



Lawrence Berkeley National Laboratory

An Inverse Approach to Solving Zone Air Infiltration Rate and People Count using Indoor Environmental Sensor Data

Han Li¹, Tianzhen Hong¹, Marina Sofos²

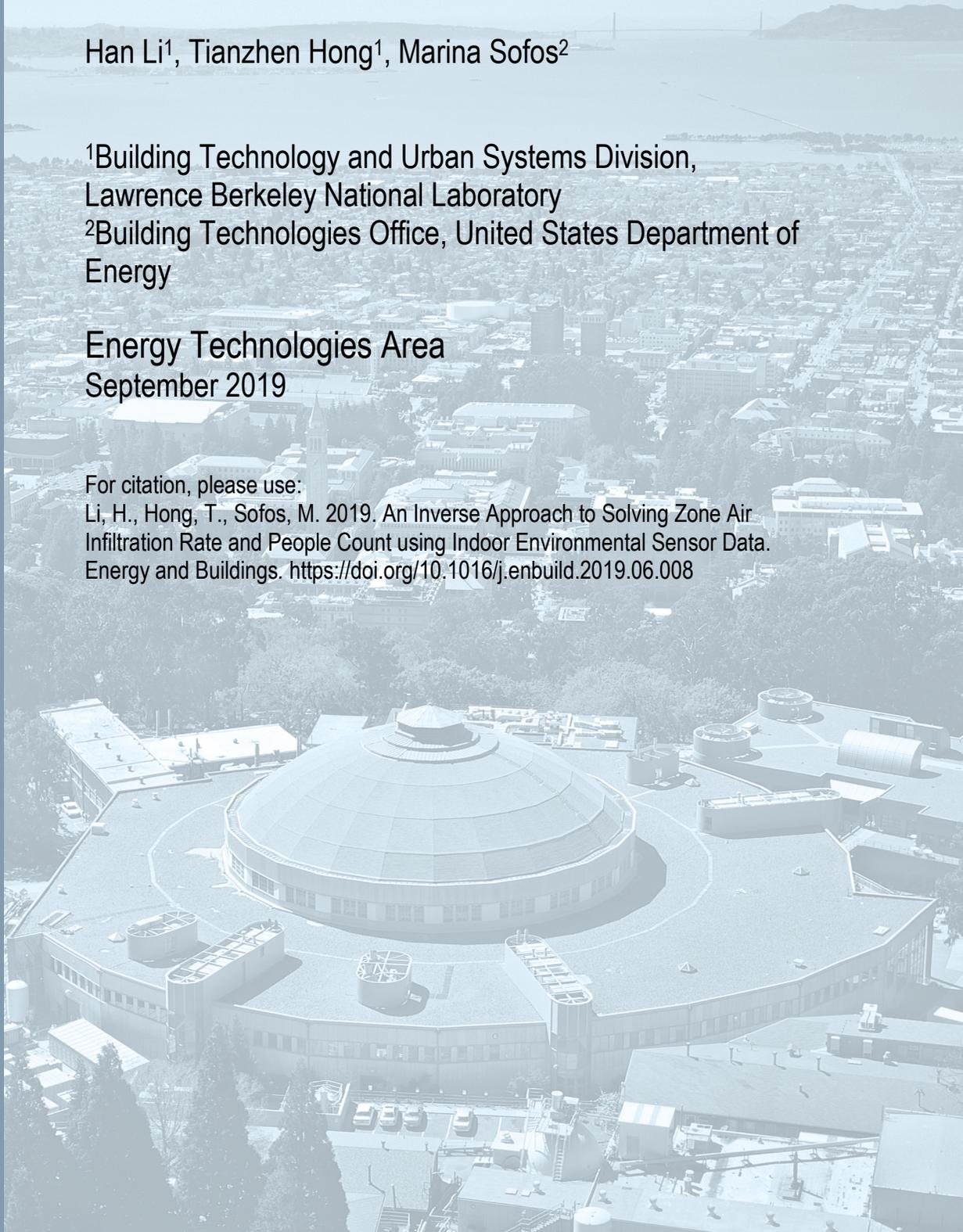
¹Building Technology and Urban Systems Division,
Lawrence Berkeley National Laboratory

²Building Technologies Office, United States Department of
Energy

Energy Technologies Area
September 2019

For citation, please use:

Li, H., Hong, T., Sofos, M. 2019. An Inverse Approach to Solving Zone Air Infiltration Rate and People Count using Indoor Environmental Sensor Data. Energy and Buildings. <https://doi.org/10.1016/j.enbuild.2019.06.008>



Disclaimer:

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.

An Inverse Approach to Solving Zone Air Infiltration Rate and People Count using Indoor Environmental Sensor Data

Han Li¹, Tianzhen Hong^{1*}, Marina Sofos²

¹Building Technology and Urban Systems Division, Lawrence Berkeley National Laboratory

²Building Technologies Office, United States Department of Energy

*Corresponding author: T. Hong. thong@lbl.gov, 1(510)4867082

Abstract

Physics-based simulation of energy use in buildings is widely used in building design and performance rating, controls design and operations. However, various challenges exist in the modeling process. Model parameters such as people count and air infiltration rate are usually highly uncertain, yet they have significant impacts on the simulation accuracy. With the increasing availability and affordability of sensors and meters in buildings, a large amount of measured data has been collected including indoor environmental parameters, such as room air dry-bulb temperature, humidity ratio, and CO₂ concentration levels. Fusing these sensor data with traditional energy modeling poses new opportunities to improve simulation accuracy. This study develops a set of physics-based inverse algorithms which can solve the highly uncertain and hard-to-measure building parameters such as zone-level people count and air infiltration rate. A simulation-based case study is conducted to verify the inverse algorithms implemented in EnergyPlus covering various sensor measurement scenarios and different modeling use cases. The developed inverse models can solve the zone people count and air infiltration at sub-hourly resolution using the measured zone air temperature, humidity and/or CO₂ concentration given other easy-to-measure model parameters are known.

Keywords: Inverse problems; sensor data; EnergyPlus; zone air parameters; infiltration; people count

Nomenclature

C_z – currnt zone air sensible heat capacity multiplier

T_z – current zone air temperature [K]

\dot{Q}_i – ith internal sensible heat gain rate of the current zone [W]

h_i – convective heat transfer coefficient of ith internal surface of the zone [$W/(m^2 \cdot K)$]

A_i – area of ith internal surface of the zone [m^2]

T_{si} – temperature of ith internal surface of the zone [K]

\dot{m}_{i_zone} – mass flow rate of air from ith nearby zone [kg/s]

C_{pi} – specific heat capacity of the air from ith nearby zone [J/kg]

T_{zi} – temperature of the air from ith nearby zone [K]

\dot{m}_{inf} – infiltration air mass flow rate [kg/s]

T_∞ – outdoor air temperature [K]

\dot{Q}_{sys} – sensible heat transfer rate due to HVAC system supply [W]

T_{sup} – HVAC system supply air temperature [K]

C_w – currnt zone air humidity capacity multiplier

W_z – currnt zone air humidity ratio [kg_{water}/kg_{dry_air}]

$m_{w,i}$ – ith internal moisture gain of the current zone [kg/s]

h_{mi} – moisture transfer coefficient of ith internal surface [m/s]

W_{si} – humidity ratio of the air near the ith surface of the zone [kg_{water}/kg_{dry_air}]

\dot{m}_{i_zone} – mass flow rate of the air from ith nearby zone [kg/s]

W_{zi} – humidity ratio of the air from the ith nearby zone [kg_{water}/kg_{dry_air}]

W_∞ – outdoor air humidity ratio [kg_{water}/kg_{dry_air}]

\dot{m}_{sys} – HVAC system supply air mass flow rate [kg/s]

W_{sup} – HVAC system supply air humidity ratio [kg_{water}/kg_{dry_air}]

C_{CO_2} – currnt zone air CO_2 capacity multiplier

C_z – currnt zone air CO_2 concentration [ppm]

$m_{CO_2,i}$ – sum of scheduled internal carbon dioxide loads of the current zone [$pm \cdot kg/s$]

C_{zi} – ith nearby zone air CO_2 concentration [ppm]

C_∞ – outdoor air CO_2 concentration [ppm]

C_{sup} – HVAC system supply air CO_2 concentration [ppm]

$\dot{Q}_{except_people,i}$ – internal sensible heat gain rate from ith sources except people [W]

\dot{Q}_{single} – sensible heat dissipation rate of a single person [W]

$\dot{m}_{w_except_people,i}$ – internal moisture gain rate from ith sources except people [kg/s]

m_{w_single} – moisture dissipation rate of a single person [kg/s]

$\dot{m}_{CO_2_except_people_i}$ – internal CO_2 gain rate from i th sources except people [$pm \cdot kg/s$]

$m_{CO_2_single}$ – CO_2 dissipation rate of a single person [$pm \cdot kg/s$]

Superscripts

X^t – the value of X at current timestamp

$X^{t-\delta t}$ – the value of X at one timestep before current timestamp

$X^{t-2\delta t}$ – the value of X at two timestep before current timestamp

$X^{t-3\delta t}$ – the value of X at three timestep before current timestamp

1. Introduction

1.1 Building energy modeling

Building energy modeling plays a critical role in researches and applications such as model predictive controls (MPC) [1] and operations [2], [3] and building energy retrofit analyses [4]. Common building energy modeling approaches include physics-based approaches, aka white-box approaches, and data-driven approaches, aka the black-box approaches.

Physics-based or forward modeling approaches explicitly model the interactions among weather conditions, building geometry, envelope, service systems, occupants, control strategies, and energy performance. The physics-based models can be further classified into reduced order models [5], [6] and dynamic models [7], [8] based on complexity. A reduced order model is usually a set of resistor–capacitor (RC) networks and is more computationally efficient than the detailed dynamic models. Therefore it is often used in situations when a short simulation time is critical, such as MPC [9]. On the other hand, detailed dynamic models describe the energy flows among building energy systems with physics laws and are solved with differential equations to provide more accurate results. The dynamic models are often more time consuming to build and solve [2]. With decades of research and development in the building energy modeling field, powerful tools, such as EnergyPlus, TRNSYS [10], ESP-r [11] have been developed and improved to model detailed buildings and systems with complex occupant behaviors and control settings. The increasing computational power also makes building energy modeling more and more widely adopted. However, detailed physics-based building energy models can be very sensitive to model assumptions. In reality, the discrepancies between design and actual building characteristics, the simplification of the building geometry, occupancy and system operation schedules, and the errors in computation can all negatively impact the model's accuracy of reflecting the real situations. Therefore, researchers have focused on calibrating the building energy models to match the measurements, so that the model can be used for predictive controls or evaluation of optimization strategies [12], [13].

Data-driven approaches gained lots of attentions in recent years. They often require a large amount of measurement data to train a reliable model that can reasonably represent the building's energy performance or other characteristics under different operation conditions. A list of literature reviewed the data-driven building energy prediction approaches [2], [14]–[16]. The main steps of data-driven approaches include data collection, data pre-processing, model training, and model testing [14]. A bunch of factors can affect the accuracy and scalability of the model built purely on data. First of all, the data collection process can be challenging. Some important sensor and meter data needed by the data-driven approaches may not be available in every building. Data quality is another issue; data from different building systems usually have different measurement periods and temporal resolutions. Missing data is always a main barrier to train and test models. Moreover, the measurements may not cover all the operation schemes, leading to a lack of full coverage of real operations by the trained models.

1.2 Measurements of uncertain model inputs

Among the building energy model inputs, air infiltration and people count are two variables that are highly uncertain and hard to measure directly at the zone level. Studies have shown that they have significant impacts on the energy simulation results [13], [17]–[19]. In most cases, air infiltration and people count are set as fixed schedules, which do not reflect the dynamic reality. Various methods have been developed to directly measure or indirectly calculate the zone air infiltration rates and people counts.

For the air infiltration rate, the most commonly used methods are tracer gas method and blower door method. Tracer gas method has been widely used to measure the infiltration rates in buildings since the 1980s. There are three categories of the method – dilution, constant injection, and constant concentration [20]. The fundamental of this method is the mass conservation of the tracer gas. The air infiltration rates can be calculated by monitoring

the relationship of tracer gas injection and the concentration change [21], [22]. The blower door method is another widely used method to measure the air airtightness of buildings. This method is also known as the fan pressurization method, which employs a large door-mounted fan to blow air into the building to quantify the air infiltration rate at a certain indoor-outdoor pressure difference [23]. Both tracer gas and blower door methods have been validated and used in enormous research and industrial application conditions. However, those methods have limitations such as the requirements of special devices, the disruptions to occupants, and potentially high time and labor costs. Some novel methods have been proposed in recent years to avoid those drawbacks while measuring the air infiltration rates. Examples are CFD-based approaches which use infrared images along with indoor and outdoor air parameter measurements and fluid mechanics analysis to quantify the air infiltration rate and pin-point the location of the air leakage [24]. However, the CFD-based methods usually need expertise in the geometry modeling, meshing, and simulation assumptions, which is hard to scale up.

Occupant behavior is a critical input in building performance modeling. The high uncertainty of occupants' presence and behavior have significant impacts on building energy modeling [25]. There is a wide spectrum of studies in detecting occupancy in buildings. The common methods of direct occupancy detection include (1) motion sensors, (2) vision-based technologies, and (3) radio-frequency localization technologies [26]. Some of the limitations of those technologies include the high cost and maintenance effort of the sensors, low accuracy in shared spaces, and privacy concerns. In recent years, some methods have been proposed to infer the occupancy and people count with advanced analytical and machine learning approaches. Wang et al. [27] developed a method to predict occupancy with fused environmental sensing and Wi-Fi sensing data using machine learning techniques. Candanedo et al. [28] developed a statistical learning model with light, temperature, humidity and CO₂ measurements to detect occupancy. Those methods showed good agreements with the ground truth. But one limitation is they require dedicated data processing and feature engineering to train reliable models.

1.3 Solving hard-to-measure parameters with inverse models

Decades of effort have been put into the development and refinement of physics-based building energy simulation tools. The integrated simulation engine EnergyPlus [29] has been through numerous tests and validations. It is now widely used in both research and commercial applications. At the same time, the cost of environmental sensing in buildings has declined. Many modern buildings are equipped with monitoring and control systems that can easily measure the HVAC system supply air and zone level air temperature, humidity, and CO₂ concentration. Traditionally, building energy simulation is used to predict the building's energy consumption and environmental parameters. However, backed by the well-tested physical principles, the building models can theoretically solve unknown parameters with reasonable model assumptions and accurate environmental measurements. Lee and Hong [30] proposed an inverse modeling approach, which uses the measured zone air temperature as the model inputs and solves the zone thermal mass or air infiltration rate. The thermal mass and air infiltration rate solved by the proposed method is implemented in EnergyPlus and validated with field measurements [30] under the free-floating condition.

In this study, a set of new inverse modeling algorithms are developed. The algorithms are based on the air sensible heat, humidity, or CO₂ conservation equations. Section 2 describes the theoretical fundamentals of the inverse balance equations. Section 3 presents a simulation-based case study, which shows the application of the inverse modeling algorithms with different availability of environmental measurements and model assumptions. The interpretation of the results is also discussed. Section 4 summarizes the case study results and discusses the pros and cons of the proposed inverse modeling algorithms. Section 5 gives the conclusions.

2. Methodology

The methodology of this study consist of three parts – (1) derivation of the inverse models, (2) implementations in EnergyPlus, and (3) verification case study. **Figure 1** shows the overall methodology map.

1. Derivation of Inverse Models

Invert the three balance equations:

$$\rho_{air} V_z C_T \frac{dT_z}{dt} = \sum_i Q_i + Q_{occ} + \sum_i h_i A_i \rho_{air} (T_{si} - T_z^t) + \sum_i \dot{m}_i C_p (T_{zi} - T_z^t) + \dot{m}_{inf} C_p (T_o - T_z^t) + \sum_i Q_{sys}$$

$$\rho_{air} V_z C_W \frac{dW_z}{dt} = \sum_i M_{W,in} + M_{W,occ} + \sum_i h_{mi} A_i \rho_{air} (W_{si} - W_z^t) + \sum_i \dot{m}_i (W_{zi} - W_z^t) + \dot{m}_{inf} (W_o - W_z^t) + \sum_i M_{W,sys}$$

$$\rho_{air} V_z C_{CO_2} \frac{dC_z}{dt} = \sum_i M_{CO_2,in} + M_{CO_2,occ} + \sum_i m_i (C_{zi} - C_z^t) + \dot{m}_{inf} (C_o - C_z^t) + \sum_i M_{CO_2,sys}$$

to solve:

Infiltration rate or Number of Occupants



2. Implementation in EnergyPlus



New feature - HybridModel:Zone

```
HybridModel:Zone,
A1 , \field Name
A2 , \field Zone Name
A3 , \field Calculate Zone Internal Thermal Mass
A4 , \field Calculate Zone Air Infiltration Rate
A5 , \field Calculate Zone People Count
A6 , \field Zone Measured Air Temperature Schedule Name
A7 , \field Zone Measured Air Humidity Ratio Schedule Name
A8 , \field Zone Measured Air CO2 Concentration Schedule Name
A9 , \field Zone Input People Activity Schedule Name
A10 , \field Zone Input People Sensible Heat Fraction Schedule Name
A11 , \field Zone Input People Radiant Heat Fraction Schedule Name
A12 , \field Zone Input People CO2 Generation Rate Schedule Name
A13 , \field Zone Input Supply Air Temperature Schedule Name
A14 , \field Zone Input Supply Air Mass Flow Rate Schedule Name
A15 , \field Zone Input Supply Air Humidity Ratio Schedule Name
A16 , \field Zone Input Supply Air CO2 Concentration Schedule Name
N1 , \field Begin Month
N2 , \field Begin Day of Month
N3 , \field End Month
N4 , \field End Day of Month
```

New output variables

```
Output:Variable,
*,
Zone Infiltration Hybrid Model Air Change Rate,
Timestep;

Output:Variable,
*,
Zone Hybrid Model People Count,
Timestep;
```



3. Verification Case Study

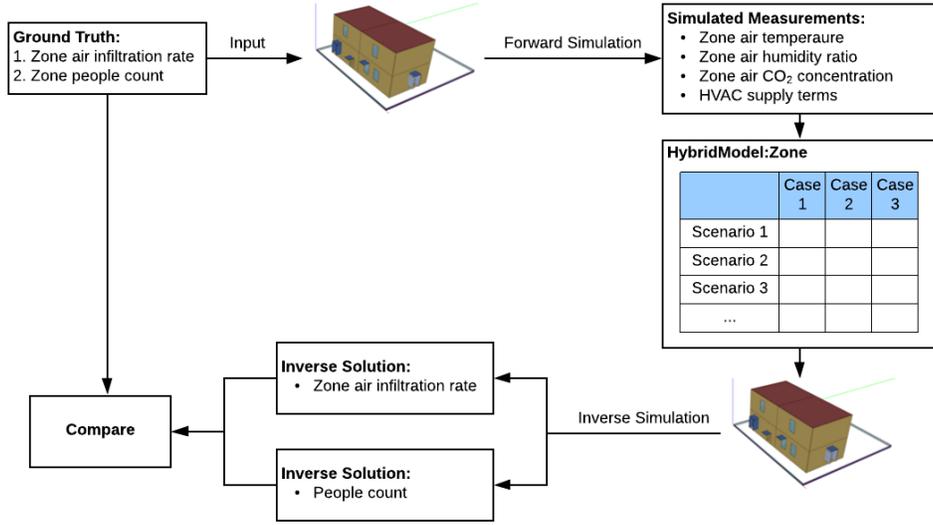


Figure 1. Overall Methodology

2.1 The zone air balance equations

The physics-based zone air heat, moisture, and contaminant equations [29] serve as the basis of the inverse modeling algorithms. The forward balance equations take into account the effect of internal heat gains (e.g., lighting system, electrical equipment, people, etc.), heat/mass exchanges with surfaces, connected zone air, outdoor air infiltration, as well as HVAC system supply air. The relationship between zone air sensible heat change and heat transfers from various sources can be expressed as the following:

$$C_z \frac{dT_z}{dt} = \sum_{i=1}^{N_{sl}} \dot{Q}_i + \sum_{i=1}^{N_{surfaces}} h_i A_i (T_{si} - T_z) + \sum_{i=1}^{N_{zones}} \dot{m}_{i_zone} C_{pi} (T_{zi} - T_z) + \dot{m}_{inf} C_p (T_{\infty} - T_z) + \dot{Q}_{sys} \quad (1)$$

Where C_z is the zone air total sensible heat capacity multiplier, $\sum_{i=1}^{N_{sl}} \dot{Q}_i$ is the sum of convective internal heat gains, $\sum_{i=1}^{N_{surfaces}} h_i A_i (T_{si} - T_z)$ is sum of convective heat gains from interior surfaces, $\dot{m}_{inf} C_p (T_{\infty} - T_z)$ is the convective heat gain from outdoor air infiltration, and \dot{Q}_{sys} is the convective heat transfer from the HVAC systems.

Similarly, the zone air moisture balance equation can be expressed as:

$$C_w \frac{dW_z}{dt} = \sum_{i=1}^{N_{sl}} \dot{m}_{w,i} + \sum_{i=1}^{N_{surfaces}} A_i h_{mi} \rho_{air} (W_{si} - W_z) + \sum_{i=1}^{N_{zones}} \dot{m}_{i_zone} (W_{zi} - W_z) + \dot{m}_{inf} (W_{\infty} - W_z) + \dot{m}_{sys} (W_{sup} - W_z) \quad (2)$$

Where C_w is the zone air moisture capacity multiplier, $\sum_{i=1}^{N_{sl}} \dot{m}_{w,i}$ is the sum of internal moisture gains, $\sum_{i=1}^{N_{surfaces}} A_i h_{mi} \rho_{air} (W_{si} - W_z)$ is the sum of moisture gains from the interior surfaces, $\sum_{i=1}^{N_{zones}} \dot{m}_{i_zone} (W_{zi} - W_z)$ is the sum of moisture gains from the connected zones, $\dot{m}_{inf} (W_{\infty} - W_z)$ is the moisture gain from outdoor air infiltration, and $\dot{m}_{sys} (W_{sup} - W_z)$ is the moisture gain from the HVAC systems.

And the zone air CO₂ mass balance equation can be expressed as:

$$C_{CO_2} \frac{dC_z}{dt} = \sum_{i=1}^{N_{sl}} \dot{m}_{CO_2,i} + \sum_{i=1}^{N_{zones}} \dot{m}_{i_zone} (C_{zi} - C_z) + \dot{m}_{inf} (C_{\infty} - C_z) + \dot{m}_{sys} (C_{sup} - C_z) \quad (3)$$

Where C_{CO_2} is the zone air CO₂ capacity multiplier, $\sum_{i=1}^{N_{sl}} \dot{m}_{CO_2,i}$ is the sum of internal CO₂ gains, $\sum_{i=1}^{N_{zones}} \dot{m}_{i_zone} (C_{zi} - C_z)$ is the CO₂ gains from connected zones, $\dot{m}_{inf} (C_{\infty} - C_z)$ is the CO₂ gains from outdoor air infiltration, and $\dot{m}_{sys} (C_{sup} - C_z)$ is the CO₂ gains from the HVAC systems.

2.2 The inverse modeling algorithms

The inverse modeling algorithms are developed to solve the zone air balance equations in their ordinary differential format. In this study, EnergyPlus is used as the simulation engine which implements these inverse models. But the methodology is generic and can be applied to other physics-based simulation engines. Depending on the model assumptions and available measured zone parameters, the inverse modeling algorithms can be used to solve different unknown parameters such as people count, air infiltration rate, zone internal thermal mass, and HVAC supply airflow rate. Zone level people count and air infiltration are two influential model parameters yet hard to measure. Thus, this study implemented the inverse algorithms in EnergyPlus to solve people count and air infiltration rate using easily measurable zone parameters such as air temperature, humidity and/or CO₂ concentration. **Figure 2** shows the solution of those two unknown parameters with three indoor environmental parameter measurements under various scenarios. The system supply terms can be ignored when the HVAC is off since there is no sensible heat/moisture/CO₂ transfer between the HVAC system supply air and zone air. But they must be provided when the HVAC is on during the measurements.

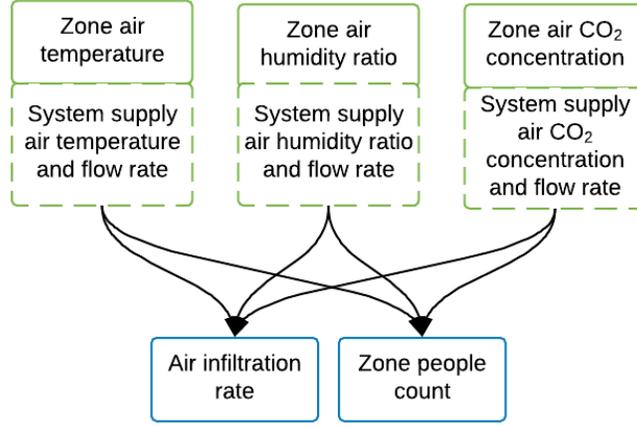


Figure 2. Relationship between measured parameters and inversely solvable unknown parameters

The ordinary differential equations (1), (2), and (3) can be solved with the finite difference approach which requires time-series measurements of zone air temperature, humidity ratio, or CO₂ concentration. With the smart sensor network, the measurements are easily accessible in modern buildings. EnergyPlus uses third-order backward approximation [8] to solve dry-bulb temperature, humidity ratio, or CO₂ concentration with the balance equations (1), (2), and (3) in its zone predictor-corrector [8] solution. It was proved to provide sufficient accuracy. Therefore, the proposed inverse algorithms also adopt the third-order backward approximation approach. With the third-order backward approximation, equations (1), (2), and (3) can be inversely re-written as (4), (5), and (6), respectively:

$$C_z \frac{\frac{11}{6}T_z^t - 3T_z^{t-\delta t} + \frac{3}{2}T_z^{t-2\delta t} - \frac{1}{3}T_z^{t-3\delta t}}{\delta t} = \sum_{i=1}^{N_{sl}} \dot{Q}_i + \sum_{i=1}^{N_{surfaces}} h_i A_i (T_{si}^t - T_z^t) + \sum_{i=1}^{N_{zones}} \dot{m}_{i,zone} C_p (T_{zi}^t - T_z^t) + \dot{m}_{inf} C_p (T_{\infty}^t - T_z^t) + \dot{m}_{sys} C_p (T_{sup}^t - T_z^t) \quad (4)$$

$$C_{wz} \frac{\frac{11}{6}W_z^t - 3W_z^{t-\delta t} + \frac{3}{2}W_z^{t-2\delta t} - \frac{1}{3}W_z^{t-3\delta t}}{\delta t} = \sum_{i=1}^{N_{sl}} m_{CO_2,i} + \sum_{i=1}^{N_{surfaces}} A_i h_{mi} \rho_{air} (W_{si}^t - W_z^t) + \sum_{i=1}^{N_{zones}} \dot{m}_{i,zone} (W_{zi}^t - W_z^t) + \dot{m}_{inf} (W_{\infty}^t - W_z^t) + \dot{m}_{sys} (W_{sup}^t - W_z^t) \quad (5)$$

$$C_{CO_2} \frac{\frac{11}{6}C_z^t - 3C_z^{t-\delta t} + \frac{3}{2}C_z^{t-2\delta t} - \frac{1}{3}C_z^{t-3\delta t}}{\delta t} = \sum_{i=1}^{N_{sl}} m_{w,i} + \sum_{i=1}^{N_{zones}} \dot{m}_{i,zone} (C_{zi}^t - C_z^t) + \dot{m}_{inf} (C_{\infty}^t - C_z^t) + \dot{m}_{sys} (C_{sup}^t - C_z^t) \quad (6)$$

The superscript notations of T_z , W_z , and C_z represent the timestamp of the measurements. For example, T_z^t is the measured zone air dry-bulb temperature at the current timestamp, while $T_z^{t-\delta t}$ is the measured air dry-bulb temperature at one time step earlier than the current timestamp.

From Equation (4), (5), and (6), it can be inferred that it is critical to model other terms accurately to use the inverse algorithms, because the inversely solved air infiltration and people count will be overfitted if other terms in the balance equations are highly uncertain. The inverse modeling algorithms work under the following conditions (the exact conditions vary depending on which parameter is used as input):

- 1) The zone air sensible thermal mass, total humidity capacity, or total CO₂ concentration capacity is known and fixed.
- 2) If the system supply air temperature, humidity ratio, or CO₂ concentration is not measured, the inverse algorithms are only valid under the free-floating (HVAC system is off) mode.
- 3) The zone internal sensible heat gains, moisture gains, or CO₂ gains are modeled at a reasonable accuracy.
- 4) The inter-zone air exchange is modeled at a reasonable accuracy.
- 5) The convective heat, moisture, or CO₂ transfer between zone surfaces and zone air are modeled at a reasonable accuracy.
- 6) The sensible heat generation rate, moisture and CO₂ dissipation rate of a single person are known.

2.2.1 Inverse modeling algorithms to solve zone air infiltration

With the measured zone air parameters, the air infiltration mass flow rate \dot{m}_{inf} can be solved with Equation (7), (8), or (9) as shown below. For example, Equation (7) calculates the sensible heat gain (or loss) rate from air infiltration with the zone air sensible heat balance equation, and then solves the infiltration mass flow rate with the infiltration heat capacity and outdoor-indoor air temperature difference. If the HVAC is on during the measurements, the system supply air mass flow rate and supply air temperature also need to be measured.

$$\dot{m}_{inf} = \frac{C_z \frac{11}{6} T_z^t - 3T_z^{t-\delta t} + \frac{3}{2} T_z^{t-2\delta t} - \frac{1}{3} T_z^{t-3\delta t}}{\delta t} - \frac{[\sum_{i=1}^{N_{sl}} \dot{Q}_i + \sum_{i=1}^{N_{surfaces}} h_i A_i (T_{sl}^t - T_z^t) + \sum_{i=1}^{N_{zones}} \dot{m}_{i_zone} C_p (T_{zi}^t - T_z^t) + \dot{m}_{sys} C_p (T_{sup}^t - T_z^t)]}{C_p (T_{\infty}^t - T_z^t)} \quad (7)$$

$$\dot{m}_{inf} = \frac{C_{wz} \frac{11}{6} W_z^t - 3W_z^{t-\delta t} + \frac{3}{2} W_z^{t-2\delta t} - \frac{1}{3} W_z^{t-3\delta t}}{\delta t} - \frac{[\sum_{i=1}^{N_{sl}} m_{w,i} + \sum_{i=1}^{N_{surfaces}} A_i h_{mi} \rho_{air} (W_{sl}^t - W_z^t) + \sum_{i=1}^{N_{zones}} \dot{m}_{i_zone} (W_{zi}^t - W_z^t) + \dot{m}_{sys} (W_{sup}^t - W_z^t)]}{W_{\infty}^t - W_z^t} \quad (8)$$

$$\dot{m}_{inf} = \frac{C_{CO_2} \frac{11}{6} C_z^t - 3C_z^{t-\delta t} + \frac{3}{2} C_z^{t-2\delta t} - \frac{1}{3} C_z^{t-3\delta t}}{\delta t} - \frac{[\sum_{i=1}^{N_{sl}} m_{CO_2,i} + \sum_{i=1}^{N_{zones}} \dot{m}_{i_zone} (C_{zi}^t - C_z^t) + \dot{m}_{sys} (C_{sup}^t - C_z^t)]}{C_{\infty}^t - C_z^t} \quad (9)$$

2.2.2 Inverse modeling algorithms to solve zone people count

With the measured zone air parameters, the zone people count N_{occ} can be solved with the following pairs of equations. For instance, Equation (10) solves the zone total internal heat gain rate, $\sum_{i=1}^{N_{sl}} \dot{Q}_i$. Then, Equation (11) solves the number of occupants in the zone by dividing the total sensible heat gain rate from people, $\sum_{i=1}^{N_{sl}} \dot{Q}_i - \sum_{i=1}^{N_{sl}} \dot{Q}_{except_people_i}$, to the sensible heat generation rate of a single person, \dot{Q}_{single} . Similar to the algorithms solving air infiltration rate, the system supply air mass flow rate and supply air temperature need to be measured if the HVAC system is on. Equation (12) and (13) solve the people count with measured humidity ratio. Equation (14) and (15) solve the people count with measured CO₂ concentration.

$$\sum_{i=1}^{N_{sl}} \dot{Q}_i = C_z \frac{11}{6} T_z^t - 3T_z^{t-\delta t} + \frac{3}{2} T_z^{t-2\delta t} - \frac{1}{3} T_z^{t-3\delta t} - \left[\sum_{i=1}^{N_{surfaces}} h_i A_i (T_{sl}^t - T_z^t) + \sum_{i=1}^{N_{zones}} \dot{m}_{i_zone} C_p (T_{zi}^t - T_z^t) + \dot{m}_{sys} C_p (T_{sup}^t - T_z^t) \right] \quad (10)$$

$$N_{occ} = \frac{\sum_{i=1}^{N_{sl}} \dot{Q}_i - \sum_{i=1}^{N_{sl}} \dot{Q}_{except_people_i}}{\dot{Q}_{single}}$$

(11)

$$\sum_{i=1}^{N_{sl}} \dot{m}_{w,i} = C_{wz} \frac{\frac{11}{6} W_z^t - 3W_z^{t-\delta t} + \frac{3}{2} W_z^{t-2\delta t} - \frac{1}{3} W_z^{t-3\delta t}}{\delta t} - \left[\sum_{i=1}^{N_{surfaces}} A_i h_{mi} \rho_{air} (W_{si}^t - W_z^t) + \sum_{i=1}^{N_{zones}} \dot{m}_{i,zone} (W_{zi}^t - W_z^t) + \dot{m}_{sys} (W_{sup}^t - W_z^t) \right] \quad (12)$$

$$N_{occ} = \frac{\sum_{i=1}^{N_{sl}} \dot{m}_i - \sum_{i=1}^{N_{sl}} \dot{m}_{w,except_people,i}}{\dot{m}_{w_single}} \quad (13)$$

$$\sum_{i=1}^{N_{sl}} \dot{m}_{CO_2,i} = C_{CO_2} \frac{\frac{11}{6} C_z^t - 3C_z^{t-\delta t} + \frac{3}{2} C_z^{t-2\delta t} - \frac{1}{3} C_z^{t-3\delta t}}{\delta t} - \left[\sum_{i=1}^{N_{zones}} \dot{m}_{i,zone} (C_{zi}^t - C_z^t) + \dot{m}_{sys} (C_{sup}^t - C_z^t) \right] \quad (14)$$

$$N_{occ} = \frac{\sum_{i=1}^{N_{sl}} \dot{m}_i - \sum_{i=1}^{N_{sl}} \dot{m}_{CO_2,except_people,i}}{\dot{m}_{CO_2_single}} \quad (15)$$

2.3 Convergence

There can be many factors affecting the convergence when trying to solve the differential equation numerically with the third-order backward approximation. The most common issue is the overflow. The latest version of EnergyPlus code is written in C++. Just as any other language, it overflows when the result from an operation exceeds a certain range. For the inverse modeling algorithms, overflow can happen when calculating the air infiltration rate. For instance, the indoor-outdoor air temperature difference term ($T_{\infty}^t - T_z^t$) can be a very small number when the two temperatures are very close. Overflow will happen if the program tries to calculate the air infiltration rate by dividing the denominator of Equation (7) by $C_p(T_{\infty}^t - T_z^t)$. Therefore, conditional checks are needed when implementing the algorithm in the code. In this case, a threshold of 0.05 °C or greater temperature difference must be met to calculate the infiltration rate at one timestamp. Similarly, thresholds are added for the algorithms using humidity ratio and CO₂ concentration. In practice, the thresholds implemented in EnergyPlus routines don't have significant impacts on its ability to solve the unknown parameters in our tests.

In addition, EnergyPlus uses a zone predictor-corrector mechanism to calculate the heating or cooling needs of a zone on the HVAC system, and update the zone air parameters based on the calculated amount of heating or cooling the HVAC system provides to a zone. The uncertainties such as truncation errors in those predictor-corrector routines can cause an anomaly in the inverse modeling routine. Therefore, thresholds for infiltration and people count calculation are applied to the code. For infiltration, a valid value must be within the range of 0 to 10 air changes per hour. For people count, the lower bound is zero, and the upper-bound is the total possible internal heat/moisture/CO₂ gain divided by the heat/moisture/CO₂ generation rate.

3. Case Study

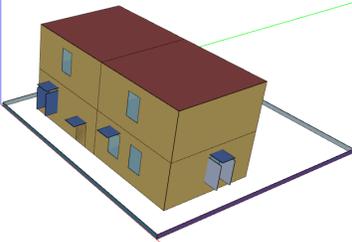
To verify that the inverse modeling algorithms are correctly implemented in EnergyPlus and to demonstrate the use of the new EnergyPlus feature, a simulation-based case study was conducted. This section presents model settings, solution scenarios, and results of the case study.

3.1 Model settings

An EnergyPlus building model is used in the case study. The model represents a two-story building with two zones on each floor with a 1600 m² total floor area. Three locations are considered to cover typical hot, cold, and mild

climate. There are two rounds of simulations. The first round is the forward simulation, where the air infiltration rate and people count are provided as model inputs. The forward simulation is used to generate the virtual measurements of the zone air and system supply air parameters. Then in the second round of simulations, the virtual measurements are provided as the inputs to the inverse modeling algorithms to solve zone air infiltration or people count. Since only one unknown variable can be solved at a time, the people count should be provided when solving the air infiltration, and vice versa. **Table 1** shows the model setting details.

Table 1. Model settings of the case study

Model settings	Forward simulation	Inverse simulation 1: solving air infiltration	Inverse simulation 2: solving people count
Purpose	Get the virtual measurements (i.e., zone air and system supply air parameters)	Use the virtual measurements to inversely solve air infiltration rates	Use the virtual measurements to inversely solve zone people counts
Building geometry sketch			
Locations	Chicago, Houston, San Francisco		
Interior lighting power density	9.69 W/m ²		
Electric equipment power density	6.78 W/m ²		
HVAC system type	Ideal air load system: the HVAC system can meet the space heating and cooling loads as long as they are below the system capacity.		
Air infiltration	Fixed schedule (ground truth)	NA	Fixed schedule
Occupancy density	10 m ² /person (ground truth)	10 m ² /person	NA

In the forward simulation, air infiltration is modeled with the maximum air change rate and a schedule of the fractions of the maximum value at different hours of a day. Similarly, the zone people count is modeled with the maximum number of people and a schedule indicating the fractions of the maximum number of people at different hours. The forward simulation uses the infiltration rate schedule from DOE prototype small office building [31]. People’s behavior and movements in real buildings are hard to predict, which affect the presence of people in building spaces. Chen et al. [32] developed an agent-based algorithm to simulate occupant movements using Markov-chain model. Based on the study, an application was developed. In this case study, a stochastic occupant schedule generated by the application is used as the ground truth to mimic the high uncertain people movements in real buildings. **Figure 3** and **Figure 4** shows the air infiltration and example people count schedule for a day.

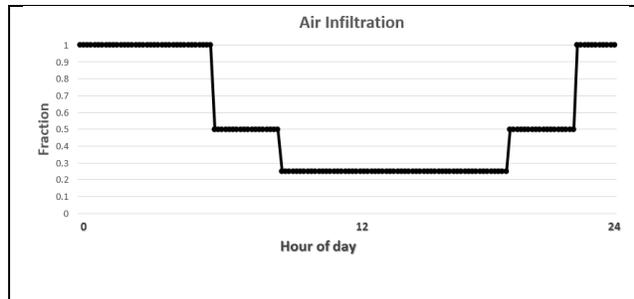


Figure 3. Air infiltration schedule

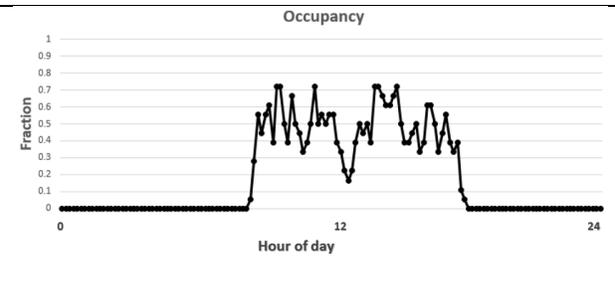


Figure 4. Occupancy schedule

3.2 Inverse solution scenarios

There can be different use cases and solution scenarios with the inverse modeling algorithms depending on which measured parameters and model details are available in the inverse simulation. Thus, experiments with different level of details of measurements and model assumptions are carried out in the case study.

- 1) The simplest use case is when the building's HVAC system is off, and the building is at free-floating mode during the zone air measurements and the HVAC system is not modeled in both the forward simulation and inverse simulation. This case is most suitable when limited measurements and limited building model details are available. However, it requires the building's HVAC system be turned off. For example, this case can be used to solve air infiltration rate when HVAC system is off during unoccupied hours.
- 2) A more complex use case is when the HVAC is on during the zone air measurements, but no HVAC detail is modeled in the inverse simulation. In this case, the HVAC system is not modeled in the forward simulation, but both zone air parameters and the HVAC system supply air parameters are measured and used in the inverse simulation. However, since HVAC is not modeled in the forward simulation, its effects on the zones interior surfaces are not accounted. This use case is most beneficial when the HVAC supply parameters can be easily measured, but the detailed system configurations are hard to be modeled (due to lack of information).
- 3) The most complicated use case is when the HVAC is on during the zone air measurements, and HVAC details are modeled in the inverse simulation. This case requires not only the measurements but also the detailed HVAC information for the inverse model. It is most beneficial when both measurements of the HVAC supply parameters and the modeling of HVAC details are achievable.

Figure 5 shows the different model inputs and measurements for the three use cases.

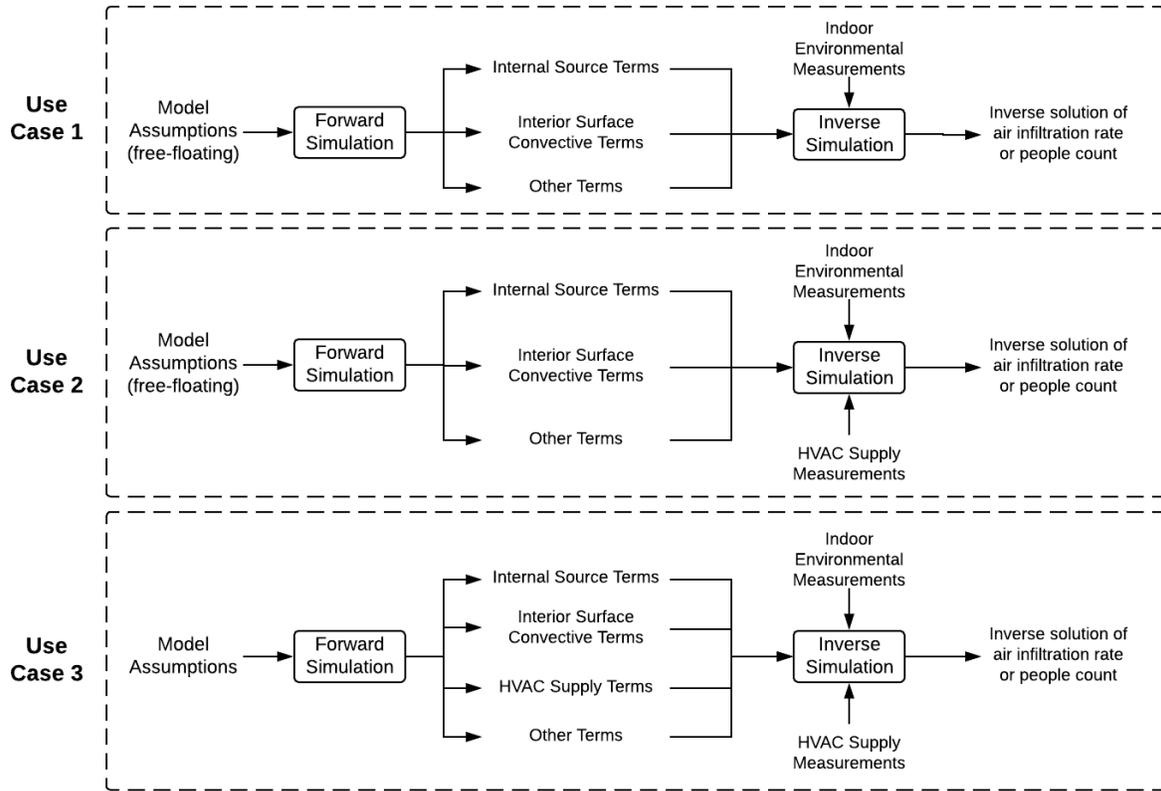


Figure 5. Air infiltration schedule

Table 2 shows the required measurements and model assumptions for different use cases and solution scenarios.

Table 2. Inverse solution use cases and scenarios

Use cases		Case 1			Case 2			Case 3		
Scenarios		S1	S2	S3	S4	S5	S6	S7	S8	S9
HVAC status during measurement		Off	Off	Off	On	On	On	On	On	On
HVAC is modeled		No	No	No	No	No	No	Yes	Yes	Yes
Climate zones		Chicago, Houston, San Francisco								
Measured Parameter(s)	zone air temperature	x			x			x		
	zone air humidity ratio		x			x			x	
	zone air CO ₂ concentration			x			x			x
	supply air temperature				x			x		
	supply air humidity ratio					x			x	
	supply air CO ₂ concentration						x			x
	supply air mass flow rate				x	x	x	x	x	x
Note		HVAC is off during measurements; no HVAC is modeled in the inverse simulation			HVAC is on during measurements; no HVAC is modeled in the inverse simulation			HVAC is on during measurements, HVAC is modeled in the inverse simulation		

3.3 Results

Based on the previous discussion, there are 216 combinations (2 unknown parameters x 4 zones x 3 locations x 3 measurements x 3 uses cases) in the case study. To illustrate the results, this section first presents time-series comparison examples between the ground truth and the inverse solutions. Then it presents the statistical metrics of the inverse solutions and summarizes the applicability of different use cases.

Time-series charts can help visually inspect the alignments between the inverse solution and the ground truth. **Figure 6** through **Figure 8** show the ground truth and the inverse solution of the air infiltration rate at one zone in the model for three use cases. The results from Chicago are selected since it covers hot summer and cold winter.

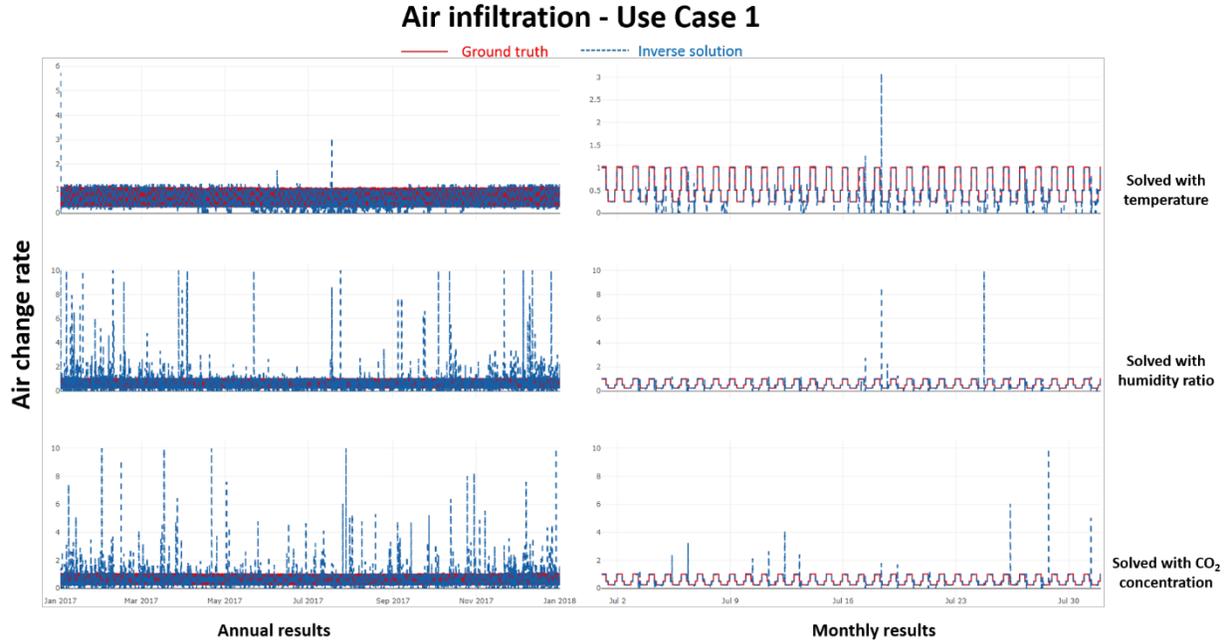


Figure 6. Use Case 1 time-series comparison of the inverse solution and the ground truth of air infiltration rates

Since there is more diversity in occupant count schedule, an annual comparison and a weekly comparison are used in the plots. **Figure 9** through **Figure 11** show the ground truth and the inverse solution of people count at one zone in the model for three use cases.

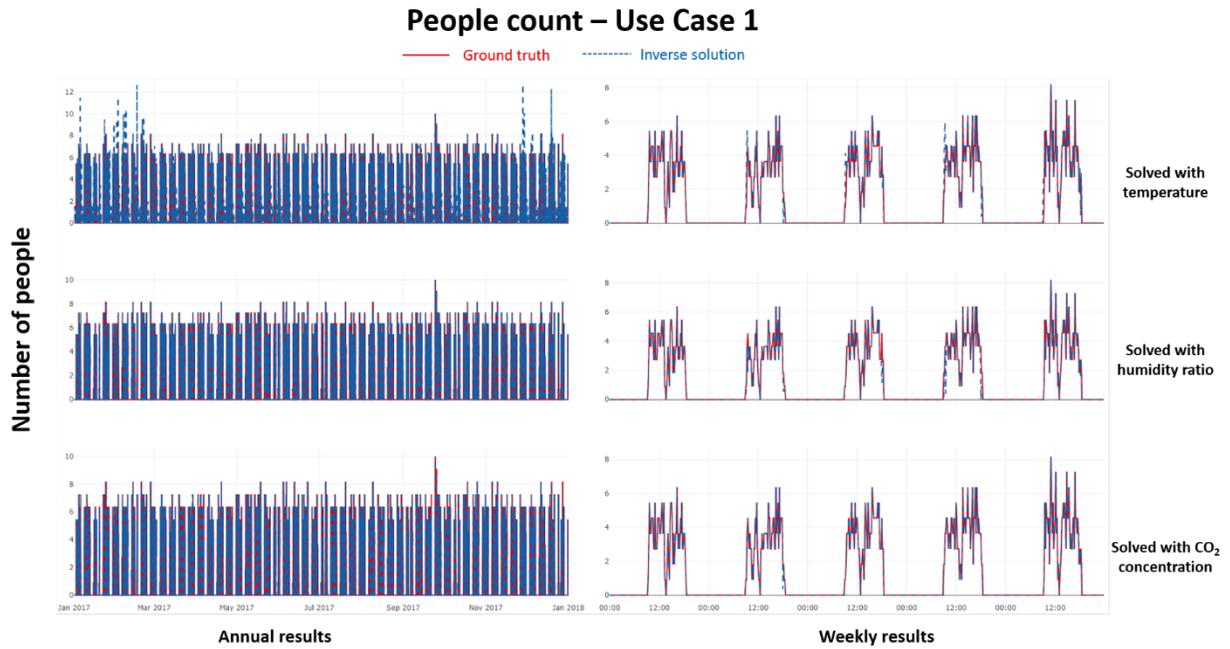


Figure 9. Use Case 1 time-series comparison of the inverse solution and the ground truth of people count

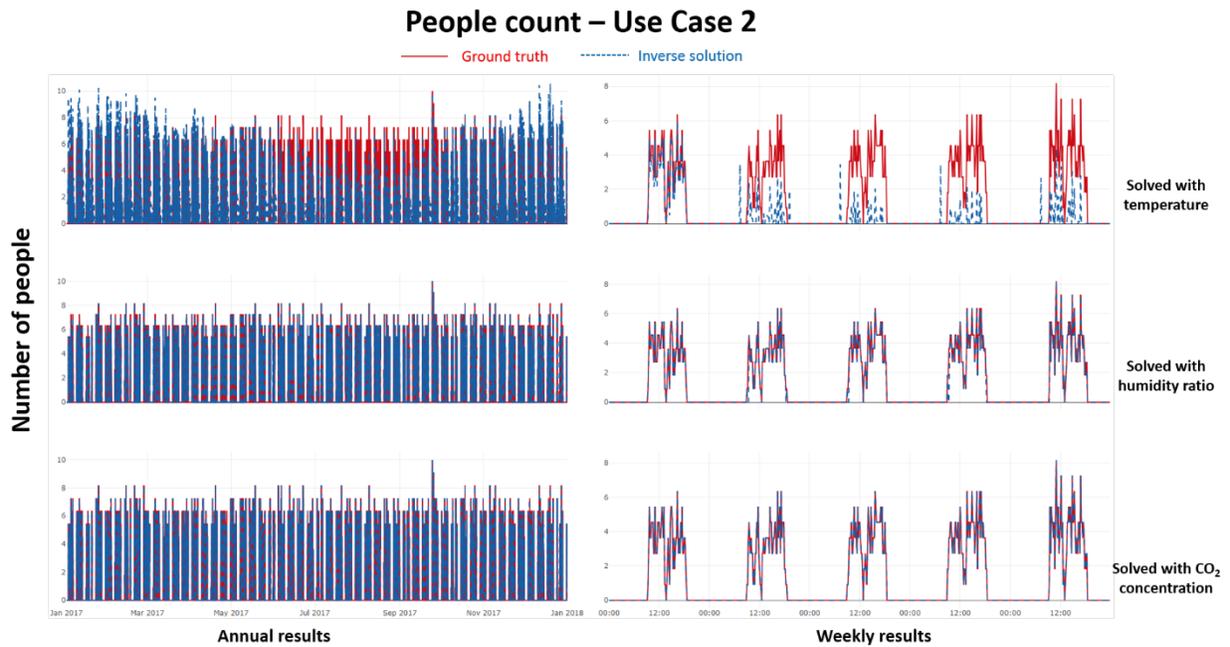


Figure 10. Use Case 2 time-series comparison of the inverse solution and the ground truth of people count

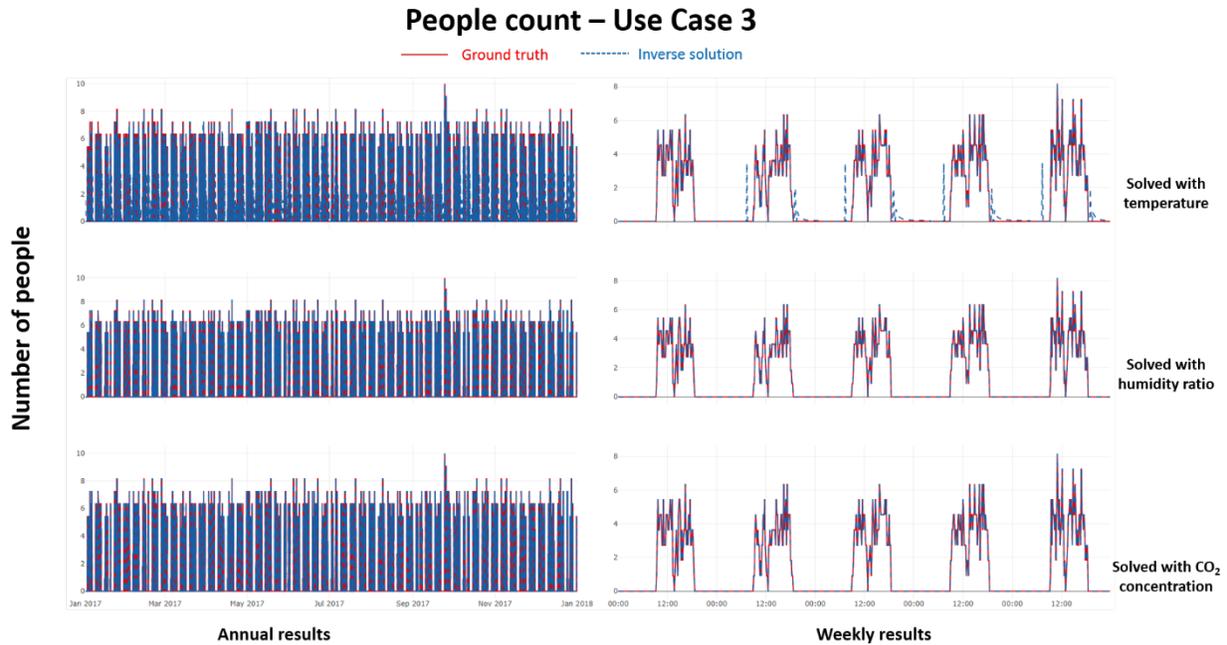


Figure 11. Use Case 3 time-series comparison of the inverse solution and the ground truth of people count

Similar to the inverse solutions of air infiltration rates, the accuracy of the inverse solutions vary by use case and measurements. However, there is an apparent discrepancy between the inverse solution and the ground truth for Case 2, when the measurement is zone air temperature. During the cooling season, the inverse solution of people count is smaller than the ground truth. The reason for the discrepancies will be discussed shortly.

The time-series comparisons between the ground truth and the inverse solutions give a snapshot of how the inverse algorithm work overall. Comparison of the probability density between the inverse solution and the ground truth can provide a statistical view of how the inverse modeling algorithms perform in solving the unknown air infiltration rate or people count. **Figure 12** and **Figure 13** show the probability density distributions of the ground truth and the inverse solutions of different use cases in three experimental locations for a single zone in the modeled building. The density violin plots (smoothed by the kernel density estimator) use the data aggregated from 10-minute time interval values for a whole year. The plots reflect the full distribution of the ground truth and inverse solutions. In the facet plot grid, each row corresponds to one use case (see details in **Table 2**) and each column corresponds to a location. There are four traces in each child plot – one ground truth and three solution scenarios with the measured air temperature, measured air humidity ratio, and measured CO₂ concentration, respectively.

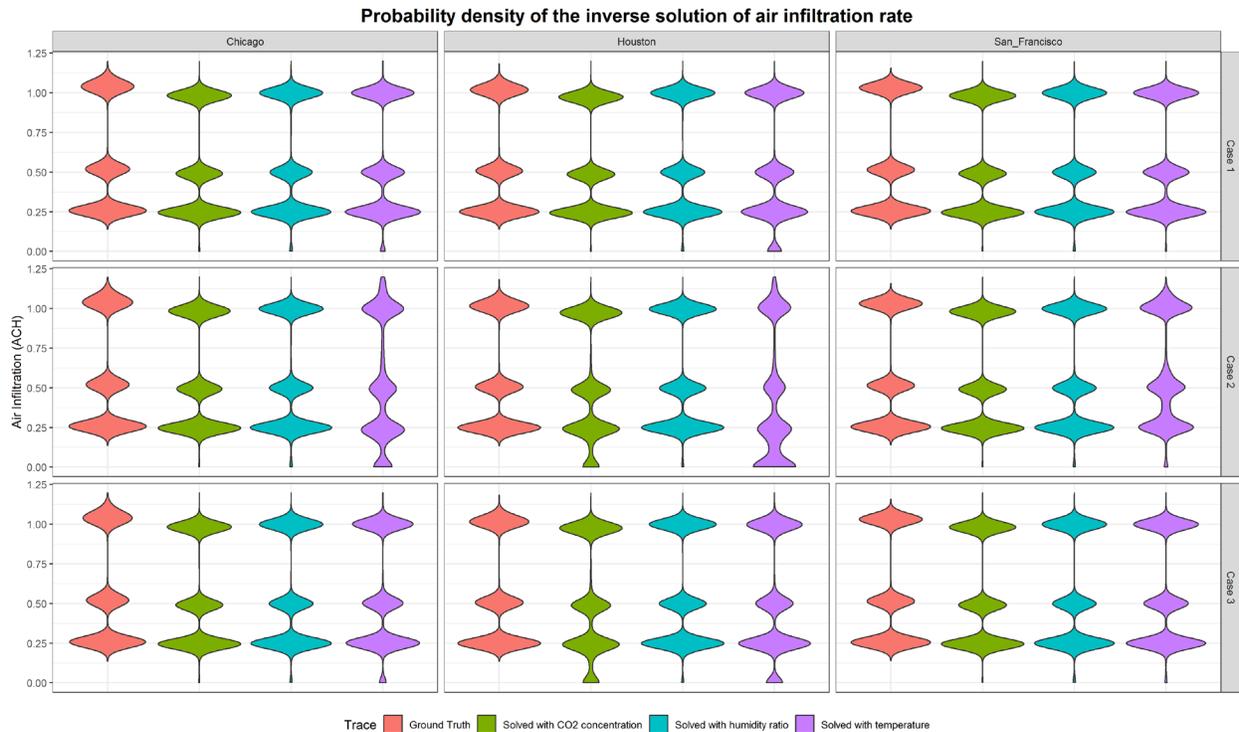


Figure 12. Probability density plots of the ground truth and the inverse solution of air infiltration rates

For example, in **Figure 12**, the sub-plot in row one and column one shows the probability density distribution of the ground truth of the air infiltration rate and the inverse solutions for Use Case 1 in Chicago. There are three bulks in the violin plot where each bulk reflects the local average of the value while the width reflects the frequency. In this case, the ground truth has three typical values (1, 0.5, and 0.25) of air infiltration rate as shown in **Figure 3**. The inverse solutions have similar violin plots with the ground truth, which means the solution matches the ground truth well. It can be seen from the figures that in general the probability density distribution of the inverse solution aligns well with the ground truth. There is one exception for the solution with Use Case 2 when solving with measured temperature (row 2, column 2 in **Figure 12**). In this case, the solution's probability density distribution shows there are many times when the solution differ from the ground truth, especially when the infiltration air change rate is below 0.25 (indicated by the large area at the bottom of the violin plot). The reasons for the poor performance of this case are discussed in the last two paragraphs of Section 4.

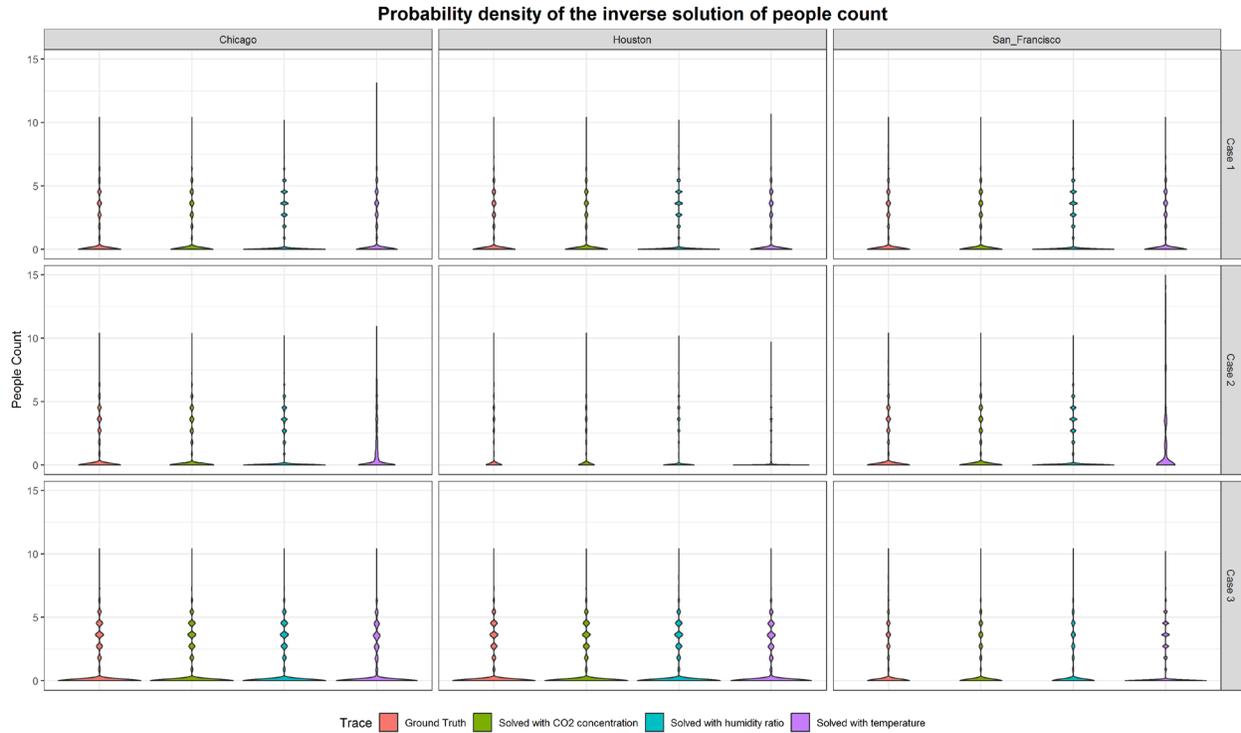


Figure 13. Probability density plots of the ground truth and the inverse solution of people count

The interpretation of **Figure 13** is similar to that of **Figure 12**. The beads-like violin plots show the probability density distributions of the ground truth and inverse solutions of the people count. Each “bead” in the plot represents a local average of the people count in the schedule. Like the air infiltration rate, the solutions of people count with temperate in Use Case 2 differ from the ground truth, which is also indicated in **Figure 10**.

Coefficient of Variance of the Root Mean Square Deviation CV(RMSD) is a commonly used index to quantify how well the predictions describe the variability of the ground truth. **Table 3** shows the CV(RMSD) between the inverse solutions and the ground truth. Smaller values of CV(RMSD) suggest better alignments between the inverse solution and the ground truth. The values are color-coded in the tables where green stands for a small value and red stands for a large value to better visualize the performance of different use cases and scenarios.

Table 3. CV(RMSD) of the inverse solutions

Location				Chicago				Houston				San Francisco			
Zone				Zone 1	Zone 2	Zone 3	Zone 4	Zone 1	Zone 2	Zone 3	Zone 4	Zone 1	Zone 2	Zone 3	Zone 4
Use case	Measured Parameter(s)	HVAC Status	HVAC model	CV(RMSD) of air infiltration solution											
Case 1	S1	Off	No	14.89	15.39	13.35	12.31	30.02	27.92	26.22	24.57	10.98	11.47	9.42	9.22
	S2	Off	No	63.72	64.33	42.71	44.14	60.85	60.44	37.41	39.46	65.32	70.05	37.55	33.81
	S3	Off	No	24.03	29.92	35.08	39.58	20.86	19.73	31.11	37.34	22.42	24.37	33.43	38.15
Case 2	S4	On	No	51.48	50.19	51.70	49.63	74.90	72.87	74.59	72.66	22.33	20.69	22.24	20.29
	S5	On	No	35.03	38.09	29.56	26.26	32.42	29.52	29.02	26.55	46.24	49.67	30.99	30.23

	S6	On	No	26.9 1	33.8 4	28.87	29.9 5	33.6 2	29.0 3	45.9 6	51.0 8	27.78	31.20	30.28	32.74
Case 3	S7	On	yes	31.7 6	35.3 1	31.93	34.8 6	52.9 3	54.0 8	52.8 8	54.7 1	13.13	13.44	12.67	11.90
	S8	On	yes	34.3 9	36.2 7	29.23	25.0 8	32.0 0	31.5 7	28.0 3	27.4 0	43.10	48.01	32.04	31.93
	S9	On	yes	26.8 2	33.0 1	28.81	29.7 9	33.4 1	28.7 2	46.0 8	51.2 3	27.70	31.10	30.22	32.71
Use case	Measured Parameter(s)	HVAC Status	HVAC model	CV(RMSD) of people count solution											
Case 1	S1	Off	No	22.1 3	15.8 2	39.08	23.8 2	20.0 4	15.2 1	30.6 0	19.2 3	20.29	15.39	31.30	19.00
	S2	Off	No	23.2 8	20.8 3	24.02	21.3 0	25.2 0	22.2 0	25.1 8	22.1 1	21.48	18.70	19.98	16.56
	S3	Off	No	9.59	9.57	14.31	14.2 9	9.16	9.14	11.8 8	11.8 4	7.60	7.58	10.31	10.26
Case 2	S4	On	No	41.8 4	23.6 9	101.6 4	55.8 6	44.8 5	26.0 7	94.3 2	55.8 3	138.6 6	122.9 3	300.8 0	196.1 9
	S5	On	No	34.1 7	30.3 6	24.31	23.4 8	34.1 6	30.0 4	23.4 5	21.5 8	28.08	26.15	19.81	18.29
	S6	On	No	4.01	3.97	6.28	6.19	7.77	7.64	10.0 8	9.15	5.08	4.86	5.59	5.43
Case 3	S7	On	yes	15.3 1	10.4 8	35.62	21.5 0	14.9 4	10.3 6	32.4 7	20.0 5	23.28	20.83	24.02	21.30
	S8	On	yes	13.6 9	12.6 9	9.92	10.0 7	16.8 3	14.0 6	13.1 0	12.5 7	18.99	17.71	11.45	11.70
	S9	On	yes	3.60	3.52	5.91	5.88	7.83	7.72	10.1 0	9.16	5.07	4.85	5.56	5.41

It can be seen from the table that the accuracy of the inverse solutions varies by solution scenarios and locations. In general, the solutions from Case 1 and Case 3 have higher accuracy than Case 2. For example, Scenario 1 (free-floating, solved with measured temperature) shows the lowest CV(RMSD) among the solutions of air infiltration. Case 3 (Scenario 7 ~ 9) shows better accuracy than Case 2 (Scenario 4~6). Similar results can be seen from the solutions of people count where the accuracy of Case 1 and Case 3 are better than Case 2.

EnergyPlus uses a predictor-corrector mechanism to simulate the relationship between the HVAC system and the zone air. In the “predictor” step, the HVAC system load is estimated from the zone heat gains. In the “corrector” step, the zone air and related terms are updated with the actual simulated HVAC system supplies.

The reasons for the worse accuracy of Case 2 when trying to solve with the sensible heat balance equations include: (1) the inaccuracy from uncertain zone internal thermal mass, and (2) the convective heat transfer between zone interior surfaces and the zone air may not be accounted correctly. In Case 1, the HVAC system is off during both the measurement and the solution period. In the solution period, HVAC kept off (achieved by the dual setpoints thermostat control logic with extremely low cooling setpoint and extremely high setpoints in EnergyPlus). The solution reflects the actual zone air heat balances. However, in Case 2, HVAC is on during the measurement period while it is not simulated in the solution period. Although the zone air parameters and system supply terms are provided in the inverse balance equations, the effects of HVAC supplies on the interior surface temperature and thermal mass are not simulated in the “corrector” step. Thus, the solutions might not reflect the actual zone air heat balances when the thermal mass and interior surface convective heat transfer account for a significant portion in the balance equations. In this case study, Chicago and Houston have more extreme weather conditions than San Francisco, which causes more drastic changes of the zone thermal mass and surface temperatures. When the HVAC system is not simulated in the solution period, the thermal mass and interior surface convective heat transfer are not accurately represented, which leads to the inaccurate inverse solutions. In Case 3, since HVAC is

on during both measurement and solution periods, the balance equations are close to the real zone air heat balances. Thus, the inverse solutions are more accurate than Case 2.

Another finding is that the solution with moisture and CO₂ balance equations are more accurate than the solution with the sensible heat balance equations. That is due to the very small impacts of the interior surface moisture and CO₂ transfer on the corresponding balance equations.

4. Discussion

Traditionally, building performance simulations are used to predict building energy use and environmental performance with known or assumed building characteristics, system operation strategies and control logic, and occupancy schedules. However, some of the model inputs such as the air infiltration rate and occupancy schedule are highly uncertain and hard to measure on the per-zone basis. As discussed in the introduction, extensive research has been conducted to directly or indirectly measure those unknown variables. The limitations of those approaches include the high cost of measurement devices, the disturbance of normal building operation, privacy concerns, and the cost of data collection and analytics. At the meanwhile, environmental sensing technologies become cheaper and more prevalent in modern buildings. The novelty of this study is that it marries the physics-based building energy model with the building environmental measurements to inversely solve the highly uncertain zone-level air infiltration rates and people count in buildings.

The inverse modeling algorithms are verified in the simulation-based case study. Validation of the inverse models in EnergyPlus using laboratory experiments and measured data was conducted in another study and results are to be published in a separate paper. In the case study, normal (forward) simulations are used to generate virtual measurements of HVAC system supply and zone air parameters including dry-bulb temperature, humidity ratio, CO₂ concentration, system supply air temperature and flow rate. Then, the virtual measurements are used as the inputs of the inverse simulation to solve the unknown air infiltration rate or people count. Finally, the solution is compared with the ground truth. Nine solution scenarios (grouped into three use cases) are developed to mimic the different level of measurement availability and model assumptions. The simplest use case is when the HVAC system is off during the measurements. And there is no need to model the HVAC system in the inverse simulation either. This case is suitable when the building is not conditioned, or the HVAC configurations are unknown for creating a building energy model. The second use case is when the HVAC system is operating during the measurement, the system supply terms are considered in the inverse balance equations, but HVAC is not modeled in the inverse simulations. This case is suitable when the measurements of both space and HVAC supply air parameters can be measured, but the HVAC details are unknown to create the building model. The most comprehensive use case is when the HVAC system is operating during the measurements, and the HVAC system details are known and modeled in the inverse simulation. This use case requires the most amount of measurements and the knowledge to model the HVAC system. The modeling parameters can vary significantly from building to building. The inverse modeling method is proposed for real case measurements. In the future, it is important to evaluate the different use cases and measurement scenarios by comparing the efforts and accuracy of the traditional approach against the inverse modeling approach.

It is found that those use cases have different accuracies depending on what the environmental measurement is and what the unknown parameter is. For air infiltration, Case 1 when solving with measured zone air temperature and Case 3 when solving with humidity ratio or CO₂ concentration have better accuracy. For people count, overall, Case 1 and Case 3 have better accuracy than Case 2 when the measurement is temperature or humidity ratio. But the solution with measured CO₂ concentration has good accuracy in all three cases. The differences in the accuracies are caused by the different level of sensitivity of the air parameters to the model settings. In the case study, zone air temperature and humidity ratio can be affected by more factors than zone air CO₂ concentration. Thus, the inverse solution with measured CO₂ concentration has better agreements with the ground truth.

Although the inverse modeling algorithms can solve the uncertain zone air infiltration rate and people count, they are subject to some limitations. First, like the normal simulations, the inverse simulation requires accurate model inputs of building geometry, thermal zoning, lighting and equipment settings. The zone air balance equations can correctly solve the unknown parameters only when the known terms are input correctly. In future studies, we are interested in quantifying the sensitivities of the inverse model results due to the uncertainties of those assumed parameters, which will inform how the inverse models can be applied, for example, at what stage of the model calibration to gain the maximum value, or how to combine with the traditional uncertainty analysis to improve the model accuracy. Secondly, the current algorithms assume the occupants have a constant sensible and latent heat generation rate, as well as a constant CO₂ dissipation rate. This assumption may not be accurate when there is a variety of occupant type and activities. Thirdly, the environmental measurements play an important role in the inverse solution. Data processing such as aligning the measurements to the same time interval with the simulation is necessary. Lastly, there can be computational errors in the inverse simulation. For example, when the indoor and outdoor air temperatures are too close, the program may not solve the correct value of that timestamp because of the overflowing issues. One potential future improvement is to couple the temperature, humidity, and CO₂ inverse algorithms. This way, the inverse solutions from different measured parameters could be used to validate each other at each timestep.

5. Conclusions

This study develops a novel inverse modeling method to solve hard-to-measure building parameters such as zone air infiltration and people count using easy-to-measure zone air temperature, humidity and CO₂ concentration. The inverse method integrates the physics-based building performance models with sensor data, posing a new opportunity of sensor data application in building performance simulation field. The new inverse modeling feature developed in EnergyPlus can improve the simulation accuracy of existing buildings as they reduce the uncertainty in model inputs. Although the inverse models are implemented in EnergyPlus, the algorithms are generic and can be adopted by other building performance simulation engines.

The inverse models should be used with caution in building simulations as they require other building model parameters to be reasonable or tuned to avoid overfitting of the calculated zone air infiltration or people count. Therefore, it is suggested the inverse models be used in later (when most model parameters are corrected or tuned) rather than early stages (when most parameters are of uncertainty) of building energy modeling workflow.

Future research can extend the inverse models to simultaneously solve the two unknown zone parameters (infiltration rate and people count) using two measured zone air parameters (selecting two from air temperature, humidity, and CO₂ concentration). The three inverse models are publicly available in EnergyPlus version 9.1 in 2019. Validation of the inverse models in EnergyPlus using measured data from real buildings is also an important future work.

Acknowledgment

This research was supported by the Assistant Secretary for Energy Efficiency and Renewable Energy, Office of Building Technologies of the United States Department of Energy, under Contract No. DE-AC02-05CH11231. The views expressed in the article do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

References

- [1] S. Prívara, J. Cigler, Z. Váňa, F. Oldewurtel, C. Sagerschnig, and E. Žáčková, "Building modeling as a crucial part for building predictive control," *Energy Build.*, vol. 56, pp. 8–22, 2013.

- [2] X. Li and J. Wen, "Review of building energy modeling for control and operation," *Renew. Sustain. Energy Rev.*, vol. 37, pp. 517–537, 2014.
- [3] T. Hong, J. Langevin, and K. Sun, "Building simulation: Ten challenges," *Build. Simul.*, vol. 11, no. 5, pp. 871–898, 2018.
- [4] S. H. Lee, T. Hong, M. A. Piette, and S. C. Taylor-Lange, "Energy retrofit analysis toolkits for commercial buildings: A review," *Energy*, vol. 89, pp. 1087–1100, 2015.
- [5] E. Iso and others, "13790: 2008 Energy performance of buildings–Calculation of energy use for space heating and cooling," *Int. Stand. Organ.*, 2008.
- [6] D. Kim and J. E. Braun, "A general approach for generating reduced-order models for large multi-zone buildings," *J. Build. Perform. Simul.*, vol. 8, no. 6, pp. 435–448, 2015.
- [7] D. B. Crawley, J. W. Hand, M. Kummert, and B. T. Griffith, "Contrasting the capabilities of building energy performance simulation programs," *Build. Environ.*, vol. 43, no. 4, pp. 661–673, Apr. 2008.
- [8] U. S. DoE, "EnergyPlus engineering reference: the reference to EnergyPlus calculations," *Lawrence Berkeley National Laboratory*. 2009.
- [9] V. S. K. V. Harish and A. Kumar, "Reduced order modeling and parameter identification of a building energy system model through an optimization routine," *Appl. Energy*, vol. 162, pp. 1010–1023, Jan. 2016.
- [10] U. of W. S. E. Laboratory, "TRNSYS : a transient system simulation program." University of Wisconsin, Solar Energy Laboratory, Madison, Wis, 1990.
- [11] "ESP-r, Multi-platform Building Energy Software Tool." [Online]. Available: http://www.esru.strath.ac.uk/Programs/ESP-r_overview.htm. [Accessed: 14-May-2019].
- [12] D. Coakley, P. Raftery, and M. Keane, "A review of methods to match building energy simulation models to measured data," *Renew. Sustain. Energy Rev.*, vol. 37, pp. 123–141, 2014.
- [13] E. M. Ryan and T. F. Sanquist, "Validation of building energy modeling tools under idealized and realistic conditions," *Energy Build.*, vol. 47, pp. 375–382, 2012.
- [14] K. Amasyali and N. M. El-Gohary, "A review of data-driven building energy consumption prediction studies," *Renew. Sustain. Energy Rev.*, vol. 81, pp. 1192–1205, Jan. 2018.
- [15] H. X. Zhao and F. Magoulès, "A review on the prediction of building energy consumption," *Renew. Sustain. Energy Rev.*, vol. 16, no. 6, pp. 3586–3592, 2012.
- [16] A. S. Ahmad *et al.*, "A review on applications of ANN and SVM for building electrical energy consumption forecasting," *Renew. Sustain. Energy Rev.*, vol. 33, pp. 102–109, May 2014.
- [17] I. Gaetani, P. Hoes, and J. L. M. Hensen, "Occupant behavior in building energy simulation : Towards a fit-for-purpose modeling strategy," *Energy Build.*, vol. 121, pp. 188–204, 2016.
- [18] B. Yan, X. Li, A. M. Malkawi, and G. Augenbroe, "Quantifying uncertainty in outdoor air flow control and its impacts on building performance simulation and fault detection," *Energy Build.*, vol. 134, pp. 115–128, 2017.
- [19] C. Younes and C. A. Shdid, "Air infiltration through building envelopes : A review," 2011.
- [20] R. Niemelä, E. Toppila, and A. Tossavainen, "A multiple tracer gas technique for the measurement of airflow patterns in large industrial premises," *Build. Environ.*, vol. 22, no. 1, pp. 61–66, Jan. 1987.
- [21] P. Stabat, D. Marchio, and S. Cui, "CO 2 tracer gas concentration decay method for measuring air change rate," vol. 84, pp. 162–169, 2015.

- [22] T. Lu, A. Knuutila, M. Viljanen, and X. Lu, "A novel methodology for estimating space air change rates and occupant CO₂ generation rates from measurements in mechanically-ventilated buildings," *Build. Environ.*, vol. 45, no. 5, pp. 1161–1172, 2010.
- [23] H. E. Feustel, "Measurements of Air Penetration in Multizone Buildings," vol. 14, pp. 103–116, 1990.
- [24] W. Liu, X. Zhao, and Q. Chen, "Energy & Buildings A novel method for measuring air infiltration rate in buildings," vol. 168, pp. 309–318, 2018.
- [25] C. M. Clevenger and J. Haymaker, "THE IMPACT OF THE BUILDING OCCUPANT ON ENERGY MODELING SIMULATIONS," pp. 1–10, 2001.
- [26] S. Salimi and A. Hammad, "Critical review and research roadmap of office building energy management based on occupancy monitoring," *Energy Build.*, vol. 182, pp. 214–241, 2019.
- [27] W. Wang, J. Chen, and T. Hong, "prediction through machine learning and data fusion of environmental sensing and Wi-Fi sensing in buildings," *Autom. Constr.*, vol. 94, no. March, pp. 233–243, 2018.
- [28] L. M. Candanedo and V. Feldheim, "Accurate occupancy detection of an office room from light, temperature, humidity and CO₂ measurements using statistical learning models," *Energy Build.*, vol. 112, pp. 28–39, 2016.
- [29] D. B. Crawley *et al.*, "EnergyPlus: creating a new-generation building energy simulation program," *Energy Build.*, vol. 33, no. 4, pp. 319–331, Apr. 2001.
- [30] S. H. Lee and T. Hong, "Leveraging Zone Air Temperature Data to Improve Physics-Based Energy Simulation of Existing Buildings Lawrence Berkeley National Laboratory, Berkeley, United States," pp. 1223–1230, 2017.
- [31] M. Deru *et al.*, "U.S. Department of Energy commercial reference building models of the national building stock," *Tech. Rep.*, no. February 2011, pp. 1–118, 2011.
- [32] Y. Chen, T. Hong, and X. Luo, "An agent-based stochastic Occupancy Simulator," *Build. Simul.*, vol. 11, no. 1, pp. 37–49, 2018.