Build Simul (2023) 16: 185–203
https://doi.org/10.1007/s12273-022-0935-7

Validation of virtual sensor-assisted Bayesian inference-based in-situ sensor calibration strategy for building HVAC systems

Guannan Li[1], Jiahao Xiong[1], Shaobo Sun[2], Jian Chen[1] (✉)

1. School of Urban Construction, Wuhan University of Science and Technology, Wuhan 430065, China
2. Department of Building Environment and Energy Engineering, The Hong Kong Polytechnic University, Hong Kong, China

Abstract

For building heating, ventilation and air-conditioning systems (HVACs), sensor faults significantly affect the operation and control. Sensors with accurate and reliable measurements are critical for ensuring the precise indoor thermal demand. Owing to its high calibration accuracy and in-situ effectiveness, a virtual sensor (VS)-assisted Bayesian inference (VS-BI) sensor calibration strategy has been applied for HVACs. However, the application feasibility of this strategy for wider ranges of different sensor types (within-control-loop and out-of-control-loop) with various sensor bias fault amplitudes, and influencing factors that affect the practical in-situ calibration performance are still remained to be explored. Hence, to further validate its in-situ calibration performance and analyze the influencing factors, this study applied the VS-BI strategy in a HVAC system including a chiller plant with air handle unit (AHU) terminal. Three target sensors including air supply (SAT), chilled water supply (CHS) and cooling water return (CWR) temperatures are investigated using introduced sensor bias faults with eight different amplitudes of [−2 °C, +2 °C] with a 0.5 °C interval. Calibration performance is evaluated by considering three influencing factors: (1) performance of different data-driven VSs, (2) the influence of prior standard deviations σ on in-situ sensor calibration and (3) the influence of data quality on in-situ sensor calibration from the perspective of energy conservation and data volumes. After comparison, a long short term memory (LSTM) is adopted for VS construction with determination coefficient R-squared of 0.984. Results indicate that σ has almost no impact on calibration accuracy of CHS but scanty impact on that of SAT and CWR. The potential of using a prior standard deviation σ to improve the calibration accuracy is limited, only 8.61% on average. For system within-control-loop sensors like SAT and CHS, VS-BI obtains relatively high in-situ sensor calibration accuracy if the data quality is relatively high.

Keywords

heating, ventilation and air-conditioning (HVAC);
in-situ sensor calibration;
Bayesian inference (BI);
virtual sensor (VS);
influencing factor;
energy conservation (EC)

Article History

Received: 13 June 2022
Revised: 03 August 2022
Accepted: 19 August 2022

1 Introduction

1.1 Background

In buildings, heating, ventilation and air-conditioning (HVAC) system is important to maintain the indoor thermal comfort and air quality (Yoon 2020; Gao et al. 2022; Liu et al. 2022). Sensors are essential to achieve the online monitoring and control of HVACs, which can make the systems operate reliably and safely (Yoon et al. 2019; Zhang et al. 2021; Li et al. 2022b). If there exist some sensor faults (i.e., sensor biases), reliability and precision of the sensor measured data cannot be guaranteed. The HVAC system operation and control strategies depend heavily on the affected by the quality (reliability and precision) of the sensor measured data (Li and Hu 2018). The system may deviate from the normal working conditions by faulty sensors, which further results in energy penalty and indoor thermal comfort degradation (Choi and Yoon 2020; Yan 2021; Zhang et al. 2022). Hence, it is very necessary to detect, diagnose and in-situ calibrate faulty sensors (Li and Hu 2019; Luo and Fong 2020; Zhou et al. 2021) and ensure HVAC operation (Du and Jin 2007; Sun et al. 2010; Yang et al. 2014), which is very important for maintaining indoor thermal comfort

E-mail: chenjian@wust.edu.cn

List of symbols

$\boldsymbol{b}$	offset vectors
c	storage unit
$\tilde{c}$	updated storage unit
C_w	specific heat capacity of water
d	humidity ratio
$D(x)$	distance function
E	power consumption of chiller
EC	energy conservation
e	residual of the polynomial
F	forget gate
FA	sensor bias fault amplitudes
G	Gaussian probability density function
g	correction function
h	enthalpy
I	input gate
J	mean square error function
k_{t-1}	previous layer output in LSTM
M	mass flow rate
ME	calibration results (mean of posterior distribution)
O	output gate
P	posterior distribution function
$P(x\|Y)$	posterior distribution
$P(Y)$	normalizing constant
$P(Y\|x)$	likelihood function
Q_{en}	results of the heat transfer capacity of the cooling water minus the power consumption of chiller
$Q_{chw,cap}$	cooling capacity from the chilled water at the chiller side
$Q_{cw,cap}$	heat transfer capacity of the cooling water at the cooling tower side
sig	sigmoid activation functions
T	temperature
tanh	hyperbolic tangent activation function
U	unknown variables, target variables, dependent variables
Vr_{in}	input variable involved in the VS model construction process
$\boldsymbol{W}$	weight matrix
x	pre-assumed calibration result
y	actual value of the target variable
$\hat{y}$	prediction value of constructed VS models
$\overline{y}$	average of target variable actual value
$\hat{\hat{y}}$	average of VS prediction value
Y_{ca}	corrected value of target sensor
Y_{me}	measuring system model value
Y_{se}	benchmark of sensor model
Y_{sy}	reliable system model value
Z_0	initial parameter in MCMC
Z_j^*	sampling candidate parameter in iteration j
z	random value
ΔZ_j	random variable in MCMC
ΔE	energy exchange in the heat transfer process
α	acceptance rate
$\varepsilon(V_{vir})$	evaluation index for VS construction accuracy
ε(EC)	deviation between the actual measured value and the reliable system reference value
η	learning rate in MLR-GD
θ_0	intercept of regression model in MLR
θ_i	coefficient of input variables in MLR
θ_{old}	iteration coefficient result of the previous layer
θ_{new}	iteration coefficient result of the present layer
$\pi(x)$	prior distribution
σ	standard deviation of prior distribution
ξ_{ca}	calibration accuracy

Subscripts

a	air side (cooling coil air)
chw	chilled water side
cw	cooling water side
in	input variable
i	the i-th variable of MLR input variables; the i-th sample of testing data
j	the j-th iteration in MCMC
l	the l-th system model in distance function
m	the m-th sensor in distance function
mix	cooling coil inlet
n	total number of MLR input variables; total number of testing data sample
N	normal values
r	number of unknown variables
ret	return water including chiller water return and cooling water return
sup	outlet including cooling coil outlet, cooling tower outlet, and evaporator outlet
t	timestamp
vir	virtual sensor variables
w	water side (cooling coil water)

Abbreviations

AHU	air handling unit
BI	Bayesian inference
CAV	constant-speed air volume terminal
CHS	chilled water supply temperature
CWR	cooling water return temperature
HVAC	heating, ventilation and air-conditioning
LSTM	long short term memory
MCMC	Markov chain Monte Carlo

MLR	multiple linear regression	VS-BI	virtual sensor-assisted Bayesian inference
MLR-LS	MLR-least squares	VSCHS	virtual chilled water supply temperature sensor
MLR-GD	MLR-gradient descent	VSd_mix	virtual humidity ratio sensor on coil inlet
SAT	air supply temperature	VSd_sup	virtual humidity ratio sensor on coil outlet
SCO	sensitivity coefficient optimization method	VSh_mix	virtual air enthalpy sensor on coil inlet
VIC	virtual in-situ calibration	VSh_sup	virtual air enthalpy sensor on coil outlet
VRF	variable refrigerant flow	VSSAT	virtual air supply temperature sensor
VS	virtual sensor		

and building energy efficiency (Papadopoulos and Azar 2016; Guo et al. 2017; Li et al. 2021a; Li et al. 2021b).

1.2 Summary of sensor fault calibration studies in building HVAC systems

Generally, there are two main types of sensor faults (Papadopoulos et al. 2022): hard fault and soft fault. The former refers to completely invalid sensors after irreversible damages. The latter is mainly caused by improper installation, changeable and harsh environments. Soft sensor faults include bias, offset, drift and precision degradation which can be addressed effectively by in-situ sensor calibration. For building HVAC systems, as surveyed by previous studies (Pan et al. 2007; Coakley et al. 2011; Li and Huang 2013), the traditional manual sensor calibration strategies are time-consuming and laborious.

1.2.1 Studies on virtual in-situ calibration (VIC) and Bayesian inference (BI) in building HVACs

For building HVAC systems, online calibration of sensors is effective to address the soft faults using data reconstruction methods based on the correlations between erroneous readings and other data. Yu and Li (2015) firstly proposed a virtual in-situ calibration (VIC) method which is capable of calibrate faulty sensors by obtaining the sensor benchmark based on physical models or statistical methods. Both Bayesian inference (BI) (Yoon and Yu 2017b) and genetic algorithm (GA) (Ramos Ruiz et al. 2016; Baba et al. 2022) can be used for online sensor calibration in building HVAC systems. Furthermore, Yoon and Yu (2017a; 2017c) compared the performance of GA and BI in sensor calibration for building energy systems, and results indicated that estimating a proper searching space for the variables in a deterministic method as GA is difficult. Owing to its advantages of convenient calculation and fewer parameters, BI has been widely used to solve various calibration problems in complex building energy systems (Li et al. 2015; Yuan et al. 2017; Liu et al. 2021). Chong and Menberg (2018) used BI to calibrate building energy models. Hou et al. (2021) elaborated the development status of BI in building energy model calibration. Mokhtari et al. (2020) used BI to calibrate the wind speed sensor of cooling tower in a thermal power plant. Sun et al. (2022) quantified the flow uncertainty of a central cooling system with multiple water-cooled chillers using BI. However, BI with only sensor measurement item requires a lot of time to calibrate each sensor individually in building HVAC systems with a complex sense network. Moreover, it could be very hard to evaluate whether the target sensor is affected by the associated sensors using BI without the system model item. To extend the calibration coverage for multiple sensors, Yoon and Yu (2017b) proposed an improved BI-based VIC by considering both the system model item and sensor measurement item for model extension. To solve the problem of low sensor redundancy in actual systems, Choi and Yoon (2020) further proposed an enhanced VIC method by combining the virtual sensor (VS) with BI to effectively solve the problem of insufficient sensor information when establishing the system model item.

1.2.2 Studies on influencing factors for Bayesian inference (BI) in-situ calibration in building HVACs

The performance of BI in-situ sensor calibration is affected by various influencing factors. Many studies have been conducted to reduce or remove the impacts of these influencing factors. Yoon and Yu (2018a, 2018b, 2018c) summarized the negative factors and put forward effective strategies to address them. Wang et al. (2019) used a sensitivity coefficient optimization (SCO) method to realize the reviving calibration strategy in VIC. Yoon and Yu (2017a, 2017c) compared the performance of GA and BI on in-situ sensor calibration, and quantitatively discussed their differences. Li et al. (2020) divided the AHU system into various partitions and discussed the accuracy of different calibration models (whole calibration, local calibration, component calibration) under these partitions. Wang et al. (2021) calibrated various physical sensors in an AHU system under different faulty working conditions. Yoon et al. (2022) used BI to optimize the energy balance calculation in integrated air conditioner and improve the performance.

1.3 Challenges

The combinations between the sizes of bias and random errors were discussed in the previous study (Yoon and Yu 2017c). The random errors were considered the standard deviations of priors. According to the relative sizes of mean and standard deviation in the prior distribution, there are two prior types including the proportional and disproportional priors. In this study, the impacts of the standard deviation were studied as the prior effect. Besides, some variables used in BI-based in-situ sensor calibration are difficult to measure directly due to the data limitations in practical building HVAC systems, but they can be obtained by constructing virtual sensors. Moreover, the energy conservation equation is also used as an item in the distance function for BI-based modelling process. In practice, the building HVAC systems measured data may not always satisfy the energy conservation equation well owing to the low data quality result from system operational dynamics, the introduced sensor bias fault amplitudes and the service control response. This may lead to certain errors in the BI-based in-situ sensor calibration model based on the energy conservation law. With concerns on the aforementioned issues, it is still challenging to apply the virtual sensor-assisted Bayesian inference based in-situ sensor calibration method in the following aspects.

(1) There is a lack of proper way to guidance how to select the proper priori distribution standard deviation value if no enough prior information can be found in information-poor buildings.
(2) The influence of different VS construction methods to assist the BI calculation on the construction accuracy still remains uncertain and requires performance evaluation.
(3) The way to investigate influences of data quality on in-situ sensor calibration performance also remains uncertain and needs further exploration.
(4) There is also a lack of comprehensive validation of the VS-BI in-situ sensor calibration method for a wide application range of various types of sensors and sensor faults.

1.4 Research contents of this study

Hence, to address these issues, this study systematically investigated the virtual sensor-assisted Bayesian inference (VS-BI)-based in-situ sensor calibration strategy and validated it in a chiller-AHU air conditioning system with typical sensor bias faults.

The research contents are as follows:

(1) Validation of the VS-BI strategy for various types of sensors (SAT and CHS sensors within-feedback-control-loop and CWR sensor out-of-feedback-control-loop) and sensor faults amplitudes (bias faults of [−2 °C, +2 °C] with a 0.5 °C interval).
(2) Performance evaluation and comparison of VS model constructed by different data-driven methods (i.e., long short term memory (LSTM), multiple linear regression (MLR) based least square optimization method, and MLR based on gradient descent optimization method).
(3) Evaluation on the influence of prior standard deviations on in-situ sensor calibration accuracy for various sensor faults with different amplitudes.
(4) Evaluation on the influence of data quality on in-situ sensor calibration accuracy from the perspective of energy conservation and changing data volumes.

2 Methodology

For building HVAC systems, the VS-BI in-situ sensor calibration strategy has two main parts as described in Sections 2.1 and 2.2.

2.1 Principle of virtual sensor (VS)

Virtual sensor (VS) is a functional model, which is developed based on other cheap or easy-to-install physical sensors to obtain the variables which are difficult to measure or the phenomena which are difficult to observe in the systems (Kim et al. 2021). In building energy systems, to make up the system missing information, VSs have been applied to building energy prediction (Sendra-Arranz and Gutiérrez 2020; Markovic et al. 2021), fault detection and diagnosis (Liu et al. 2017; Li et al. 2018; Kim and Lee 2021), construction of VS system (Hong et al. 2021), VIC calibration of sensors (Choi and Yoon 2020), etc. Generally, VSs are used for sensor replacement, backup of existing sensors, and observation of variables or phenomena that are difficult to measure (Kim et al. 2021). For the chiller-AHU air conditioning system concerned in this study, there are some hard-to-obtain measurements, such as the enthalpy, humidity ratio at the inlet and outlet of the coil, the heat exchange rate and heat exchange efficiency in heating systems, etc.

In this study, VS is used to construct such hard-to-obtain measured variables that are required in model development and sensor calibration process of the BI calibration strategy for HVACs. Figure 1 shows the basic structure of VSs, they are constructed by three different data-driven methods, LSTM, Least squares and Gradient descent.

2.1.1 LSTM

As a deep recurrent neural network, LSTM can analyze time series (Bedi and Toshniwal 2019). Although LSTM, it was also used for the developing VS model (Mercante and

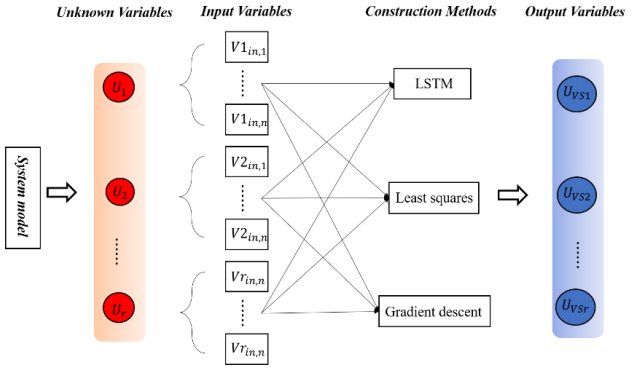


Fig. 1 Construction process of VSs (U represents the unknown variable; V represents the input variable; U_{VS} represent the VS-based output variables)

Netto 2022). LSTM is mainly composed of three gates and a unit. The former three gates are input gate I_t, forgetting gate F_t, output gate O_t, as denoted in Eqs. (1)–(3). The latter storage unit c_t reflects the internal structure state as denoted in Eqs. (4)–(5). The output of current moment layer is k_t as denoted in Eq. (6).

The structure of LSTM is shown in Figure 2. The forgetting gate uses Vr_{in} and previous layer output k_{t-1} as inputs in this layer to determine the information that should be abandoned. The input gate uses Vr_{in} and k_{t-1} as inputs to determine which parts of $\tilde{c}_t$ should be updated to the information of storage unit c_t. The output gate adjusts the output of the LSTM unit based on the superposition of sigmoid and tanh functions. The output result k_t is used as the input of the next layer LSTM structure.

$$F_t = \text{sig}\left(\boldsymbol{W}_{\text{F}} \cdot \left[k_{t-1}, Vr_{\text{in},t}\right] + \boldsymbol{b}_{\text{F}}\right) \tag{1}$$

$$I_t = \text{sig}\left(\boldsymbol{W}_{\text{I}} \cdot \left[k_{t-1}, Vr_{\text{in},t}\right] + \boldsymbol{b}_{\text{I}}\right) \tag{2}$$

$$O_t = \text{sig}\left(\boldsymbol{W}_{\text{O}} \cdot \left[k_{t-1}, Vr_{\text{in},t}\right] + \boldsymbol{b}_{\text{O}}\right) \tag{3}$$

$$\tilde{c}_t = \tanh\left(\boldsymbol{W}_{\text{c}} \cdot \left[k_{t-1}, Vr_{\text{in},t}\right] + \boldsymbol{b}_{\text{c}}\right) \tag{4}$$

$$c_t = F_t * c_{t-1} + I_t * \tilde{c}_t \tag{5}$$

$$k_t = O_t * \tanh\left(c_t\right) \tag{6}$$

where, sig and tanh are sigmoid and the hyperbolic tangent

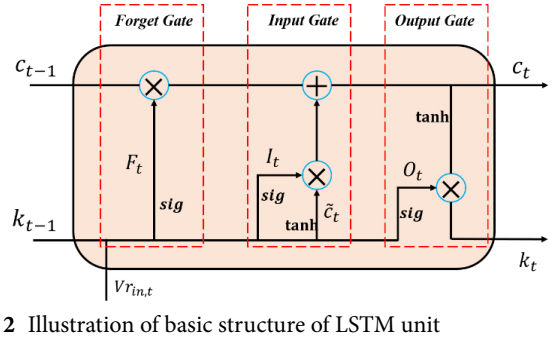


Fig. 2 Illustration of basic structure of LSTM unit

activation functions, respectively. $Vr_{\text{in},t}$ are the input variables of the current time series. $\boldsymbol{W}_{\text{F}}$, $\boldsymbol{W}_{\text{I}}$, $\boldsymbol{W}_{\text{O}}$, $\boldsymbol{W}_{\text{c}}$ are the corresponding weight matrixes, respectively. $\boldsymbol{b}_{\text{F}}$, $\boldsymbol{b}_{\text{I}}$, $\boldsymbol{b}_{\text{O}}$, $\boldsymbol{b}_{\text{c}}$ are the corresponding offset vectors, respectively.

2.1.2 MLR

Multiple linear regression (MLR) can be used to develop predictive model by correlated two or more independent variables to fit a target variable (Yuan et al. 2018). Its basic form can be described as shown in Eq. (7).

$$U_r = \theta_0 + \sum_{i=1}^{n} \theta_i Vr_{\text{in},i} + e \tag{7}$$

where U_r indicates the unknown variable in the system model and it is also the target variable of virtual sensor construction. $Vr_{\text{in},i}$ indicates the input variable involved in the construction process. θ_0 indicates the intercept of the fitting polynomial. n indicates the number of input variables. i indicates the serial number of input variables. θ_i (i = 1, 2, ..., n) indicates the coefficient of the i-th input variable and e indicates the residual of the polynomial.

Both least squares (LS) (Arsie et al. 2017) and gradient descent (GD) (Chen and Miao 2020) methods can be used to obtain the near-optimal MLR predictive model. In this study, both the MLR-LS and MLR-GD are considered to develop the candidate VS models.

MLR-LS: The LS method can find the optimal MLR model by minimizing the sum of squared residuals (Carey and Richardson 2006). Equation (8) is the objective function of LR–mean square error function. Based on Eq. (9), the coefficients and residuals of the MLR model can be directly calculated to complete the modeling of VSs.

$$J\left(\theta_i\right) = \sum_{i=1}^{n} \left(U_r - \left(\theta_0 + \sum_{i=1}^{n} \theta_i Vr_{\text{in},i}\right)\right)^2 \tag{8}$$

$$\frac{\partial J}{\partial \theta_i} = 0 \tag{9}$$

MLR-GD: The GD method has been widely used to optimize machine learning models (Vijayalakshmi et al. 2022) like MLR. GD uses the same objective function as that of LS to minimize the sum of the squared differences between the actual values and the regression values as shown in Eq. (8). Equation (10) shows the regression coefficient solution equation of GD:

$$\theta_{\text{new}} = \theta_{\text{old}} - \eta \cdot \frac{\partial}{\partial \theta_{\text{old}}} J\left(\theta_i\right) \tag{10}$$

where J is mean square error function, η is the learning rate, θ_{new} is the iteration coefficient result of current epoch

layer, and θ_{old} is the iteration coefficient result of the previous epoch layer.

2.2 Principle of Bayesian inference-based in-situ calibration method

BI is a data inference method that modifies the prior probability with the help of new information and converts it into a posterior probability. BI can be used to statistically derive the results in order to minimize the distance function of the calibration problem. After BI-based calibration, the mean value of the posterior distribution should be used to replace the fault amplitude, which is the calibration result.

According to BI, $D(x)$ in Eq. (13) is the distance function (Yoon and Yu 2017b; Choi and Yoon 2020), and the posterior distribution $P(x|Y)$ is the probability density function that minimizes the distance function $D(x)$ under the condition of prior distribution. The specific structure of the distance function (Yoon and Yu 2017b) is shown in Eqs. (14)–(17). Equation (14) includes system model calibration item and sensor calibration item, wherein the system model calibration item represents the difference between the measured value Y_{me} of the actual system and the benchmark Y_{sy} of the reliable system, the sensor calibration item represents the difference between calibration value Y_{ca} calculated by Eq. (17) and sensor benchmark Y_{se} obtained by sensor model. Equation (15) is the calculation function of the reliable system model and Eq. (16) is the calculation function of the sensor model. Both f_1 and f_2 need to calibrate the relevant variables $Y_{re,n}$ and unknown variables U_n through the correction function $g(x)$.

$$P(x|Y)=\frac{P(Y|x)\times\pi(x)}{P(Y)} \tag{11}$$

$$P(Y)=\int P(Y|x)\pi(x)\mathrm{d}x \tag{12}$$

$$P(Y|x)=\prod_{t=1}^{T}\frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{1}{2\sigma^2}D(x)\right] \tag{13}$$

$$D(x)=\underbrace{\sum_{l}^{L}\left(Y_{sy,l}-Y_{me,l}\right)^2}_{\text{System model calibration item}}+\underbrace{\sum_{m}^{M}\left(Y_{se,m}-Y_{ca,m}\right)^2}_{\text{Sensor calibration item}} \tag{14}$$

$$Y_{sy}=f_1\left(Y_{re,1},Y_{re,2},\ldots,Y_{re,n},U_1,U_2,\ldots,U_r\right) \tag{15}$$

$$Y_{se}=f_2\left(Y_{re,1},Y_{re,2},\ldots,Y_{re,n},U_1,U_2,\ldots,U_r\right) \tag{16}$$

$$Y_{ca}=g(O,x) \tag{17}$$

$$\pi(x)=\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{x^2}{2\sigma^2}} \tag{18}$$

where, x is the preset calibration result. $\pi(x)$ is the prior distribution of x, usually choose zero-mean Gaussian distribution as shown in Eq. (18), σ is the standard deviation of prior distribution determines the form of the prior distribution. $P(x|Y)$ is the posterior distribution, $P(Y)$ is the normalization constant, $D(x)$ is the distance function, Y_{sy} is the benchmark of reliable system model, l is the number of system models, Y_{se} is the benchmark of sensor model, m is the number of sensors, Y_{me} is the measured value of system model, Y_{re} is the corrected value of model related variables, U_r is the corrected value of unknown variables in the model, Y_{ca} is the calibrated value of the target sensor through the preset calibration result x, and O is the original measured value of the target sensor.

It is usually difficult to calculate the integral in Eq. (12) directly. Therefore, the prior distribution is defined as normal distribution according to the central limit theorem (Dudley 1978), and then the Markov chain Monte Carlo (MCMC) method is used to solve the normalization constant $P(Y)$ (Gilks et al. 1996; Sinha 2009). Metropolis Hastings algorithm is a widely used MCMC sampling method (Hastings 1970; Huang et al. 2016). It is used for BI sampling to generate posterior distribution samples of preset calibration results x, so as to obtain the statistical characteristics of posterior distribution (mean, standard deviation, etc.). The basic steps of Metropolis Hastings algorithm are as follows.

Step 1: The initial parameter Z_0 of Markov chain should be preset from a prior distribution. For example, Z_0 can be defined as the mean of a prior distribution. If a zero-mean Gaussian distribution is used, then Z_0 should be 0 in the beginning of the entire iteration process.

Step 2: After Step 1, a Gaussian probability density function $G(Z_j^*\,|\,Z_{j-1})$ can be preset. The center of Gaussian probability density curve should be sampling parameter Z_{j-1} in the previous (j–1)-th iteration. For the j-th iteration, the sampling candidate parameter Z_j^* could be determined using the preset probability density function, which is obtained by adding a random variable ΔZ_j to the previous sampling parameter Z_{j-1} as denoted in Eq. (19).

$$Z_j^*=Z_{j-1}+\Delta Z_j,\quad \Delta Z_j\in[-z,z] \tag{19}$$

where z is a random value.

Step 3: An acceptance rate of candidate parameters α should be calculated based on Eq. (20).

$$\alpha=\min\left\{1,\frac{P(Z_j^*\,|\,Y)G(Z_{j-1}\,|\,Z_j^*)}{P(Z_{j-1}\,|\,Y)G(Z_j^*\,|\,Z_{j-1})}\right\} \tag{20}$$

where $P(Z_j^*\,|\,Y)$ is the posterior distribution function of candidate parameter Z_j^*, $P(Z_{j-1}\,|\,Y)$ is the posterior distribution function of the previous parameter Z_{j-1}, $G(Z_{j-1}\,|\,Z_j^*)$ represents the Gaussian probability density

function with the center of Z_j^*, $G(Z_j^* \mid Z_{j-1})$ represents the Gaussian probability density function with the center of Z_{j-1}.

Step 4: This step is to judge whether the candidate parameters should be received or not by comparing the acceptance rate α with another random number θ which is randomly generated in the range of [0, 1] as described in Eq. (21).

$$Z_j = \begin{cases} Z_j^* & \alpha \geq \theta \\ Z_{j-1} & \alpha < \theta \end{cases} \tag{21}$$

Step 5: Steps 2 to 4 should be repeated until the set maximum sampling value is reached. After the repeat process, both the samples of posterior distribution and probability density function curve can be obtained. Finally, the calibrated sensor bias x can be determined using the posterior distribution and the density function.

In this study, the BI calibration and Metropolis Hastings algorithm are implemented in pycharm2019 Version 3.3, using the pymc3 module. Section 4.2.1 presents an example illustration of the calibration steps for the SAT sensor fault sample with −2 °C bias (10:30 a.m., Aug. 1).

3 Research framework

Figure 3 shows the research framework of this study consisting of three main parts.

(1) Strategy validation

The VS-BI in-situ sensor calibration strategy is validated using a chiller-AHU system with three target faulty sensors. For each faulty sensor, eight different fault amplitudes of biases are used. The validation process and the validation results are present in Sections 4.2 and 5.2, respectively.

(2) Construction of VSs

VS models of water/air temperatures and humidity ratios are constructed by three different data-driven methods. Performance evaluation and the hyper-parameters optimization are conducted so as to prepare proper VS models for the BI-based in-situ sensor calibration process.

(3) Discussion of calibration parameter and data quality

The influence of prior standard deviation on VS-BI-based in-situ sensor calibration results is evaluated by setting different standard deviations. The influence of data quality on VS-BI-based in-situ sensor calibration results is evaluated from the perspective of energy conservation (EC) equation and data volumes. Deviation degrees of the EC equation can be simulated by changing the data volumes of the selected data set. Hence the influences of data quality on the calibration results can be investigated.

4 Case study

Figure 4 shows the target Chiller-AHU system in the case study. The system contains a constant-speed air volume (CAV) system, including reheat coil, cooling coil, fan, etc. For the cooling condition, the outdoor air is mixed with the indoor return air, then the mixed air flows into the cooling coils that are cooled by the chilled water. The cooled air is afterward supplied to each air-conditioned area. The air flowrate is controlled by the fan. The control unit regulates the coil water flow so that air supply temperature (SAT) maintains around its setpoint. In summer, the indoor air temperature setpoint is 24 °C and the SAT setpoint is 14 °C. As shown in Figure 4, the Chiller-AHU system services for a five-zone water-cooled office building located in Chicago (annual average temperature is 9.99 °C, annual

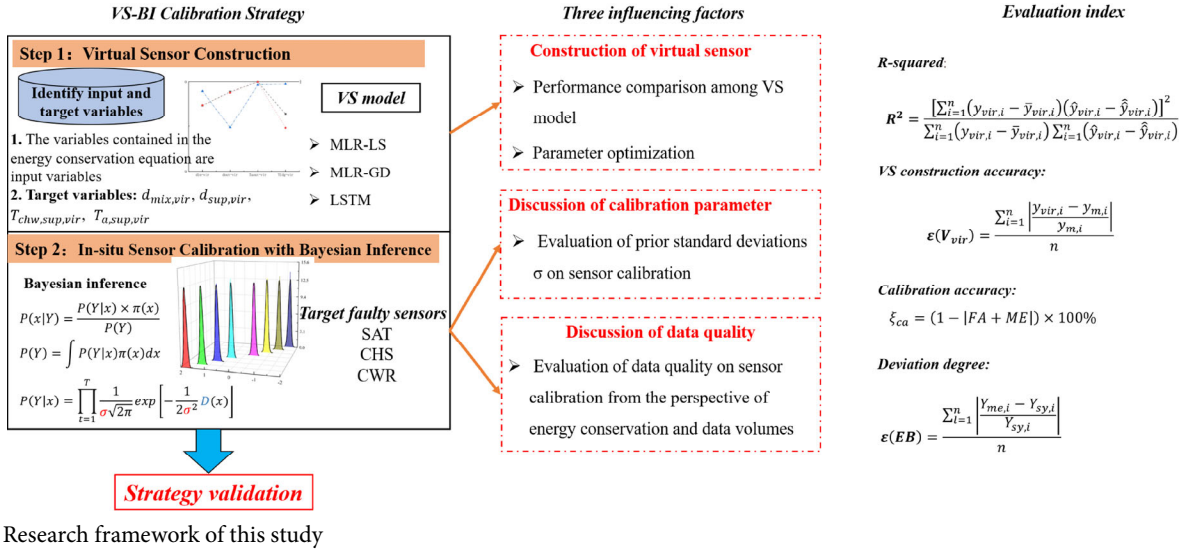


Fig. 3 Research framework of this study

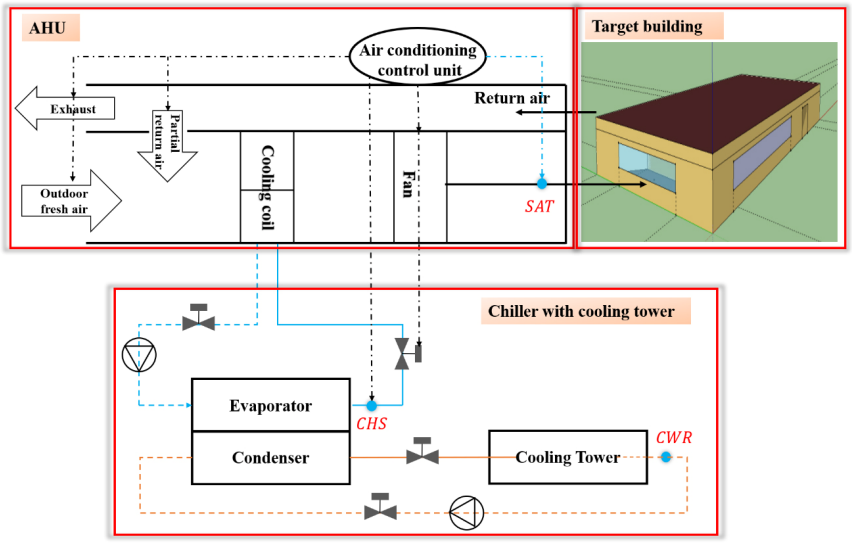


Fig. 4 Illustration of the target building and the chiller-AHU system with target sensors

average humidity is 70.34%) which is an official example in EnergyPlus example files (EnergyPlus 2010). The building is a single-story rectangular building (30.48 m × 2.44 m), which is composed of four external areas and one internal area. A 0.6 m high return air chamber is set above each area to ensure indoor air circulation, and each outer wall was equipped with windows. Details of building envelope are shown in Table 1.

Refer to Choi and Yoon's study (Choi and Yoon 2020), the data collected in July of the typical summer cooling season are used to develop VSs and the data collected in the first 15 days of August are used to test the developed VSs. As shown in Table 2 and Figure 5, all data were collected from 10:30 a.m. to 17:30 p.m. in weekdays with an interval of half an hour. Both the data collected from the system shutdown state and the dynamical data collected at the beginning of startup should be preliminarily eliminated. In this study, we used the same testing data set for VS model performance evaluation. As shown in Figure 4, the three target sensors concerned include SAT and chilled water supply temperature (CHS) involved in the system feedback control loop, and the cooling water return temperature sensor (CWR) that is not involved in the system feedback control loop.

Table 1 Description of building envelope

Enclosure structure	Material	Thermal conductivity (W/(m·K))
Roof	Outside: polyurethane extruded board (100 mm)	0.245
	Inside: reinforced concrete (150 mm)	1.95
Exterior wall	Outside: polyurethane extruded board (75 mm)	0.245
	Inside: reinforced concrete (150 mm)	1.95
Floor	Outside: polyurethane extruded board (30 mm)	0.245
	Inside: reinforced concrete (150 mm)	1.95
Interior wall	Aerated concrete block (200 mm)	0.26
Window	Ordinary glass (3 mm) × two	0.9
	Air layer	/
Door	Grey glass (3 mm)	0.9

Table 2 Data information

Building energy system working timetable		Data set information
Jul. 1–Jul. 31	Working days: only weekdays; System working hours: 10:30–17:30; Sampling interval: 30 minutes	Training set: 490 samples for 3 VS construction methods and the BI-based sensor calibration method
Aug. 1–Aug. 15		Testing set: 165 samples for 3 VS construction methods and the BI-based sensor calibration method

	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
Jul	25	26	27	28	29	30	1
	2	**3**	**4**	**5**	**6**	**7**	8
	9	**10**	**11**	**12**	**13**	**14**	15
	16	**17**	**18**	**19**	**20**	**21**	22
	23	**24**	**25**	**26**	**27**	**28**	29
Aug	30	**31**	**1**	**2**	**3**	**4**	5
	6	**7**	**8**	**9**	**10**	**11**	12
	13	**14**	**15**				

Fig. 5 Illustration of working days of the HVAC system

4.1 Setup of the target sensor bias faults

This study validates the VS-BI in-situ sensor calibration strategy in two aspects: different types of sensors and different fault amplitudes of sensor bias faults. Table 3 shows the

Table 3 Settings of sensor faults in this study

Target sensor	Fault amplitudes (°C)	Interval (°C)	Fault type
SAT	−2 to +2	0.5	Single sensor fault
CHS	−2 to +2	0.5	Single sensor fault
CWR	−2 to +2	0.5	Single sensor fault

setups of the three target faulty sensors. The sensor bias faults are from −2 °C to +2 °C with an interval of 0.5 °C, both positive and negative biases are considered.

4.2 Calibration process of target sensors

In the calibration process, for system within-control-loop sensors like SAT and CHS, it is necessary to construct both the system model item and the sensor measurement calibration item. For the system out-of-control-loop CWR sensor, only the sensor measurement calibration item is required.

4.2.1 Calibration process of SAT

SAT is located at the outlet side of the cooling coil. The system model is constructed according to the cooling coil energy conservation, as shown in Eqs. (22)–(26). Equation (22) represents the heat exchange at the air side, which can be obtained based on the actual measurements. Equation (23) represents the heat exchange at the water side, which is the benchmark of the system. Their difference is defined as the distance function of the system model calibration item. Meanwhile, the distance function of the sensor calibration item is constructed based on the correction function proposed in Section 2, as shown by Eq. (24). The sum of the two items is the final distance function, as shown in Eq. (25). The enthalpy h in Eq. (22) and the humidity in Eq. (26) cannot be directly obtained, so the data-driven method is used to calculate the d_{vir} (including the humidity ratio of the coil air supply outlet $d_{sup,vir}$ and the humidity ratio of the coil air inlet $d_{mix,vir}$) and further the enthalpy h by Eq. (26). Figure 6 shows an example illustration of the calibration steps for the SAT sensor fault sample with −2 °C bias (10:30 a.m., Aug. 1). This study set the sampling value as 16000.

$$\Delta E_a = \left(h_{a,sup} - h_{a,mix}\right) \times M_a \tag{22}$$

$$\Delta E_w = \left(T_{w,ret} - T_{w,sup}\right) \times M_w \times C_w \tag{23}$$

$$T_{a,sup,ca} = T_{a,sup} + x_{a,sup} \tag{24}$$

$$D\left(x_{a,sup}\right) = \underbrace{\left(\Delta E_a - \Delta E_w\right)^2}_{\text{System model calibration item}} + \underbrace{\left(T_{a,sup,vir} - T_{a,sup,ca}\right)^2}_{\text{Sensor calibration item}} \tag{25}$$

$$h = 1.01 \times T + d_{vir} \times \left(2501 + 1.85 \times T\right) \tag{26}$$

where ΔE_a represents the heat exchange at the air side. ΔE_w represents the heat exchange at the water side. $h_{a,sup}$ and $h_{a,mix}$ represent the enthalpy of the supply and mixed air, respectively. M_a and M_w represent the air and chilled water mass flowrate, respectively. $T_{w,sup}$ and $T_{w,ret}$ represent the chilled water inlet and outlet temperatures, respectively. C_w is the specific heat capacity of water (i.e., 4.186 kJ/(kg·°C)).

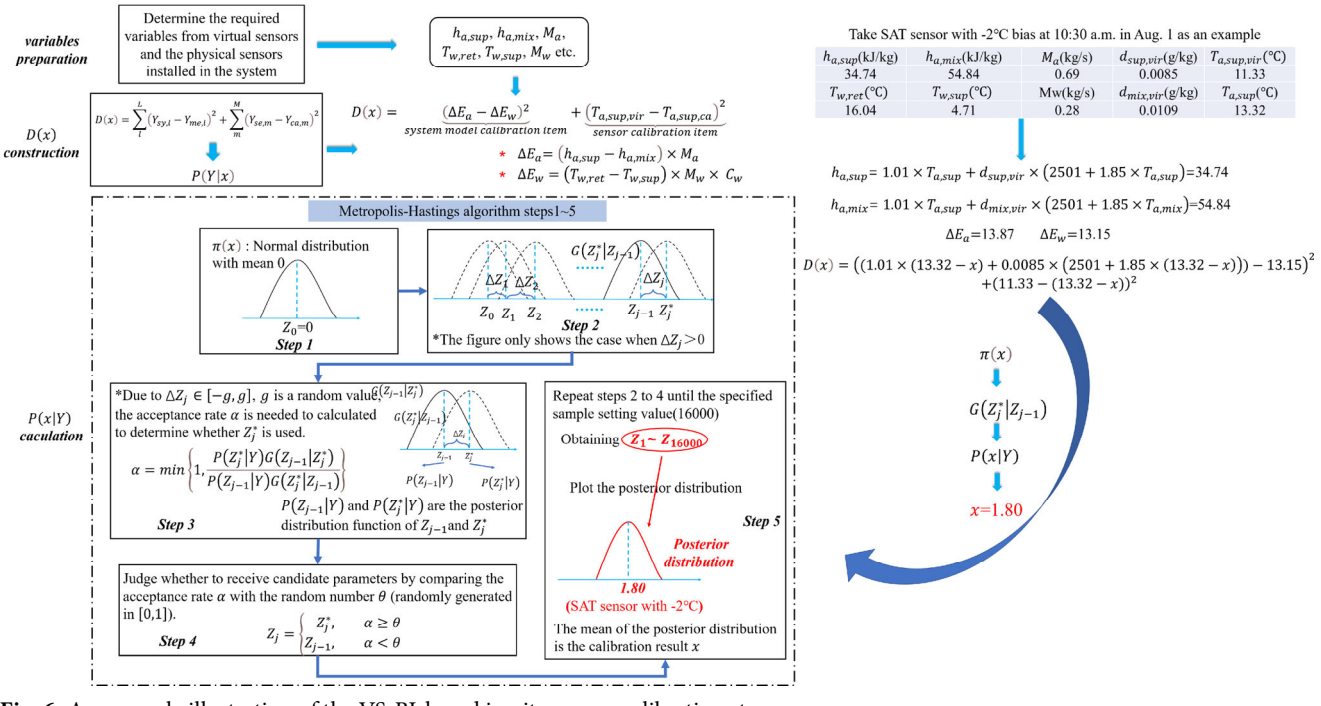


Fig. 6 An example illustration of the VS-BI-based in-situ sensor calibration steps

$T_{a,sup,ca}$ and $T_{a,sup}$ represent the calibrated and measured air supply temperature, respectively. $x_{a,sup}$ represents the pre-assumed SAT calibration value which is converted to the mean value of a posterior distribution. $T_{a,sup,vir}$ represents the SAT constructed by virtual sensor.

Both data-driven and grey-box methods are used to construct the VS models for $h_{a,sup}$, $h_{a,mix}$, $d_{sup,vir}$, $d_{mix,vir}$, $T_{a,sup,vir}$. Details of the VS models (i.e., VSH_sup, VSh_mix, VSd_sup, VSd_mix, and VSSAT) are shown in Table 4.

4.2.2 Calibration process of CHS

CHS is located at the outlet of the evaporator. For a chiller and cooling tower system, the chiller cooling capacity should be equal to the heat exchange of the cooling tower minus the chiller power consumption. Q_{en} is the result of the heat exchange of cooling tower minus the power consumption of chiller, and take Q_{en} as the benchmark of the reliable system, as shown in Eq. (27). Equation (28) is the refrigerating capacity of chilled water defined as $Q_{chw,cap}$, which is the measured value of the actual system. Equation (29) is the heat exchange of cooling water defined as $Q_{cw,cap}$. The difference between $Q_{chw,cap}$ and Q_{en} is defined as the system model calibration item. The sensor calibration item is constructed according to the correction function, as shown in Eq. (30). The sum of these two items is the final distance function, as shown in Eq. (31).

$$Q_{en} = Q_{cw,cap} - E \tag{27}$$

$$Q_{chw,cap} = \left(T_{chw,ret} - T_{chw,sup}\right) \times M_{chw} \times C_w \tag{28}$$

$$Q_{cw,cap} = \left(T_{cw,ret} - T_{cw,sup}\right) \times M_{cw} \times C_w \tag{29}$$

$$T_{chw,sup,ca} = T_{chw,sup} + x_{chw,sup} \tag{30}$$

$$D\left(x_{chw,sup}\right) = \underbrace{\left(Q_{chw,cap} - Q_{en}\right)^2}_{\text{System model calibration item}} + \underbrace{\left(T_{chw,sup,vir} - T_{chw,sup,ca}\right)^2}_{\text{Sensor calibration item}} \tag{31}$$

where $T_{chw,ret}$ and $T_{chw,sup}$ represent the return and supply water temperatures of chilled water, respectivley. M_{chw} represents the chilled water flow. C_w is the specific heat capacity of water (4.186 kJ/(kg·°C)). $T_{cw,ret}$ and $T_{cw,sup}$ represent the measured return and supply water temperatures of cooling water, respectivley. M_{cw} represents the cooling water flow. $T_{chw,sup,ca}$ represents the calibrated chilled water supply temperature. $x_{chw,sup}$ represents the pre-assumed chilled water supply temperature calibration value which is converted to the mean value of a posterior distribution. $T_{chw,sup,vir}$ represents the CHS constructed by the VSCHS model as shown in Table 4.

Table 4 Detailed information of the virtual sensor models

Target VS	Input variables	Output variable	Method
VSd_sup	φ_{sup}, $T_{a,sup}$	$d_{sup,vir}$	Data driven
VSd_mix	φ_{mix}, $T_{a,mix}$	$d_{mix,vir}$	Data driven
VSh_sup	$d_{sup,vir}$, $T_{a,sup}$	$h_{a,sup}$	Grey-box
VSh_mix	$d_{mix,vir}$, $T_{a,mix}$	$h_{a,mix}$	Grey-box
VSSAT	M_a, $T_{a,mix}$, M_w, $T_{w,sup}$, $T_{w,ret}$	$T_{a,sup,vir}$	Data driven
VSCHS	E, M_{chw}, $T_{chw,ret}$, M_{cw}, $T_{cw,ret}$, $T_{cw,sup}$	$T_{chw,sup,vir}$	Data driven

4.2.3 Calibration process of CWR

Since CWR is not involved in the system feedback control loop, only the sensor measurement calibration item is required. Equations (32) and (33) present the distance function and the correction function, respectively.

$$D\left(x_{cw,ret}\right) = \underbrace{\left(T_{cw,ret,N} - T_{cw,ret,ca}\right)^2}_{\text{Sensor calibration item}} \tag{32}$$

$$T_{cw,ret,ca} = T_{cw,ret} + x_{cw,ret} \tag{33}$$

where $x_{cw,ret}$ represents the calibration constant of the CWR sensor. $T_{cw,ret,N}$, $T_{cw,ret,ca}$, and $T_{cw,ret}$, represent the normal, calibrated and measured CWR temperatures, respectively.

5 Results and discussion

5.1 Virtual sensor results

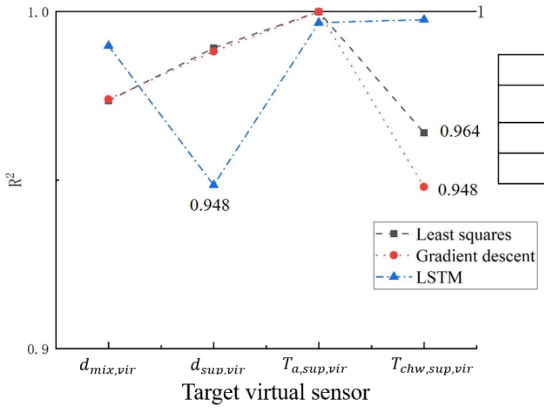
5.1.1 Comparative results of three virtual sensor methods

In this section, three methods (i.e., MLR-LS, MLR-GD, LSTM) are used to construct virtual sensors. The VS model performance is evaluated by evaluation index R^2 in Eq. (34).

$$R^2 = \frac{\left[\sum_{i=1}^{n}\left(y_i - \overline{y}_i\right)\left(\hat{y}_{vir,i} - \hat{\overline{y}}_{vir,i}\right)\right]^2}{\sum_{i=1}^{n}\left(y_i - \overline{y}_i\right)\sum_{i=1}^{n}\left(\hat{y}_{vir,i} - \hat{\overline{y}}_{vir,i}\right)} \tag{34}$$

where y_i and $\hat{y}_{vir,i}$ represent the actual value of the target variable and the virtual sensor construction value, respectively. $\overline{y}_i$ and $\hat{\overline{y}}_{vir,i}$ represent the average actual value and average construction value of the target variable, respectively. n represents the total number of samples of the virtual sensor construction testing set and i represents the i-th sample of testing set. R^2 closer to 1 means higher construction accuracy.

Figure 7 shows that all three methods have R^2 around 0.95. Comparative results indicates that LSTM shows slightly higher average construction accuracy in terms of R^2. For the MLR-LS model, this study used the "statsmodels.api" (Schleich et al. 2016) in Python. Figures 8(a)–(d) show the distributions of residuals for the four VS models respectively. The sum squared residuals of four VS models are 1.37×10^{-5},



Method	Mean results of R^2
MLR-LS	0.981
MLR-GD	0.977
LSTM	0.984

Fig. 7 Performance evaluations of four VS models developed by three methods

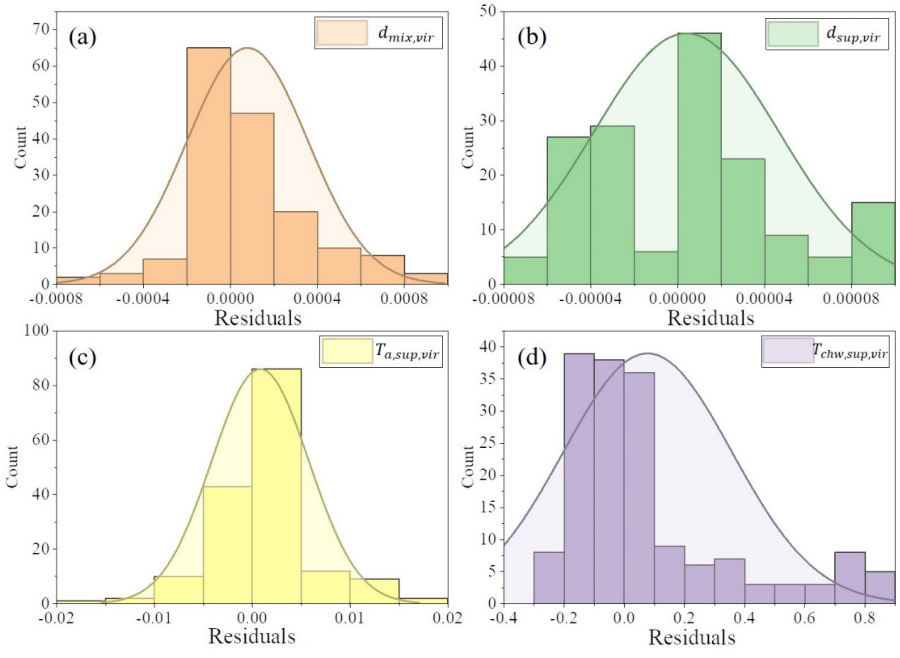


Fig. 8 Distributions of residuals for four MLR-LS VS models: (a) $d_{\mathrm{mix,vir}}$, (b) $d_{\mathrm{sup,vir}}$, (c) $T_{\mathrm{a,sup,vir}}$, and (d) $T_{\mathrm{chw,sup,vir}}$

3.01×10^{-7}, 4.2×10^{-3} and 13.8, respectively. For the MLR-GD model, this study used "gradientDescentMulti" (Kasai 2018) from the Matlab R2018 as the main function for predictive modeling. The learning rate for gradient descent is 0.1, the number of iterations is 1000.

5.1.2 Hyper-parameter optimization results of the selected LSTM-based VS models

From Section 5.1.1, LSTM should be selected as the VS construction method. Hyper-parameters optimization is conducted in two main steps (Fan et al. 2017; Wang et al. 2020; Li et al. 2022a) using another performance evaluation index mean relative error $\varepsilon(V_{\mathrm{vir}})$ by Eq. (35).

$$\varepsilon(V_{\mathrm{vir}}) = \frac{\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_{\mathrm{vir},i}}{y_i}\right|}{n} \tag{35}$$

where, y_i and $\hat{y}_{\mathrm{vir},i}$ represent the actual and the VS-based predictive values of the target variable, respectively. n represents the total number of samples of the testing set and i represents the i-th sample of testing set. Smaller $\varepsilon(V_{\mathrm{vir}})$ means higher VS construction accuracy of the target variable.

Step 1: Conduct the grid search on the LSTM-based VS model. A combination of various hyper-parameters is prepared for grid search with given candidate values in a relatively wide range. For the structure of the LSTM model, this study considers the number of LSTM layer, the activation function and the optimizer. Four hyper-parameters are selected for optimization including Learning rate, Batch size, number of neurons in the hidden layer (Units) and number of iterations (Epoch).

Step 2: The grid search process aims to obtain the VS predictive performance in terms of the evaluation index $\varepsilon(V_{vir})$.

Table 5 shows the hyper-parameters tuning results after a comprehensive grid search on 81 combinations of four hyper-parameters for all four data-driven VS models.

Radar plot in Figure 9 shows an example of the optimization results for the VS $d_{sup,vir}$ of nine combinations of hyper-parameters (9 from 81). After predictive performance comparison in terms of $\varepsilon(V_{vir})$, the optimal combination of hyper-parameters for the LSTM-based VSd_sup is determined with red color. The optimal LSTM model has parameters as follows: the number of LSTM layer is 1, the activation function is ReLU and the optimizer is Adam; learning rate is 0.01, Batch size is 32, number of neurons in the hidden layer (Units) is 19 and number of iterations (Epoch) is 150.

5.2 In-situ sensor calibration strategy validation results

Figures 10–12 show the calibration results of three target sensors under different fault amplitudes. The mean value of samples in the posterior distribution should be the calibration result. Standard deviation of the posterior distribution indicates the stability of calibration results. Height of the posterior distribution represents the probability of calibration results. Here, a calibration accuracy ξ_{ca} is defined to determine the accuracy of the calibration results, as in Eq. (36).

$$\xi_{ca} = (1 - |\mathrm{FA} + \mathrm{ME}|) \times 100\% \tag{36}$$

where FA represents different fault amplitudes and ME represents calibration results.

The average calibration accuracy of all three sensors are around 72%. For calibration results of SAT in Figure 10, the calibration accuracy increases as the absolute fault amplitude decreases. The maximum probability density is over 12.5 and standard deviation is around 0.06. For a normal posterior distribution, more concentrated distribution means higher reliability of calibration results. For calibration results of CHS, Figure 11 shows that the posterior distribution is much more concentrated. For CRW sensor that does not

Table 5 Results of LSTM hyper-parameters optimization

Target variable of the VS model	Hyper-parameters combination							
	Layer of LSTM [1, 2, 3]	Activation function [ReLU, sigmoid, tanh]	Optimizer [Adam, RMSprop, Adamax]	Learning rate [0.001, 0.005, 0.01]	BatchSize [32, 64, 96]	Unit [6, 19, 32]	Epoch [50, 100, 150]	$\varepsilon(V_{vir})$
$d_{sup,vir}$	1	ReLU	Adam	0.01	32	19	150	0.22%
$d_{mix,vir}$	1	ReLU	Adam	0.01	64	32	150	0.34%
$T_{a,sup,vir}$	1	ReLU	Adam	0.01	32	6	150	0.08%
$T_{chw,sup,vir}$	1	ReLU	Adam	0.01	96	32	150	0.37%

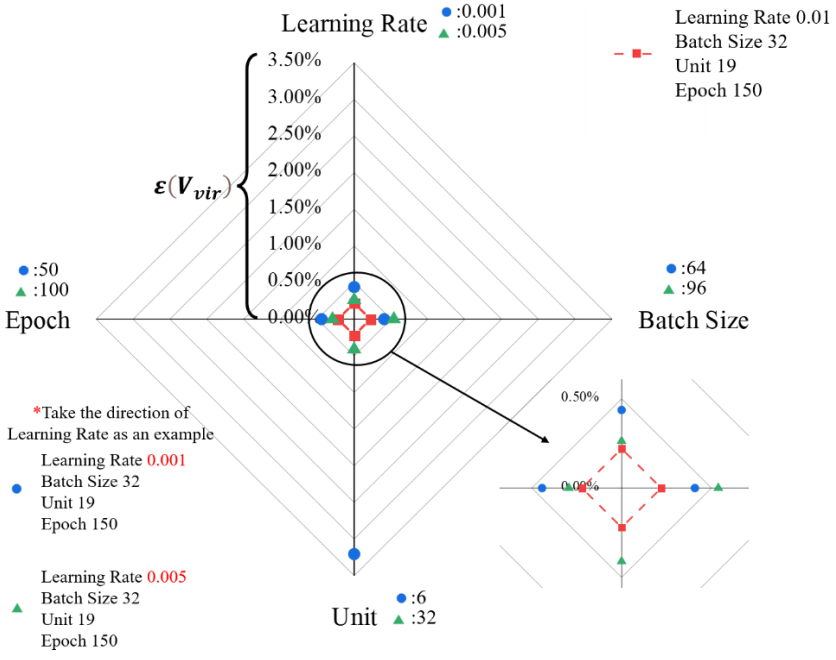
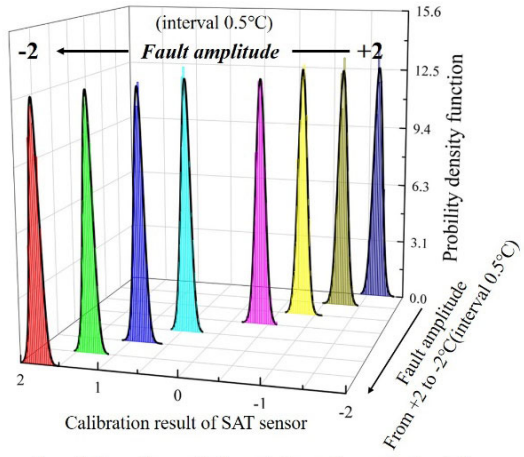
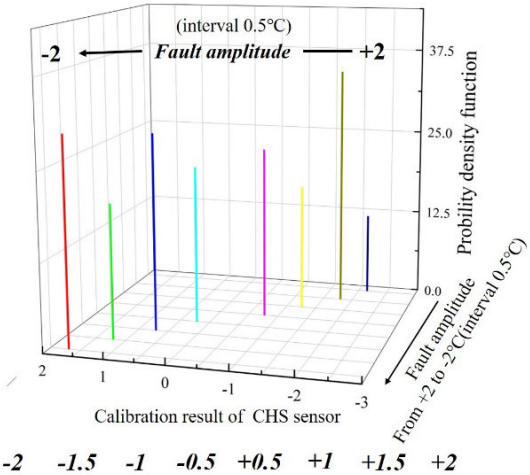


Fig. 9 An example of the optimization results for $d_{sup,vir}$



Fault amplitude :	*-2*	*-1.5*	*-1*	*-0.5*	*+0.5*	*+1*	*+1.5*	*+2*
Calibration results :	1.80	1.35	0.89	0.45	-0.40	-0.84	-1.27	-1.66
Standard deviation of Posterior distribution :	0.063	0.062	0.063	0.062	0.063	0.061	0.063	0.063
Results of Calibration accuracy:	79.97%	84.50%	88.47%	95.18%	89.65%	83.7%	76.95%	66.36%

Fig. 10 Posterior probability distribution and calibration results of SAT (Note: for example, for the SAT sensor fault with −2 °C bias, its BI-based calibration results are presented in normal distribution which is a type of Gaussian function with the red area under the Gaussian curve. The red color is only used to separate the −2 °C bias from the other biases. The BI-based calibration results are presented not only in the form of a Gaussian curve but also providing its Gaussian distribution characteristics with mean = 1.80 °C and standard deviation = ±0.063 °C)



Fault amplitude :	*-2*	*-1.5*	*-1*	*-0.5*	*+0.5*	*+1*	*+1.5*	*+2*
Calibration results :	1.695	1.204	0.710	0.217	-0.771	-1.260	-1.799	-2.147
Standard deviation of Posterior distribution :	1.26×10^{-5}	1.45×10^{-5}	1.27×10^{-5}	1.55×10^{-5}	1.98×10^{-5}	1.80×10^{-5}	1.66×10^{-5}	1.67×10^{-5}
Results of Calibration accuracy:	69.55%	70.37%	71.03%	71.68%	72.93%	73.98%	70.07%	76.42%

Fig. 11 Posterior probability distribution and calibration results of CHS

involve in system feedback control loop, the average calibration accuracy is 99.3% as shown in Figure 12.

5.3 Discussion on influencing factors

5.3.1 Calibration parameter—prior standard deviation σ

The prior distribution standard deviation is a major parameter affecting the in-situ sensor calibration results. This section mainly discusses the effect of different prior standard deviation on the calibration results. Table 6 lists the candidate prior standard deviations in the range of 0.2, 0.4, 0.8, 1 and 1.5. For different prior standard deviations, Figures 13 and 14 show the in-situ sensor calibration results of SAT and CWR, respectively. Different fault amplitudes correspond to different colors. The influence of different prior standard deviations is expressed in the form of color gradation changes. For each fault amplitude, darker color means

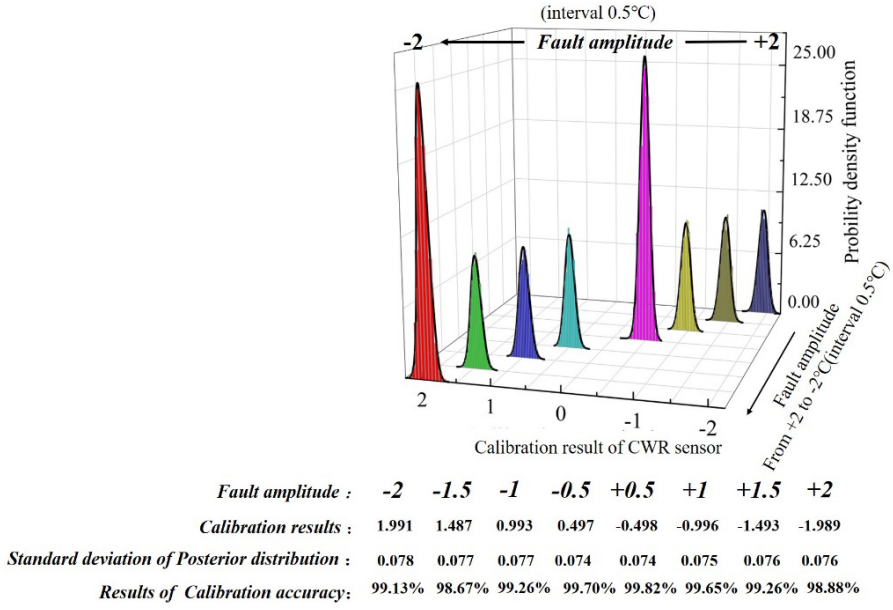


Fig. 12 Posterior probability distribution and calibration results of CWR

Table 6 Calibration results for different prior standard deviations

Sensor	Fault amplitude (°C)	Different values of σ				
		0.2	0.4	0.8	1	1.5
SAT	+2	−1.512	−1.621	−1.655	−1.664	−1.659
	+1.5	−1.156	−1.244	−1.265	−1.269	−1.270
	+1	−0.765	−0.820	−0.835	−0.837	−0.840
	+0.5	−0.364	−0.391	−0.397	−0.396	−0.402
	−0.5	0.409	0.439	0.445	0.452	0.447
	−1	0.810	0.869	0.888	0.885	0.891
	−1.5	1.225	1.316	1.339	1.345	1.346
	−2	1.640	1.765	1.797	1.800	1.806
CHS	+2	−2.236	−2.236	−2.236	−2.236	−2.236
	+1.5	−1.799	−1.799	−1.799	−1.799	−1.799
	+1	−1.260	−1.260	−1.260	−1.260	−1.260
	+0.5	−0.771	−0.771	−0.771	−0.771	−0.771
	−0.5	0.217	0.217	0.217	0.217	0.217
	−1	0.710	0.710	0.710	0.710	0.710
	−1.5	1.204	1.204	1.204	1.204	1.204
	−2	1.695	1.695	1.695	1.695	1.695
CRW	+2	−1.736	−1.928	−1.979	−1.989	−1.993
	+1.5	−1.302	−1.444	−1.486	−1.493	−1.494
	+1	−0.867	−0.962	−0.990	−0.996	−0.995
	+0.5	−0.434	−0.483	−0.497	−0.498	−0.499
	−0.5	0.435	0.482	0.497	0.497	0.501
	−1	0.868	0.961	0.991	0.993	0.997
	−1.5	1.300	1.444	1.488	1.487	1.493
	−2	1.737	1.929	1.983	1.991	1.994
Average calibration accuracy		76.33%	82.74%	84.55%	84.80%	84.94%

higher calibration accuracy. When prior standard deviation σ = 0.2, the average calibration accuracy is only 76.33%. As prior standard deviation σ increases, the calibration accuracy can be improved to a certain extent. But the potential of using σ to optimize the calibration accuracy is scant. When σ increases to a certain extent, it is not obvious to improve the calibration accuracy. Table 6 shows the detailed calibration result of three target sensors for different standard deviations. It seems that the prior standard deviation has almost no effect on the calibration results of CHS sensor.

The standard deviation of the prior distribution generally depends on the prior knowledge (Yoon and Yu 2018a). But for most cases, the standard deviation of the prior distribution σ can hardly be calculated directly using the given data information from the building systems. Many researchers (Yoon and Yu 2017c, 2018a, 2018c) have explored the impact of prior distributions on the in-situ sensor calibration performance. In these studies, random error is regarded as the standard deviation of the prior distribution. However, the aforementioned studies did not provide enough details on how to select a proper value of the prior standard deviation. In practice, there are two main situations. The first is that the prior information cannot be determined by the information-poor building system. The second is that information-rich building system can easily calculate the standard deviation of the prior distribution. For practical applications of the first situation, if there is no additional prior information, the standard normal distribution usually can be used as the pre-assumed prior distribution (Wang et al. 2020), which means that σ = 1 is recommended. In this study, as shown in Figures 13 and 14, high calibration

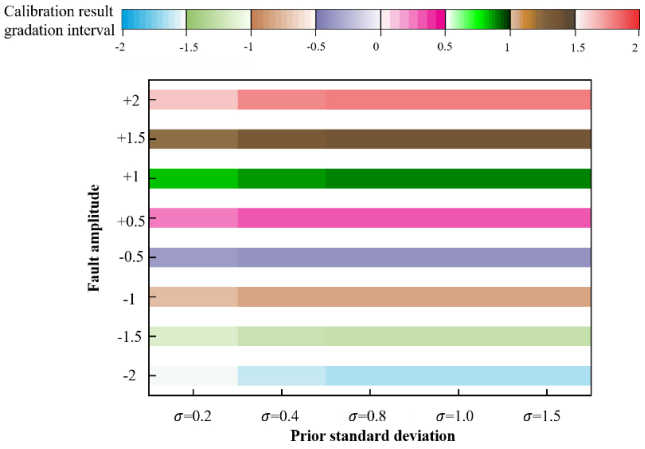


Fig. 13 Calibration results of SAT sensor for different prior standard deviations

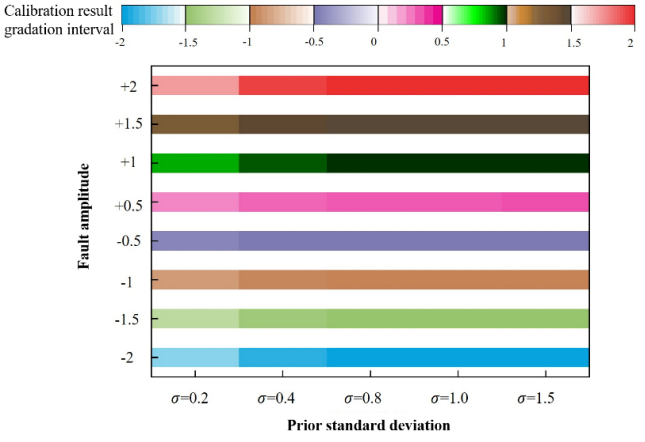


Fig. 14 Calibration results of CWR sensor for different prior standard deviations

accuracy can be achieved when $\sigma = 1$. But for the case of $\sigma > 1$, results show no more enhanced calibration accuracy. Further if $\sigma < 1$, it can be found that the calibration accuracy decreases as the calibration parameter σ drops. Especially for the case of calibration parameter $\sigma = 0.2$, results show that the lowest calibration accuracies were obtained which indicates that calibration parameter σ should not be too small. But for practical applications of the second situation, prior information of the building system is clearly presented, the standard deviation of the prior distribution σ can be calculated (Sun et al. 2010; Yoon and Yu 2018a). In such case, the calculated σ can be used as prior standard deviation. For example, the calibration parameter σ was set as 0.3 in Yoon and Yu's study (Yoon and Yu 2018a) while the calibration parameter σ was set as 10 in Sun et al.'s study (Sun et al. 2010) which is much larger than 1.

5.3.2 Energy conservation factor - energy conservation deviation degree ε(EC)

In practical building HVAC systems, the system measured data may not always satisfy the energy conservation equation well owing to the system operational dynamics, the introduced sensor bias fault amplitudes and the service control response. This may lead to certain errors in the calibration model based on the energy conservation law. This error is defined as the energy conservation factor ε(EC) which represents the deviation degree between the actual measured value and the reliable system reference value as shown in Eq. (37).

$$\varepsilon(\mathrm{EC}) = \frac{\sum_{i=1}^{n} \left| \frac{Y_{\mathrm{me},i} - Y_{\mathrm{sy},i}}{Y_{\mathrm{sy},i}} \right|}{n} \tag{37}$$

where $Y_{\mathrm{me},i}$ and $Y_{\mathrm{sy},i}$ represent the measured value of the system model and the benchmark of the reliable system model, respectively. n represents the total number of samples in the testing set. i represents the i-th of sample of the testing set. A small ε(EC) means the measured data mostly obeys the energy conservation equation, which means a high-level data quality.

In this section, the influence of ε(EC) on calibration results is discussed by changing the data volume of testing set. The sensors discussed are SAT and CHS involved in the system feedback control loop. By a random sampling manner, volumes of the testing sets are set as 1/3, 2/3 and 100% of the original testing set separately. It should be noticed that each of the three new testing data set should cover data information of the original testing data set as much as possible. Figures 15 and 16 show ε(EC) and calibration accuracy results of SAT and CHS sensors for three

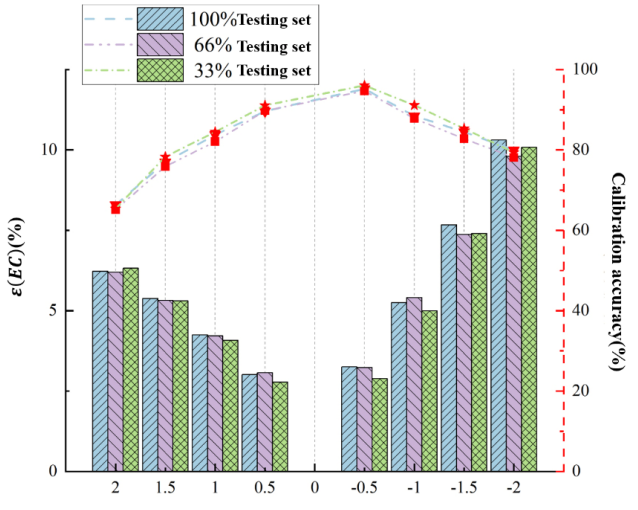


Fig. 15 Influence of different data volumes on ε(EC) and calibration accuracy for SAT (Note: left axis—ε(EC) displayed in histogram; right axis—calibration accuracy displayed in dotted lines. Horizontal axis—fault amplitudes of the sensor bias faults. Different colors in the legend represent the testing data sets with three different volumes)

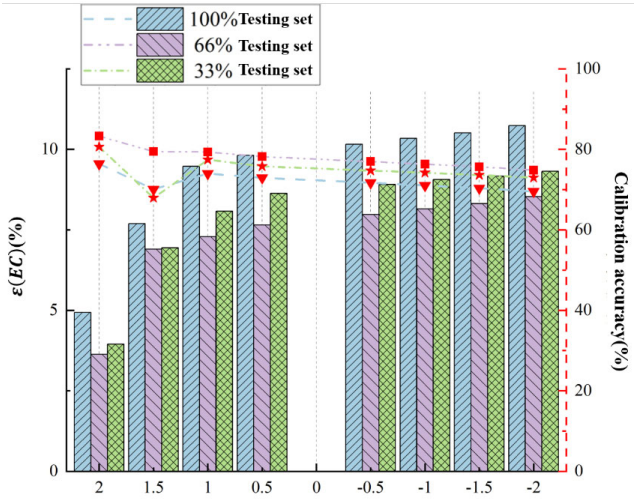


Fig. 16 Influence of different data volumes on ε(EC) and calibration accuracy for CHS

different testing sets with data volume of 1/3, 2/3 and 100% respectively.

Figure 15 shows that the data volume does not have significant impact the calibration accuracy of SAT sensor. The average calibration accuracy for three different data volumes are 83.1%, 82.02% and 83.89% respectively. For each fault amplitude, the ε(EC) could be very close to each other for three different data volumes. For SAT sensor, the data volume seems to show little influence on the ε(EC). The average ε(EC) for three different data volumes are 5.67%, 5.58% and 5.48% respectively. The fault amplitude obviously affects the ε(EC). As the absolute fault amplitude increases, ε(EC) of SAT increases to a certain extent. Also, ε(EC) of SAT shows a certain effect on the calibration accuracy. As the ε(EC) increases, calibration accuracy of SAT increases to a certain extent. Figure 16 shows that the data volume has a relatively significant impact the calibration accuracy of CHS sensor. When the data volume is 2/3 of the original testing set, the highest calibration accuracy is 78.01% on average. When the data volume is 100% of the original testing set, the calibration accuracy is relatively lower of 72% on average. Unlike the SAT sensor, different fault amplitudes have little influence on ε(EC) for the CHS sensor. Higher calibration accuracy of the CHS sensor can be obtained if there is a relatively lower ε(EC), which is consistent with the regular pattern of SAT sensor.

As shown in Figure 15 and Figure 16, the two sensors SAT and CHS show very different relations between the ε(EC) and the calibration accuracy for eight sensor bias amplitudes. The main reason could be that the SAT air temperature data are more frequently fluctuated by time than the CHS water temperature data. Specifically, for the SAT sensor, its data fluctuation phenomenon aggravates as the fault amplitude increases. From Figure 15, it can be found that the fault amplitude has a significant effect on ε(EC). ε(EC) increases with the growing absolute bias fault amplitude for the SAT sensor. But for the CHS sensor with relatively less data fluctuation, Figure 16 shows different relations between the ε(EC) and the calibration accuracy for eight sensor bias amplitudes. If small-size datasets (33%, 66%) are selected for in-situ sensor calibration, some data with relatively poor data quality may be eliminated. As a result, a smaller data set with improved data quality can be obtained for the further calibration process. This could explain why there is a big difference of the ε(EC) as the data volume changes in Figure 16.

6 Conclusions

In this study, the virtual sensor-assisted Bayesian inference (VS-BI) in-situ sensor calibration strategy is adopted to calibrate the sensor bias faults in building energy systems like HVAC. In a HVAC system including a chiller plant with air handle unit terminal, three different sensors (i.e., SAT, CHS and CWR) with a wide range of eight different levels of bias fault amplitudes are adopted to validate the strategy. The in-situ sensor calibration strategy is analyzed by considering three main influencing factors: (1) performances of virtual sensors constructed by different data-driven methods, (2) the influence of prior standard deviations σ on in-situ sensor calibration results and (3) the influence of data quality on in-situ sensor calibration from the perspective of energy conservation and data volumes. Main conclusions are as follows.

(1) Long short term memory (LSTM) is adopted for VS construction with an average determination coefficient R-squared (R^2) of 0.984 since LSTM outperforms slightly than MLR-LS and MLR-GD. Parameter optimization of LSTM can further improve the accuracy of virtual sensor.

(2) The prior standard deviation σ has a scanty influence on the sensor calibration accuracy of SAT and CWR, but has almost no effect on the calibration accuracy of CHS. The potential of using a prior standard deviation σ to improve the calibration accuracy is limited, only 8.61% on average.

(3) For SAT and CHS sensors which are involved in system feedback control loop, system operational data with relatively lower deviation degree ε(EC) means relatively higher data quality. VS-BI obtains relatively high in-situ sensor calibration accuracy if the data quality is relatively high.

Although the virtual sensor-assisted Bayesian inference in-situ sensor calibration strategy adopted in this study is validated to be capable of calibrating the sensor bias faults in building HVAC systems, there are still some limitations

of this study. Only single sensor bias faults were considered for performance evaluation. Also, the problem of low data quality was only analyzed from the perspective of energy conservation and data volumes. Considering the practical application issues, the in-situ sensor calibration strategy still needs further method validation and performance evaluation. In practical, the data quality problem may be more significant and simultaneous sensor faults are inevitable. Besides, the data limitations, algorithm computational time cost and model complexity should also be seriously considered for practical building HVAC systems. Future works would focus on the three main aspects:

(1) For practical building HVAC systems, a comprehensive strategy validation on the in-situ sensor calibration performance should be conducted by analyzing more forms of sensor faults including bias, drift, precision degradation, single and simultaneous faults, etc.

(2) Not only the in-situ sensor calibration process, but also the entire fault-tolerant process including sensor fault detection, diagnosis, reconstruction and tolerant control should be validated by simulation and experiment investigations. For simulation study, co-simulation platform should be prepared. For experiment study, both within-control-loop and out-of-control-loop types of sensors should be considered.

(3) Regarding the data limitation issue in practical building HVAC systems, it is necessary to develop a pure data-driven method to overcome the requirement of the system model item based on physical laws. By doing this, it could be promising to develop an in-situ sensor calibration strategy using only the build-in sensor measurements which can be adaptive to the practical building system situation in the field.

Acknowledgements

This work is jointly supported by the National Natural Science Foundation of China (51906181), the 2021 Construction Technology Plan Project of Hubei Province (No. 2021-83), and the Excellent Young and Middle-aged Talent in Universities of Hubei Province, China (Q20181110).

Declaration of competing interest

The authors have no competing interests to declare that are relevant to the content of this article.

Author contribution statement

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Guannan Li, Jiahao Xiong, Shaobo Sun and Jian Chen. The first draft of the manuscript was written by Jiahao Xiong and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

References

Arsie I, Cricchio A, de Cesare M, et al. (2017). Neural network models for virtual sensing of NO_x emissions in automotive diesel engines with least square-based adaptation. *Control Engineering Practice*, 61: 11–20.

Baba FM, Ge H, Zmeureanu R, et al. (2022). Calibration of building model based on indoor temperature for overheating assessment using genetic algorithm: Methodology, evaluation criteria, and case study. *Building and Environment*, 207: 108518.

Bedi J, Toshniwal D (2019). Deep learning framework to forecast electricity demand. *Applied Energy*, 238: 1312–1326.

Carey GF, Richardson WB Jr (2006). A note on least squares methods. *Communications in Numerical Methods in Engineering*, 22: 83–92.

Chen Y, Miao D (2020). Granular regression with a gradient descent method. *Information Sciences*, 537: 246–260.

Choi Y, Yoon S (2020). Virtual sensor-assisted *in situ* sensor calibration in operational HVAC systems. *Building and Environment*, 181: 107079.

Chong A, Menberg K (2018). Guidelines for the Bayesian calibration of building energy models. *Energy and Buildings*, 174: 527–547.

Coakley D, Raftery P, Molloy P, et al. (2011). Calibration of a detailed bes model to measured data using an evidence-based analytical optimisation approach. In: Proceedings of the international IBPSA Building Simulation Conference, Sydney, Australia.

Du Z, Jin X (2007). Tolerant control for multiple faults of sensors in VAV systems. *Energy Conversion and Management*, 48: 764–777.

Dudley RM (1978). Central limit theorems for empirical measures. *Annals of Probability*, 6: 899–929.

EnergyPlus (2010). EnergyPlus v9.1.0 Input Output Reference. The Encyclopedic Reference to EnergyPlus Input and Output 1996–2016.

Fan C, Xiao F, Zhao Y (2017). A short-term building cooling load prediction method using deep learning algorithms. *Applied Energy*, 195: 222–233.

Gao L, Li D, Liu X, et al. (2022). Enhanced chiller faults detection and isolation method based on independent component analysis and k-nearest neighbors classifier. *Building and Environment*, 216: 109010.

Gilks WR, Richardson S, Spiegelhalter D (1996). Introducing Markov chain Monte Carlo. In: Gilks WR, Richardson S, Spiegelhalter D (eds), Markov Chain Monte Carlo in Practic. New York: Chapman and Hall/CRC, 1995.

Guo Y, Wall J, Li J, et al. (2017). Real-time HVAC sensor monitoring and automatic fault detection system. In: Mukhopadhyay SC, Postolache OA, Jayasundera KP, et al. (eds), Sensors for Everyday Life: Environmental and Food Engineering. Cham, Switzerland: Springer International Publishing.

Hastings WK (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57: 97–109.

Hong Y, Yoon S, Kim Y-S, et al. (2021). System-level virtual sensing method in building energy systems using autoencoder: Under the limited sensors and operational datasets. *Applied Energy*, 301: 117458.

Hou D, Hassan IG, Wang L (2021). Review on building energy model calibration by Bayesian inference. *Renewable and Sustainable Energy Reviews*, 143: 110930.

Huang P, Wang Y, Huang G, et al. (2016). Investigation of the ageing effect on chiller plant maximum cooling capacity using Bayesian Markov Chain Monte Carlo method. *Journal of Building Performance Simulation*, 9: 529–541.

Kasai H (2018). SGDLibrary: A MATLAB library for stochastic gradient descent algorithms. *Journal of Machine Learning Research* 18: 215.

Kim R, Hong Y, Choi Y, et al. (2021). System-level fouling detection of district heating substations using virtual-sensor-assisted building automation system. *Energy*, 227: 120515.

Kim W, Lee J-H (2021). Fault detection and diagnostics analysis of air conditioners using virtual sensors. *Applied Thermal Engineering*, 191: 116848.

Li Z, Huang G (2013). Preventive approach to determine sensor importance and maintenance requirements. *Automation in Construction*, 31: 307–312.

Li Q, Gu L, Augenbroe G, et al. (2015). Calibration of dynamic building energy models with multiple responses using Bayesian inference and linear regression models. *Energy Procedia*, 78: 979–984.

Li G, Chen H, Hu Y, et al. (2018). An improved decision tree-based fault diagnosis method for practical variable refrigerant flow system using virtual sensor-based fault indicators. *Applied Thermal Engineering*, 129: 1292–1303.

Li G, Hu Y (2018). Improved sensor fault detection, diagnosis and estimation for screw chillers using density-based clustering and principal component analysis. *Energy and Buildings*, 173: 502–515.

Li G, Hu Y (2019). An enhanced PCA-based chiller sensor fault detection method using ensemble empirical mode decomposition based denoising. *Energy and Buildings*, 183: 311–324.

Li J, Zhao T, Wang P, et al. (2020). Effects of various partitions on the accuracy of virtual *in situ* calibration in building energy systems. *Journal of Building Engineering*, 32: 101538.

Li G, Yao Q, Fan C, et al. (2021a). An explainable one-dimensional convolutional neural networks based fault diagnosis method for building heating, ventilation and air conditioning systems. *Building and Environment*, 203: 108057.

Li G, Zheng Y, Liu J, et al. (2021b). An improved stacking ensemble learning-based sensor fault detection method for building energy systems using fault-discrimination information. *Journal of Building Engineering*, 43: 102812.

Li G, Li F, Xu C, et al. (2022a). A spatial-temporal layer-wise relevance propagation method for improving interpretability and prediction accuracy of LSTM building energy prediction. *Energy and Buildings*, 271: 112317.

Li G, Yao Q, Hu Y, et al. (2022b). Investigating thermostat sensor offset impacts on operating performance and thermal comfort of three different HVAC systems in Wuhan, China. *Case Studies in Thermal Engineering*, 31: 101788.

Liu J, Li G, Chen H, et al. (2017). A robust online refrigerant charge fault diagnosis strategy for VRF systems based on virtual sensor technique and PCA-EWMA method. *Applied Thermal Engineering*, 119: 233–243.

Liu Z, Huang Z, Wang J, et al. (2021). A novel fault diagnosis and self-calibration method for air-handling units using Bayesian Inference and virtual sensing. *Energy and Buildings*, 250: 111293.

Liu J, Li X, Li G, et al. (2022). A statistical-based online cross-system fault detection method for building chillers. *Building Simulation*, 15: 1527–1543.

Luo XJ, Fong KF (2020). Novel pattern recognition-enhanced sensor fault detection and diagnosis for chiller plant. *Energy and Buildings*, 228: 110443.

Markovic R, Azar E, Annaqeeb MK, et al. (2021). Day-ahead prediction of plug-in loads using a long short-term memory neural network. *Energy and Buildings*, 234: 110667.

Mercante R, Netto TA (2022). Virtual flow predictor using deep neural networks. *Journal of Petroleum Science and Engineering*, 213: 110338.

Mokhtari A, Ghodrat M, Javadpoor Langroodi P, et al. (2020). Wind speed sensor calibration in thermal power plant using Bayesian inference. *Case Studies in Thermal Engineering*, 19: 100621.

Pan Y, Huang Z, Wu G (2007). Calibrated building energy simulation and its application in a high-rise commercial building in Shanghai. *Energy and Buildings*, 39: 651–657.

Papadopoulos S, Azar E (2016). Optimizing HVAC operation in commercial buildings: A genetic algorithm multi-objective optimization framework. In: Proceedings of 2016 Winter Simulation Conference (WSC).

Papadopoulos PM, Lymperopoulos G, Polycarpou MM, et al. (2022). Distributed diagnosis of sensor and actuator faults in air handling units in multi-zone buildings: a model-based approach. *Energy and Buildings*, 256: 111709.

Ramos Ruiz G, Fernández Bandera C, Gómez-Acebo Temes T, et al. (2016). Genetic algorithm for building envelope calibration. *Applied Energy*, 168: 691–705.

Schleich M, Dan O, Ciucanu R (2016). Learning Linear Regression Models over Factorized Joins. In: 2016 International Conference on Management of Data.

Sendra-Arranz R, Gutiérrez A (2020). A long short-term memory artificial neural network to predict daily HVAC consumption in buildings. *Energy and Buildings*, 216: 109952.

Sinha D. (2009). Markov Chain Monte Carlo: Stochastic Simulation For Bayesian Inference (2nd Ed.). *Journal of the American Statistical Association*, 104(485): 422–423.

Sun Y, Wang S, Huang G (2010). Online sensor fault diagnosis for robust chiller sequencing control. *International Journal of Thermal Sciences*, 49: 589–602.

Sun S, Wang S, Shan K (2022). Flow measurement uncertainty quantification for building central cooling systems with multiple water-cooled chillers using a Bayesian approach. *Applied Thermal Engineering*, 202: 117857.

Vijayalakshmi K, Vijayakumar K, Nandhakumar K (2022). Prediction of virtual energy storage capacity of the air-conditioner using a stochastic gradient descent based artificial neural network. *Electric Power Systems Research*, 208: 107879.

Wang P, Yoon S, Wang J, et al. (2019). Automated reviving calibration strategy for virtual *in situ* sensor calibration in building energy systems: Sensitivity coefficient optimization. *Energy and Buildings*, 198: 291–304.

Wang Z, Hong T, Piette MA (2020). Building thermal load prediction through shallow machine learning and deep learning. *Applied Energy*, 263: 114683.

Wang P, Han K, Liang R, Ma L, Yoon S (2021). The virtual *in situ* calibration of various physical sensors in air handling units. *Science and Technology for the Built Environment*, 27: 691–713.

Yan K (2021). Chiller fault detection and diagnosis with anomaly detective generative adversarial network. *Building and Environment*, 201: 107982.

Yang X, Jin X, Du Z, et al. (2014). Optimum operating performance based online fault-tolerant control strategy for sensor faults in air conditioning systems. *Automation in Construction*, 37: 145–154.

Yoon S, Yu Y (2017a). A comparison of stochastic and deterministic optimization algorithms on virtual in-situ calibration in building systems. In: Proceedings of 2017 ASHRAE Winter Conference, Las Vegas, NV, USA.

Yoon S, Yu Y (2017b). Extended virtual *in situ* calibration method in building systems using Bayesian inference. *Automation in Construction*, 73: 20–30.

Yoon S, Yu Y (2017c). A quantitative comparison of statistical and deterministic methods on virtual *in situ* calibration in building systems. *Building and Environment*, 115: 54–66.

Yoon S, Yu Y (2018a). Hidden factors and handling strategies on virtual *in situ* sensor calibration in building energy systems: prior information and cancellation effect. *Applied Energy*, 212: 1069–1082.

Yoon S, Yu Y (2018b). Hidden factors and handling strategy for accuracy of virtual *in situ* sensor calibration in building energy systems: sensitivity effect and reviving calibration. *Energy and Buildings*, 170: 217–228.

Yoon S, Yu Y (2018c). Strategies for virtual *in situ* sensor calibration in building energy systems. *Energy and Buildings*, 172: 22–34.

Yoon S, Yu Y, Wang J, et al. (2019). Impacts of HVACR temperature sensor offsets on building energy performance and occupant thermal comfort. *Building Simulation*, 12: 259–271.

Yoon S (2020). *In-situ* sensor calibration in an operational air-handling unit coupling autoencoder and Bayesian inference. *Energy and Buildings*, 221: 110026.

Yoon S, Yu Y, Li H, et al. (2022). Improved energy balance calculation of unitary air conditioners via virtual *in situ* calibration. *Journal of Building Engineering*, 45: 103464.

Yu Y, Li H (2015). Virtual *in situ* calibration method in building systems. *Automation in Construction*, 59: 59–67.

Yuan J, Nian V, Su B (2017). A meta model based Bayesian approach for building energy models calibration. *Energy Procedia*, 143: 161–166.

Yuan T, Zhu N, Shi Y, et al. (2018). Sample data selection method for improving the prediction accuracy of the heating energy consumption. *Energy and Buildings*, 158: 234–243.

Zhang L, Leach M, Bae Y, et al. (2021). Sensor impact evaluation and verification for fault detection and diagnostics in building energy systems: A review. *Advances in Applied Energy*, 3: 100055.

Zhang B, Yan X, Liu G, et al. (2022). Multi-source fault diagnosis of chiller plant sensors based on an improved ensemble empirical mode decomposition Gaussian mixture model. *Energy Reports*, 8: 2831–2842.

Zhou Z, Li G, Chen H, et al. (2021). Fault diagnosis method for building VRF system based on convolutional neural network: Considering system defrosting process and sensor fault coupling. *Building and Environment*, 195: 107775.