Century, 2023, p. 345, https://doi.org/10.1007/978-981-99-2714-2_24. Lecture Notes in Civil Engineering.
- [8] T.S. Larsen, C. Plesner, V. Leprinse, F.R. Carrié, A.K. Bejder, Calculation methods for single-sided natural ventilation: now and ahead, In Energy Build 177 (2018) 279–289, <https://doi.org/10.1016/j.enbuild.2018.06.047>.
- [9] H.Y. Zhong, Y. Sun, J. Shang, F.-P. Qian, F.Y. Zhao, H. Kikumoto, C. Jimenez-Bescos, X. Liu, Single-sided natural ventilation in buildings: a critical literature review, Build. Env. 212 (2022) 108797, <https://doi.org/10.1016/j.buildenv.2022.108797>.
- [10] S. Batterman, Review and extension of CO₂-based methods to determine ventilation rates with application to school classrooms, Int. J. Env. Res. Public Health 14 (2) (2017) 145, <https://doi.org/10.3390/ijerph14020145>.
- [11] G. Remion, B. Moujalled, M. El Mankibi, Review of tracer gas-based methods for the characterization of natural ventilation performance: comparative analysis of their accuracy, Build. Env. 160 (2019) 106180, <https://doi.org/10.1016/j.buildenv.2019.106180>.
- [12] ASTM, ASTM D6245-18. Standard guide For Using Indoor Carbon Dioxide Concentrations to Evaluate Indoor Air Quality and Ventilation, American Society for Testing and Materials (ASTM), Pennsylvania, 2018.
- [13] D.L. Johnson, R.A. Lynch, E.L. Floyd, J. Wang, J.N. Bartels, Indoor air quality in classrooms: environmental measures and effective ventilation rate modeling in urban elementary schools, Build. Env. 136 (2018) 185–197, <https://doi.org/10.1016/j.buildenv.2018.03.040>.
- [14] A. Asif, M. Zeeshan, Indoor temperature, relative humidity and CO₂ monitoring and air exchange rates simulation utilizing system dynamics tools for naturally ventilated classrooms, Build. Env. 180 (2020) 106980, <https://doi.org/10.1016/j.buildenv.2020.106980>.
- [15] S.S. Korsavi, A. Montazami, D. Mumovic, Ventilation rates in naturally ventilated primary schools in the UK; contextual, occupant and Building-related (COB) factors, Build. Env. 181 (2020) 107061, <https://doi.org/10.1016/j.buildenv.2020.107061>.
- [16] M. Gil-Baez, J. Lizana, J.A. Becerra Villanueva, M. Molina-Huelva, A. Serrano-Jimenez, R. Chacartegui, Natural ventilation in classrooms for healthy schools in the COVID era in Mediterranean climate, Build. Env. 206 (2021) 108345, <https://doi.org/10.1016/j.buildenv.2021.108345>.
- [17] M. Shrestha, H.B. Rijal, G. Kayo, M. Shukuya, An investigation on CO₂ concentration based on field survey and simulation in naturally ventilated Nepalese school buildings during summer, Build. Env. 207 (2022) 108405, <https://doi.org/10.1016/j.buildenv.2021.108405>.
- [18] R. Duarte, M. Glória Gomes, A. Moret Rodrigues, Estimating ventilation rates in a window-aired room using Kalman filtering and considering uncertain measurements of occupancy and CO₂ concentration, Build. Env. 143 (2018) 691–700, <https://doi.org/10.1016/j.buildenv.2018.07.016>.
- [19] A. Kabirikopaei, J. Lau, Uncertainty analysis of various CO₂-based tracer-gas methods for estimating seasonal ventilation rates in classrooms with different mechanical systems, Build. Env. 179 (2020) 107003, <https://doi.org/10.1016/j.buildenv.2020.107003>.
- [20] R. Soyer, Kalman filtering and sequential bayesian analysis, WIREs Comput. Stat. 10 (5) (2018) e1438, <https://doi.org/10.1002/wics.1438>.
- [21] G. Remion, B. Moujalled, M. El Mankibi, Dynamic measurement of the airflow rate in a two-zones dwelling, from the CO₂ tracer gas-decay method using the Kalman filter, Build. Env. 188 (2021) 107493, <https://doi.org/10.1016/j.buildenv.2020.107493>.
- [22] N. Mahyuddin, H. Awbi, M. Alshitawi, The spatial distribution of carbon dioxide in rooms with particular application to classrooms, Indoor Built Environ. 23 (3) (2013) 433–448, <https://doi.org/10.1177/1420326x13512142>.
- [23] N. Mahyuddin, H. Awbi, A review of CO₂Measurement procedures in ventilation research, Int. J. Vent. 10 (4) (2012) 353–370, <https://doi.org/10.1080/14733315.2012.11683961>.
- [24] R.E. Kalman, A new approach to linear filtering and prediction problems, J. Basic Eng 82 (1) (1960) 35–45, <https://doi.org/10.1115/1.3662552>.
- [25] J. Mochnac, S. Marchevsky, P. Kocan, Bayesian filtering techniques: kalman and extended Kalman filter basics, in: 19th International Conference Radioelektronika, 2009, pp. 119–122, <https://doi.org/10.1109/radioelek.2009.5158765>. IEEE Explore.
- [26] R.R. Labbe, Kalman and Bayesian Filters in Python, 2020. Available at<, [http://github.com/rllabbe/Kalman-and-Bayesian-Filters-in-Python](https://github.com/rllabbe/Kalman-and-Bayesian-Filters-in-Python).>.
- [27] N. Chopin, O. Papaspiliopoulos, An introduction to sequential Monte Carlo. Springer Series in Statistics, Springer International Publishing, 2020, <https://doi.org/10.1007/978-3-030-47845-2>.
- [28] A. Persily, L. Jonge, Carbon dioxide generation rates for building occupants, Indoor air 27 (5) (2017) 868–879, <https://doi.org/10.1111/ina.12383>.
- [29] NOAA, 2024. Trends in atmospheric carbon dioxide (CO₂). Available at: <<https://gml.noaa.gov/ccgg/trends/>>. Accessed on 10 December 2024.
- [30] A. Ortiz Perez, B. Bierer, L. Scholz, J. Wöllenstein, S. Palzer, A wireless gas sensor network to monitor indoor environmental quality in schools, Sensors 18 (12) (2018) 4345, <https://doi.org/10.3390/s18124345>.
- [31] H. Choi, J. Lee, Y. Yi, H. Na, K. Kang, T. Kim, Deep vision-based occupancy counting: experimental performance evaluation and implementation of ventilation control, Build. Env. 223 (2022) 109496, <https://doi.org/10.1016/j.buildenv.2022.109496>.