A quantitative comparison of statistical and deterministic methods on virtual in-situ calibration in building systems

Sungmin Yoon, Yuebin Yu*

Durham School of Architectural Engineering and Construction, University of Nebraska-Lincoln, United States

Keywords: sensor calibration, virtual in-situ calibration, Bayesian MCMC, genetic algorithm, absorption refrigeration, building systems

Abstract:

Systematic and random errors of working sensors in building systems could significantly compromise the system's performance and thus indoor environmental quality. An extended virtual in-situ calibration has been suggested to solve problems regarding sensor errors and calibration. This calibration can correct these errors for all critical working sensors in building systems without removing working sensors or adding reference sensors as is done in a conventional calibration. This method is capable of estimating measurands using a parameter estimation technique based on mathematical system models. Deterministic and statistical methods can be used for conducting the estimation. In this study, genetic algorithm (GA)-based optimization is used as a deterministic method and Bayesian Markov Chain Monte Carlo (MCMC) is used as a statistical method to solve the calibration problem formulated by the extended virtual in-situ calibration. A case study of a single-effect LiBr-H₂O refrigeration system illustrates the problem formulating process and compares the accuracy distributions of calibrations derived from the two different methods.

1. Introduction

Energy use from heating, ventilation, air conditioning, and refrigeration systems (HVAC&R) to offset the heating and cooling load and provide a proper indoor environment for occupants constitutes a large portion of the total energy consumed in buildings [1,2]. Although there have been many studies on the optimal control strategies of air-conditioning systems aiming at energy conservation [3,4], issues, including a problematic or inferior control sequence and set points, equipment performance degradation, and various faults that occur in HVAC&R and building energy systems, further increase the energy use or lead to undesired indoor environmental quality. To address the challenges in the building sector, a comprehensive solution package, including continuous fine-tuning of building automation systems, automated analytical optimization and automated fault detection, diagnostics and repair, is needed. Research has been conducted from these perspectives to help harness the increasing energy use in buildings [5-7]. Most of the proposed approaches will be effective only if the information and data obtained from sensors are trustable and accurate [5]. For example, automated fault detection and diagnostics algorithms for devices and system may not even work if the key measurements from the sensors are wrong. With respect to smart buildings, the sensor's role is fundamentally and significantly important in ensuring and enhancing the building performance.

A virtual in-situ sensor calibration (VIC) [8] method has been recently studied in order to solve the practical problems of a conventional sensor calibration in building energy systems; the conventional calibration process and its limitations in building systems were described in [8]. Specifically, the problems addressed are: (1) time and monetary cost; (2) disruption of normal operation; (3) difficulty in accessing various sensors embedded in equipment; and (4) the large quantity of sensors [8,9]. Meanwhile,

because the working environments of a sensor is generally different from the controlled condition for a physical calibration, the systematic errors associated with the system working conditions cannot be solved even with a perfect conventional calibration. The proposed in-situ calibration method is able to approximate the measurement values and establish benchmark values for the calibration. This is done either statistically or by using system models without removing the working sensor or adding reference sensors, as in a conventional calibration. Since the new method is conducted in-situ, it is capable of handling the systematic errors (offset) associated with changing working conditions, as well as the random errors (noise) due to sensor degradation.

We further extended the novel in-situ sensor calibration method by casting it as a parameter estimation problem in order to calibrate all working sensors simultaneously in a circumstance where we are unable to certify which sensors are malfunctioning. This extension is beyond the scope of the existing VIC. The extended virtual in-situ calibration (EVIC) introduces reliable system output variables and multiple sets of measurements to estimate true measurements [9,10]. Deterministic and statistical approaches can be applied to solve the formulated calibration problem created by EVIC. For a model calibration in the building sector, the deterministic approach uses an optimization process to derive optimal values of unknown parameters by minimizing the objective function consisting of differences between measured and calculated values [11-13]. The statistical approach (Bayesian calibration) calculates the probability density functions of the estimated unknown parameters, reducing the difference between simulated and observed data by considering the inherently stochastic nature of input variables [14-17].

This study uses the two methods for comparing their results in EVIC: (1) GA-based optimization as the deterministic approach; and (2) Bayesian calibration as the statistical approach. With the understanding of different ways to suggest calibrated results in two methods, we focus more on an effect of the difference in prior information of two methods on their calibration results. The deterministic method uses lower and upper bounds of unknown variables for the prior information while probability density functions are used for prior distributions in the statistical calibration. These priors may be collected by measured data from working sensors. Of note, the measurements can illustrate possible random errors from standard deviations, but are unable to reveal systematic errors. This is because the systematic errors cannot be detected without ways to discover whether values are accurate. Nevertheless, there is no alternative but to use the measurements for prior information before a calibration. In such cases, we contemplate the pros and cons of each method intuitively. The disadvantage of GA is that it requires wider ranges than ranges established from measured data in order to correct the large systematic errors occurring outside the ranges of measurements, because the optimal variables are estimated within the given ranges. However, once the extended ranges are defined properly, GA is unaffected by the relationship between systematic and random errors because the ranges have an identical probability regardless of standard deviations of random errors. On the other hand, Bayesian calibration does not need any extension for prior information because it is able to explore and derive the accurate values even at a low probability of the prior distribution. In a proportional relationship between the two types of errors, the prior distributions can be more informative and thus function as related equations that help the calibration problem to be more constrained for a determined condition. But, uninformative prior distributions can adversely affect the accuracy of calibration results.

The next question of this study is how effectively calibration results are improved by the use of multiple sets of measurements within the two methods. If it is very effective, are the accuracy distributions according to different numbers of sets similar or dissimilar in different types of working

sensors such as temperatures and mass flow rates? Such an analysis helps to determine the sufficient number of measurement sets that can satisfy each reference value for accuracy by considering the different sensor types. Otherwise, how do we solve the calibration problem appropriately when the accuracy is not sufficiently improved by the multiple sets? This study has proposed a two-step calibration approach for EVIC to handle this issue.

We present the procedure of problem formulation as well as the calibration results for a single-effect LiBr-H₂O refrigeration system. The results from deterministic and statistical methods are compared with the pre-defined true values (answers) from a forward simulation method [18]. Then, accuracy distributions of the calibrated results with a different number of measurement sets are identified in the two methods in order to compare them quantitatively and the reasonable number of measurement sets in deriving the accurate values using two algorithms is discussed. Finally, a more proper method for the suggested calibration problem, if possible, is recommended.

2. Extended virtual in-situ calibration in building systems

2.1. EVIC problem formulation

The EVIC method proposed by Yoon and Yu [9] can calibrate multiple working sensors simultaneously. The EVIC problem has a distance function describing the squared sum of all working sensor errors; it represents the difference between correction functions and benchmarks for working sensors in a system, as shown in Eq. 1. It should be minimized through the suggested calibration process. Each correction function, as in Eq. 2, can be formulated by offsetting constants and/or correction coefficients along with a measurement, which will compensate for the systematic errors for a specific sensor over all working stages. The benchmarks, which are modeled by the relative variables and unknown parameters in a system model, are considered as reference values for calibration, as in Eq. 3. Since the relative variables may be determined by uncertain measurements from other working sensors, they are also defined by the corresponding correction functions in the benchmark function. Thus, the offsetting constants and correction coefficients of the formulated correction function, and the unknown parameters of the benchmark, are selected as variables of interest in this calibration problem. These variables can be estimated by minimizing the defined distance function through the EVIC process.

$$D(x) = \sum_{i=1}^N (Y_{bi} - Y_{ci})^2 \quad (1)$$

$$Y_c = g(M, x_1, x_2, \dots, x_k) \quad (2)$$

$$Y_b = f(Y_{c,v1}, Y_{c,v2}, \dots, Y_{c,vr}, x_{u1}, x_{u2}, \dots, x_{uq}) \quad (3)$$

where, D is the distance function regarding the measurement error, x represents the variables of the EVIC problem, Y_b is the corresponding benchmark, Y_c is the corrected measurement, i is the counter for working sensors, N is the number of benchmarks in distance function, g is the correction function, M is the measurement from the working sensor, k is the counter for the variables, f is the system model, v is the relative variable of the measurement M , r is the counter for the relative variables, x_u is the unknown parameter in the system model, and q is the count for the unknown parameters.

This calibration constructed by the distance function as in Eq. 1 is naturally an under-determined problem; many different combinations of the variables can be derived from the calibration even though

they may differ from their true values. In order to obtain the accurate variables, the under-determined calibration problem should be constrained by decreasing the number of unknown conditions and/or increasing the number of related equations within the calibration system. Using multiple sets of steady-state measurements (S) from various timestamps may turn the calibration problem into a determined one through adding new relative equations. To this end, a formulation of the distance function is rewritten as Eq. 4. The right term on the right side denotes a form of the general model calibration that calibrates uncertain input variables from reliable output variables. The left term on the right side represents the sensor calibration. In EVIC, the model calibration and the sensor calibration are conducted simultaneously and accurate estimates of the unknown parameters help to derive the true offsetting constants and correction coefficients and vice versa.

$$D(x) = \sum_{s=1}^S \left[\sum_{i=1}^N (Y_{bi} - Y_{ci})^2 + \sum_{h=1}^H (Y_{bh} - Y_{Rh})^2 \right] \quad (4)$$

where, Y_R is the known system output values, h is the counter for the reliable output variables, H is the number of reliable output variables, and S is the number of measurement sets.

2.2 Proportional and disproportional error types

This calibration method handles both major types of sensor errors – random errors (precision) and systematic errors (accuracy). Random errors are indicated by the difference between measurements and their mean. They are caused by external problems that affect the sensor readings or hardware noise. The random errors can be detected by their measurement data and described as probability distributions, such as Gaussian, in some cases. A random error with a large standard deviation indicates measurement inaccuracies (low precision). Systematic errors are indicated by the discrepancy between the mean of measurements and true measurement values. They cannot be identified without ways to discover true values and thus present a more challenging problem than random errors. They may occur because of the sensor's physical condition, measured phenomena, working environments, or other factors.

A working sensor has various combinations regarding random and systematic errors in a real operational system because the two different errors are not essentially related to each other. Conceptually, Figure 1 shows the possible combinations that could happen in a working sensor and two error types are defined in this study – proportional error type and disproportional error type. In a proportional error type, a magnitude of systematic errors is proportional to the size of random errors. Systematic error can be involved in the range of random errors. On the contrary, it is difficult for the systematic error to be included in the random error ranges in the disproportional error type because the systematic and random errors are inversely related. In this study, the suggested EVIC using two algorithms will be applied to these two error types and their results will be discussed, respectively, in Section 4.

Figure 1. Classification of sensor error types in EVIC.

2.3. Calibration approaches (local and two step approaches)

EVIC has suggested two calibration approaches: (1) whole calibration and (2) local and whole calibration (two-step calibration). Figure 2 illustrates the suggested calibration approaches. In the whole calibration, a distance function is formulated for an entire system. A calibration is conducted for the whole system using the defined distance function. On the other hand, several distance functions are constructed for sub-components of a whole system in the two-step calibration. The components can be defined as each local calibration area based on relationships between input and output variables of the models, so that the equation system (distance function) of each local calibration can be closed as much as possible. The results from the local calibration may be more informative than prior knowledge used for only whole calibration and each local calibration. In other words, in a two-step calibration, a role of the local calibration is to reliably update the prior information for later whole calibration by mathematically closing the equation system of each local area.

Figure 2. Diagram defining calibration domain and two different calibration approaches.

3. Methods

Figure 3 illustrates how we quantify the accuracy distributions of two different methods. Section 3.1 formulates the EVIC problem including the suggested distance functions for the whole and two-step calibration. Sections 3.2 and 3.3 apply both Bayesian calibration and GA-based optimization to the whole and two-step calibration approaches, respectively. To obtain the accuracy distributions, different numbers of measurement sets, from 1 to 10, are fed into the Bayesian MCMC and GA-based optimization algorithms. The calibration performance is then indexed based on the number of measurement sets for the comparison in Section 4.

Figure 3. Illustration of the methods used and input & output data in this study.

3.1. EVIC problem formulation for ARS

This paper introduces a simulated single-effect LiBr-H₂O absorption refrigeration system (ARS) for a calibration case, as shown in Figure 4. Various sensors for temperature, pressure, mass flow rate, etc., are used in the ARS to provide information on the system's performance. Its steady-state thermodynamic models were developed based on the previous studies (as listed in Table 1), which were established in MATLAB (ver. 2015b). This model has the following calculation process: (1) solution concentrations, (2) temperatures of states 5 and 3, (3) enthalpies except for state 3, (4) mass flow rates of solutions and refrigerants, (5) enthalpy of state 3, and (6) heat transfer rate at each component and pump work. In the previous study [9], the results calculated by these models showed a good agreement (deviations less than 5.5%) with the results from other researches [19,20].

Figure 4. Flow diagram of the single-effect LiBr-H₂O refrigeration system.

Table 1. Reference models of the ARS.

Building system	A single-effect LiBr-H ₂ O AR system
Steady-state model	Pressures of the system [21] Refrigerant (water) properties [22] Absorbent (LiBr) concentration [23] Solution (LiBr-water) properties [23] Density of strong solution [24] Solution pump work [25]
Input	Absorber temperature, T_1 Generator temperature, T_7 Condenser temperature, T_8 Evaporator temperature, T_{10} Solution heat exchanger effectiveness, ε Evaporator cooling load, $Q_e^{*1,m}$
Output	Condenser heat flow rate, $Q_c^{*1,m}$ Generator heat flow rate, $Q_g^{*2,m}$ Absorber heat flow rate, $Q_a^{*2,m}$ Solution pump energy consumption, $W_p^{*1,m}$ Temperatures, $T_3, T_5^{*2,s}$ Mass flow rates for the strong solutions, $m_{1\sim3}^{*2,s}$ Mass flow rates for the weak solutions, $m_{4\sim6}^{*2,s}$ Refrigerant mass flow rates, $m_{7\sim10}$ LiBr concentrations in strong solution, $X_{1\sim3}$ LiBr concentrations in weak solution, $X_{4\sim6}$

*1: 1st local area, *2: 2nd local area, m : benchmark in the model calibration term, and s : benchmark in the sensor calibration term

The EVIC was applied to working sensors for temperatures and mass flow rates at the inner ARS (states 1-10). In this problem, as in Table 2, various sensor errors are designed to suggest the proportional and disproportional error types; their true measurements are pre-determined by using the forward

simulation method of the mentioned steady-state system models [18]. The systematic and random errors for mass flow rates were defined in accordance with their differences in scale.

Table 2. Description of sensor errors.

Working sensor	Proportional type		Disproportional type	
	Systematic error	Random error (standard deviation)	Systematic error	Random error (standard deviation)
T_1	-5°C [Large]	2 [Large]	-5°C [Large]	0.7 [Normal]
T_3	-	0.7 [Normal]	-	1.0 [Large]
T_8	-3°C [Normal]	1.5 [Large]	-3°C [Normal]	0.7 [Normal]
Remaining temperatures	-	0.7 [Normal]	-	0.7 [Normal]
$\dot{m}_{7-10}$	0.0075 kg/s, [Normal] ^{*1}	0.005 [Large] ^{*2}	0.0075 kg/s [Normal] ^{*1}	0.0025 [Small] ^{*3}
Remaining mass flow rates	-	0.035 [Normal] ^{*4}	-	0.035 [Normal] ^{*4}

*1: The error is about 8.7%, *2: It can be considered 1.0 in the temperature scale, *3: It can be considered 0.5 in the temperature scale, and *4: It can be considered 0.7 in the temperature scale.

The entire system was divided into two parts for a local calibration based on the relations between input and output variables in the system model. The first local calibration includes the evaporator, condenser, and solution pump. The second local calibration is conducted for the absorber and generator. The distance function of each calibration area, as in Eq. 4, was comprised of the benchmarks and correction functions. The correction functions were determined as linear equations having only offsetting constants for all working sensors, as shown in Tables 3 and 4. The known coefficients (a and b) in Tables 3 and 4 were introduced into the correction functions of mass flow rates in order to match their prior distributions in scale to those of temperatures. At the first local calibration area, the benchmarks were formulated for the evaporator cooling load, condenser heat flow rate, and pump energy consumption in the model calibration term of the distance function (the second term on the right side in Eq. 4). At the second local domain, the benchmarks of the absorber heat flow rate and generator heat flow rate were defined in the model calibration term of its distance function. The benchmarks for the temperatures T_3 and T_5 , and the mass flow rates m_{1-3} and m_{4-6} were involved in the sensor calibration term (the first term on the right side in Eq. 4) at the second local domain. For the whole calibration, all of the benchmarks above were included in one distance function. Offsetting constants for all temperature and mass flow rate sensors and a solution heat exchanger effectiveness, ε (unknown parameter), were defined as the variables of this calibration. Each number of sets of measurements (S) from 1 to 10 was set in the defined distance function of each calibration domain. With each set, the calibration was conducted 10 times.

3.2. Statistical calibration using Bayesian MCMC (Bayesian calibration)

According to Bayes' theorem, as in Eq. 5, Bayesian inference derives a posterior distribution, $P(x|Y_b)$, of variables (x), including offsetting constants, correction coefficients and unknown parameters, based on a likelihood function, $P(Y_b|x)$, and prior distributions of the variables, $\pi(x)$. As in Eq. 7, the likelihood functions for each local and whole calibration are formulated by plugging the distance function of each measurement set into a Gaussian distribution individually. Then, the sum of each set, found by Eq. 4, is converted into the phi in order to express the product of the individual Gaussian distribution of each set in the likelihood function. The defined likelihood function can be maximized when the distance

functions of each set are minimized. Thus, the posterior distributions may have a higher probability if the value of the distance function is smaller.

$$P(x | Y_b) = \frac{P(Y_b | x) \times \pi(x)}{P(Y_b)} \quad (5)$$

$$P(Y_b) = \int P(Y_b | x) \pi(x) dx \quad (6)$$

$$P(Y_b | x) = \prod_{s=1}^S \frac{1}{\sigma \sqrt{2\pi}} \exp \left[-\frac{1}{2\sigma^2} \left(\sum_{i=1}^N (Y_{bi} - Y_{ci})^2 + \sum_{h=1}^H (Y_{bh} - Y_{Rh})^2 \right) \right] \quad (7)$$

where, x represents the variables, Y_b is the benchmark, $\pi(x)$ is the prior distribution of the variables, $P(x | Y_b)$ is the posterior distribution of the variables, $P(Y_b | x)$ is the likelihood function, $P(Y_b)$ is the normalizing constant, and σ is the standard deviation.

The prior distributions of the defined variables in the two calibration approaches suggested in this study were assumed to obey a normal distribution according to the Central Limit Theorem [26]. The means and standard deviations of the normal distributions can be defined by measurements from working sensors. The measurements can reveal random errors of each working sensor, but do not reveal information about systematic errors because the systematic error is not directly related to the random error as discussed in Section 2.2. However, we have no choice but to use the measurements for the prior information for the offsetting constants before a calibration. Accordingly, the standard deviations of normal distributions were defined by the standard deviation values (size of random errors) of measurements, and the means of every normal distribution were set at zero, assuming the raw measurements are reliable. For the later whole calibration, the prior distributions are defined by the posterior distributions derived from the former local calibration. Meanwhile, expert knowledge and previous information from literature can be used for the prior distributions of unknown parameters. The prior distribution of the solution heat exchanger effectiveness was defined as in Table 3.

The defined prior distributions based on the two error types have different characteristics. The prior distributions from the proportional type are more advantageous to correct the systematic errors than those of the disproportional type because the proportional relationship gives much information naturally about systematic errors. In this case study, two different types of prior distributions were defined based on each error type, as in Table 2, and used to identify different characteristics of the two cases in Bayesian calibration. This study called the two cases a proportional Bayesian and a disproportional Bayesian, respectively.

In the Bayesian inference, it is significantly difficult to calculate the normalizing constant, as in Eq. 6, using an analytical approach. This study used a MCMC method to derive a posterior distribution without calculating the integration. The MCMC method is Monte Carlo integration using Markov Chains [27,28]. To run the MCMC method, the Metropolis-Hastings algorithm [29,30] was used, which is a popular sampling technique for obtaining a sequence of random samples from joint multivariate distributions. The previous researches have showed the effectiveness of the MCMC method, in calculating the reliable posterior distributions of their Bayesian calibration, with reasonable iterations and a period of burn-in [15-17]. In this study, the number of iterations was defined as 5,000 with the burn-in

of 1,000 iterations to ignore the beginning of iterations, in which calibration variables do not converge. This process, using the Bayesian inference with MCMC method, was implemented in MATLAB.

3.3. Deterministic calibration using GA-based optimization

Alternatively, the deterministic calibration can determine the optimal values of design variables using an optimization process. The objective function for the deterministic calibration can be defined with the corresponding distance function, as shown in Eq. 8, for each local calibration of the two-step method and for the whole calibration. Design variables of the optimization problem consist of offsetting constants, correction coefficients, and unknown parameters as in Bayesian calibration.

$$\min_x D(x) = \sum_{s=1}^S \left[\sum_{i=1}^N (Y_{bi} - Y_{ci})^2 + \sum_{h=1}^H (Y_{bh} - Y_{Rh})^2 \right] \quad (8)$$

Unlike prior distributions of offsetting constants, correction coefficients, and unknown parameters in the Bayesian calibration, the ranges of these variables in the deterministic calibration are constrained by their lower and upper bounds, and the optimal values are calculated within definite ranges even though systematic errors happen outside these ranges. Accordingly, it is actually impossible to directly use the measurements from working sensors as the ranges. In this case study, the ranges are widely established regardless of the two error types to get over this issue. This is shown in Table 3. The whole calibration in the two-step calibration uses the minimum and maximum values within the results from each local calibration to obtain its more reliable ranges, as shown in Table 4.

The specific optimization algorithm such as a gradient-based or evolutionary algorithm is needed. GA is a better optimization method for a global optimum especially when a problem is not perfectly smooth and uni-modal [31,32]. For this reason, GA has frequently been applied to identify the optimum values of such a problem regarding the deterministic calibration in the field of building airflow or energy simulation having various models [11-13]. Considering the EIVC problem consisting of multiple models (nonlinear equations) for each system component, the GA can provide the better global solution (i.e. near optimal solution) for the deterministic EVIC, as in the building model calibration. This study uses the GA provided by the optimization toolbox in MATLAB.

4. Results and discussion

4.1. Effectiveness of the use of multiple measurement sets for two methods in a whole calibration approach

A calibration accuracy was evaluated by root mean square error (RMSE), which indicates the average error of calibrated measurements. The RMSE was calculated based on the difference between calibrated measurements and the pre-defined true measurements from the forward simulation method. In Bayesian calibration, the raw measurements were calibrated by the estimated variables that have the highest probability in their posterior distributions and their correction functions after the calculation of Bayesian MCMC. The accuracy distributions for the two methods were calculated from 100 calibrated results (10 times per set * 10 sets) using a boxplot that presents quartiles and bounds. The width of a boxplot can show the precision of the calibrated measurements; a lower width has a higher precision. These distributions were grouped for temperature, mass flow rate, and the unknown parameter (solution heat exchanger effectiveness) due to the consideration on their different scales, as shown in Figures 5 and 9. The sufficient number of measurement sets that can provide an acceptably high accuracy was determined

considering the error characteristics of different variables; the temperature value was $\pm 0.5^{\circ}\text{C}$ (error) and value for the mass flow rate and the unknown parameter was $\pm 5\%$ (error).

All temperatures and mass flow rates were satisfied with their reference values even at one set in the proportional Bayesian. The results of the unknown parameter were significantly accurate at one set and using more than seven sets provided the satisfactory results. In the disproportional Bayesian, the sufficient numbers in terms of the mass flow rates and the unknown parameter were three and five, respectively, while the calibrated results of temperatures were not satisfied with the reference even at ten sets. For the temperature results from GA, the reference value was also not met by the calibrated results at any number of sets. From more than two sets, most results of the mass flow rates and the unknown parameter were lower than the reference value. Considering the calibration accuracy of the three variables, the most reliable results were calculated by the proportional Bayesian based on its favorable prior distributions. In addition to that, the accuracy distributions and the sufficient number for the three variables were different from each other in all cases. Consequently, we should consider the different distributions and sufficient numbers when determining a reasonable one among the numbers for a whole system.

The RMSEs of mass flow rates and an unknown parameter in three cases were enhanced by increasing the number of multiple sets. The temperature RMSEs were not improved except for those in the disproportional Bayesian, as shown in Figure 5. These differences were related to the characteristics of system models. For example, when two unknown terms such as a mass flow rate and an enthalpy difference (or a pressure difference) are multiplied by each other and we know the calculated output values in multiple cases, we can estimate the accurate values of two terms. However, it is impossible to calculate each enthalpy or pressure, which is related to the corresponding temperature, in the difference term by only using the multiple cases without additional information about either of them. When we do not know the individual enthalpy or pressure value, the corresponding temperature cannot be estimated. That is to say, the mass flow rates and the unknown parameter that were multiplied by a temperature or by the difference term were more constrained by adding multiple measurement sets while the multiple sets were redundant for some temperatures in their difference terms according to the characteristic of building system models. As shown in Figures 6-8, all distributions of mass flow rates and the temperature distribution of T_8 in three cases converged toward their true values, whereas temperature distributions of T_4 and T_{10} were erratic or diverged from their true values. In such a case where some temperatures were unknown, the extended ranges of GA made the temperature RMSE distributions most irregular and they had a negative effect on the results of the related mass flow rates and the unknown parameter, as shown in Figures 5 and 8. This will be discussed in more detail in the following section.

Nevertheless, the temperature RMSEs of the disproportional Bayesian showed a slightly improved distribution. As shown in Figure 7, the temperature T_1 along with the temperature T_8 converged toward their true values only in the disproportional Bayesian, thus making the RMSE distribution curved for an improvement. Since many temperatures were already accurate at one set using good prior information, the temperature RMSEs of the proportional Bayesian did not represent the improved distribution. In other words, the calibration results were more enhanced through a likelihood function which includes the multiple sets in the disproportional Bayesian where the priors are not sufficiently beneficial to estimate accurate results. For this reason, RMSE distributions of temperatures, mass flow rates, and an unknown parameter in the disproportional Bayesian were more clearly improved than those in the proportional Bayesian. But, the overall accuracy of calibrated results from the proportional Bayesian was better than that of the disproportional Bayesian, particularly in temperature distributions for

both of normal and erroneous sensors. Considering various error combinations, including the proportional and disproportional error types in real operational systems, the temperature results of the disproportional Bayesian must be more accurate (at one set) and enhanced from the initial value by the use of multiple sets. Also, as shown in Figure 6, the distributions toward the erroneous values, such as T_4 and T_{10} in the proportional Bayesian, should be changed to the true values.

Figure 5. RMSE distributions from the three cases in the whole calibration.

Figure 6. Calibrated results from the proportional Bayesian in the whole calibration (Case 1).

Figure 7. Calibrated results from the disproportional Bayesian in the whole calibration (Case 2).

Figure 8. Calibrated results from the GA in the whole calibration (Case 3).

4.2. Differences among three cases (two methods) in a two-step calibration

This study used the suggested two-step calibration by combining local and whole calibration processes to overcome the limitations of whole calibration. As shown in Figure 9, in the proportional Bayesian, the RMSEs of three variables were lower than the reference values from one set and the temperature RMSEs were lower than those of the whole calibration. In the disproportional Bayesian, all calibrated results of three variables from more than three sets were satisfied with each reference unlike the unsatisfactory results of temperatures from the whole calibration approach. The RMSEs of temperatures and mass flow rates dropped sharply at two sets. With GA, using more than two sets was sufficient to get satisfactory results about the mass flow rates and the unknown parameter. For temperatures, the results

were better than those of the whole calibration, but they were still not close to the reference value. When comparing all distributions of three cases, as described in the whole calibration (section 4.1), the proportional Bayesian showed the most accurate results, especially the temperatures and mass flow rates. The well-improved distributions by the use of multiple sets were suggested in the disproportional Bayesian. In the GA, the calibration results also got better more obviously through the two-step calibration method with multiple sets. Meanwhile, the main difference between the whole and two-step methods was the improvement of temperature distributions of all cases.

Such an improvement from the two-step calibration was attributed to the more informative priors derived from each local calibration of two methods, as shown in Table 4. In the first local area, each variable, including the temperature and mass flow rate, was completely constrained to estimate the more accurate value of the variables by using multiple sets. Then, the estimated values of the first local area were used as known values of the second local calibration. These local calibration processes allowed the calibration problem of each area to be more mathematically closed, thus accurately estimating variables regardless of reliability of their priors. Finally, the later whole calibration of all cases could be conducted by the informative priors as in the proportional type.

The gap in results from the two methods was related to two differences between prior distributions and the ranges for variables. As stated before, the first is that it is impossible for the variables to be derived outside the defined ranges in the deterministic method. In order to cover the large systematic errors, which may exist outside the initial ranges that are defined by measurements (size of random error), the ranges should be more widely extended in the GA. In this study, the ranges were extended, as shown in Tables 3 and 4. Conversely, Bayesian calibration was able to estimate accurate variables even at a very low probability (almost zero) in the prior distributions without any extension (e.g. T_1 in the proportional and disproportional Bayesian). In summary, GA-based calibration is unavoidably conducted assuming that all systematic errors for working sensors may be outside the ranges of the random errors (a negative perspective), to overcome the limitation regarding the variable ranges. Bayesian calibration is based on a temporary assumption that every systematic error exists within the ranges of the random errors (a positive perspective). These different perspectives can be identified by the different distribution directions toward their true values. For example, most distributions (T_1 , m_{7-10}) were improved from zero in the Bayesian calibration, as shown in Figures 10 and 11, but from the larger absolute values in GA as shown in Figure 12.

The second difference resulting in these gaps is how informative priors are in the two methods. Prior distributions of the Bayesian method can have a probability density function where it is more beneficial to constrain the distance function than the simple ranges of GA because the prior distributions play an important role in adding the related equations into the calibration problem. Although the prior distributions are uninformative in terms of their systematic errors as in the disproportional error type, the results were more reliable than those from GA. This was because the EVIC method relied on the basic idea that most working sensors of a system are correct; the normal sensors having only random errors outnumbered the wrong sensors about systematic errors in this case study. This may be regarded as a general case in a real operational system. For these two reasons, in the under-determined calibration problem, the accuracy distributions from the GA were less obvious and reliable than those of Bayesian MCMC, especially in the normal sensors (e.g. T_3 , T_4 , T_5 , and T_{10} in Table 4 and Figures 10-12).

Finally, Tables 3 and 4 compared each result of two methods derived from the sufficient number of sets (5 sets) in order to show how different their results were.

Figure 9. RMSE distributions from the three cases in the two-step calibration.

Figure 10. Calibrated results from the proportional Bayesian in the two-step calibration (Case 4).

Figure 11. Calibrated results from the disproportional Bayesian in the two-step calibration (Case 5).

Figure 12. Calibrated results from the GA in the two-step calibration (Case 6).

Table 3. A comparison of calibrated results (median) from two methods using 5 sets in the two-step calibration in the whole calibration

Sensor	Correction function and variable x	Virtual condition ^{*3}		Proportional Bayesian ^{*3} (Case 1)					Disproportional Bayesian ^{*3} (Case 2)					GA-based optimization ^{*3} (Case 3)		
				Before a calibration (prior distribution)		After a calibration (posterior distributions)			Before a calibration (prior distribution)		After a calibration (posterior distributions)			Before a calibration	After a calibration	
		True value of sensor	True value of x	Mean	Standard deviation	Mean	Standard deviation	Calibrated (deviation)	Mean	Standard deviation	Mean	Standard deviation	Calibrated (deviation)	Ranges	Optimal value	Calibrated (deviation)
T_1^* ($=T_2$)	$T_{c,1}=T_1+x_1$	32.5	-5	0	2	-4.21	1.32	33.29 (0.79°C)	0	0.7	-1.98	0.62	35.52 (3.02°C)	[-10,10]	-4.10	33.40 (0.90°C)
T_3	$T_{c,3}=T_3+x_3$	56.94	0	0	0.7	0.08	0.78	57.01 (0.07°C)	0	1	-0.04	1.00	56.90 (0.04°C)	[-5,5]	-3.82	53.12 (3.82°C)
$T_4(=T_7)$	$T_{c,4}=T_4+x_4$	82.5	0	0	0.7	0.54	0.78	83.04 (0.54°C)	0	0.7	1.49	0.62	83.99 (1.49°C)	[-5,5]	-0.81	81.69 (0.81°C)
$T_5(=T_6)$	$T_{c,5}=T_5+x_5$	52.5	0	0	0.7	-0.15	0.79	52.35 (0.15°C)	0	0.7	0.24	0.67	52.74 (0.24°C)	[-5,5]	4.95	57.45 (4.95°C)
T_8^* ($=T_9$)	$T_{c,8}=T_8+x_8$	42.5	-3	0	1.5	-2.68	0.64	42.82 (0.32°C)	0	0.7	-2.27	0.45	43.23 (0.73°C)	[-5,5]	-2.18	43.32 (0.82°C)
T_{10}	$T_{c,10}=T_{10}+x_{10}$	9	0	0	0.7	0.41	0.81	9.41 (0.41°C)	0	0.7	1.82	0.56	10.82 (1.82°C)	[-5,5]	2.20	11.20 (2.20°C)
m_r^* ($=m_{7-10}$)	$m_{c,r}=m_r+ax_{mr}$	0.086	1.5	0	1	1.320	0.351	0.085 (1.05%)	0	0.5	1.221	0.230	0.085 (1.62%)	[-5,5]	1.336	0.085 (0.95%)
$m_1(=m_{1-3})$	$m_{c,1}=m_1+bx_{m1}$	0.795	0	0	0.5	-0.190	0.488	0.786 (1.16%)	0	0.7	-0.253	0.452	0.783 (1.55%)	[-5,5]	-0.490	0.771 (3.04%)
$m_4(=m_{4-6})$	$m_{c,4}=m_4+bx_{m4}$	0.709	0	0	0.5	0.256	0.579	0.722 (1.85%)	0	0.7	0.743	0.463	0.746 (5.28%)	[-5,5]	-0.481	0.685 (3.34%)
ε^{*2}	x_{ul}	-	0.6	0.75	0.07	0.603	0.218	0.603 (0.42%)	0.75	0.07	0.634	0.053	0.634 (5.63%)	[0,1]	0.501	0.501 (16.49%)

*: erroneous sensors, *2: unknown parameter, *3: temperatures (°C) and mass flow rates (kg/s), a is 1/200, and b is 1/20.

Table 4. A comparison of calibration results (median) from two methods using 5 sets in the two-step calibration

Sensor	Correction function and variable x	Virtual condition ^{*3}		Proportional Bayesian ^{*3} (Case 4)					Disproportional Bayesian ^{*3} (Case 5)					GA-based optimization ^{*3} (Case 6)		
				After a local calibration (prior distribution) ^{*4}		After a whole calibration (posterior distributions)			After a local calibration (prior distribution) ^{*4}		After a whole calibration (posterior distributions)			After a local calibration	After a whole calibration	
		True value of sensor	True value of x	Mean	Standard deviation	Mean	Standard deviation	Calibrated (deviation)	Mean	Standard deviation	Mean	Standard deviation	Calibrated (deviation)	Ranges ^{*4}	Optimal value	Calibrated (deviation)
T_1^* ($=T_2$) ^{L2}	$T_{c,1}=T_1+x_1$	32.5	-5	-4.76	0.31	-4.91	0.53	32.59 (0.09°C)	-4.31	0.27	-4.64	0.60	32.86 (0.36°C)	[-5.72,-0.81]	-4.70	32.80 (0.30°C)
T_3 ^{L2}	$T_{c,3}=T_3+x_3$	56.94	0	0.03	0.74	-0.02	0.39	56.92 (0.02°C)	-0.02	1.04	0.00	0.40	56.94 (0.00°C)	[-0.57,4.97]	1.55	58.48 (1.54°C)
$T_4(=T_7)$ ^{L1}	$T_{c,4}=T_4+x_4$	82.5	0	0.05	0.63	0.04	0.37	82.54 (0.04°C)	0.03	0.60	0.09	0.41	82.59 (0.09°C)	[-1.06,3.26]	1.26	83.76 (1.26°C)
$T_5(=T_6)$ ^{L2}	$T_{c,5}=T_5+x_5$	52.5	0	-0.04	0.66	-0.05	0.41	52.45 (0.05°C)	0.03	0.66	-0.01	0.41	52.49 (0.01°C)	[-1.84,2.86]	-0.19	52.31 (0.19°C)
T_8^* ($=T_9$) ^{L1}	$T_{c,8}=T_8+x_8$	42.5	-3	-3.08	0.36	-3.05	0.48	42.45 (0.05°C)	-2.79	0.30	-2.97	0.47	42.53 (0.03°C)	[-3.66,-2.99]	-3.32	42.18 (0.32°C)
T_{10} ^{L1}	$T_{c,10}=T_{10}+x_{10}$	9	0	-0.01	0.61	0.03	0.38	9.03 (0.03°C)	0.01	0.61	0.06	0.37	9.06 (0.06°C)	[-1.90,2.31]	-1.00	8.00 (1.00°C)
m_r^* ($=m_{7-10}$) ^{L1}	$m_{c,r}=m_r+ax_{mr}$	0.086	1.5	1.435	0.187	1.493	0.212	0.086 (0.04%)	1.360	0.174	1.387	0.221	0.085 (0.66%)	[1.43,1.50]	1.447	0.086 (0.31%)
$m_1(=m_{1-3})$ ^{L1}	$m_{c,1}=m_1+bx_{m1}$	0.795	0	0.085	0.360	0.066	0.343	0.799 (0.45%)	-0.217	0.303	-0.065	0.324	0.792 (0.37%)	[0.16,0.82]	0.319	0.811 (2.04%)
$m_4(=m_{4-6})$ ^{L2}	$m_{c,4}=m_4+bx_{m4}$	0.709	0	0.140	0.303	0.107	0.337	0.715 (0.80%)	0.578	0.246	0.439	0.334	0.731 (3.14%)	[0.18,0.68]	0.274	0.723 (1.98%)
$\varepsilon^{*2,L2}$	x_{ul}	-	0.6	0.620	0.045	0.614	0.035	0.614 (2.27%)	6.072	0.351	6.067	0.327	6.067 (1.12%)	[5.76,6.45]	0.617	0.617 (2.85%)

*: erroneous sensors, *2: unknown parameter, *3: temperatures (°C) and mass flow rates (kg/s), a is 1/200, and b is 1/20, *4: from local calibration results in the two-step calibration (priors of the local calibration are same with those in Table 3), L1: from 1st local calibration, L2: from 2nd local calibration.

Table 5. A summary of results from three cases in both of calibration approaches.

Methods	Whole calibration						Two-step calibration					
	Accuracy ^{*1}			Improvement ^{*1}			Accuracy			Improvement		
	T ^{*2}	M	U	T	M	U	T	M	U	T	M	U
Proportional Bayesian	G ^{*3}	G	G	B	G	N	G	G	G	N	G	G
Disproportional Bayesian	B	G	G	N	G	N	G	G	G	G	G	G
GA-based optimization	B	G	G	B	N	N	B	G	G	N	G	N

*1: Accuracy and improvement by multiple measurement sets, *2: T (temperatures), M (mass flow rates), U (unknown parameter), and *3: G (good for accuracy or improvement), N (normal accuracy or normal improvement), B (bad accuracy or bad improvement).

5. Conclusion

Based on the parameter estimation, the extended virtual in-situ calibration derives the offsetting constants and/or correction coefficients of correction functions and unknown parameters of the system model by minimizing the distance between values of correction functions and benchmarks for all working sensors in a target system in order to calibrate random and systematic errors. Such a parameter estimation for EVIC can be conducted by a deterministic or statistical method. Also, this calibration method has two approaches (whole and two-step) and uses several measurement sets to estimate the variables of interest more accurately by constraining the problem.

This study formulated the calibration problem for an ARS and solved it with both Bayesian MCMC and GA-based optimization for the whole calibration and the two-step calibration. The calibrated results from two methods were quantitatively discussed by comparing their accuracy distributions (RMSE) under different numbers of data sets. The whole calibration results indicated the different distributions and sufficient numbers according to the types of variables. Unless informative priors were used for two methods, temperature distributions calculated from two methods did not converge to the reference value even when using ten sets. This was because certain temperatures in terms of a temperature difference, which is usually included in mathematical models of building systems, were unable to be constrained by using only the multiple sets in the whole calibration approach. On the contrary, regardless of priors in the Bayesian method, the two-step approach that was suggested in this study to handle the challenge could improve significantly all accuracy distributions, even for temperatures when increasing the data sets. However, the GA-based method based on the two-step approach could not achieve the reference accuracy for temperatures due to the wide ranges that led to large calibration errors of the normal sensors for temperatures. Table 5 suggested the summary of these results. These results can vary depending on the used data sets, the type of a system and its model, correction function used in EVIC, and the different algorithms selected, but the positive and negative perspectives for deterministic and statistical approaches and their effects on the accuracy distributions are representative from the selected algorithms here. It is the difference in defining priors of the two approaches that mainly cause the different perspectives and results under the inherent characteristic of EVIC, e.g. an under-determined condition and two error types.

Considering all distributions in each method, calibration results calculated from Bayesian MCMC were more reliable and stable. Estimating a proper searching space for the variables in a deterministic method as GA is difficult. A wide range can result in large calibration errors for normal sensors, while a narrow range could lead to a failure to detect large systematic errors. With Bayesian calibration, the two-

step approach can treat combinations of random and systematic errors for real building systems that have various sensor errors. Finally, it is important to determine the reasonable number of measurement sets by considering different types of working sensors and unknown parameters in a building system.

Acknowledgments

The authors gratefully acknowledge the support of this study by the National Science Foundation under Grant No. EPS-10004094.

References

- [1] Annual Energy Outlook 2013. EIA. 2013.
- [2] Yu Y, Loftness V, Yu D. Multi-structural fast nonlinear model-based predictive control of a hydronic heating system. *Building and Environment* 2013;69:131–48.
- [3] Yao Y, Chen J. Global optimization of a central air-conditioning system using decomposition–coordination method. *Energy and Buildings* 2010;42(5):570–83.
- [4] Yao Y, Wang L. Energy analysis on VAV system with different air-side economizers in China. *Energy and Buildings*, 2010;42(8):1220–30.
- [5] Yu Y, Woradechjurnroen D, Yu D. A review of fault detection and diagnosis methodologies on air-handling units. *Energy and Buildings* 2014;82:550–62.
- [6] Li S, Wen J. Application of pattern matching method for detecting faults in air handling unit system. *Automation in Construction* 2014;43:49–58.
- [7] Du Z, Fan B, Jin X, Chi J. Fault detection and diagnosis for buildings and HVAC systems using combined neural networks and subtractive clustering analysis. *Building and Environment* 2014;73:1–11.
- [8] Yu Y, Li H. Virtual in-situ calibration method in building systems. *Automation in Construction* 2015;59:59–67.
- [9] Yoon S, Yu Y. Extended virtual in-situ calibration method in building systems using Bayesian inference. *Automation in Construction* 2017;73:20–30.
- [10] Yoon S, Yu Y. Autonomous in-situ sensor calibration in building systems using Bayesian inference. in: 9th IAQVEC Conference. Songdo, Republic of Korea, October 20–26, 2016.
- [11] Yoon S, Seo S, Cho W, Song D. A calibration method for whole-building airflow simulation in high-rise residential buildings. *Building and Environment* 2015;85:253–62.
- [12] Wang S, and Xu X. Parameter estimation of internal thermal mass of building dynamic models using genetic algorithm. *Energy Conversion and Management* 2006;47(13–14):1927–41.
- [13] Wang S, Xu X. Simplified building model for transient thermal performance estimation using GA-based parameter identification. *Int J Thermal Sci* 2006;45(4):419–32.
- [14] Heo Y, Augenbroe G, Graziano D, Muehleisen RT. Scalable methodology for large scale building energy improvement: Relevance of calibration in model-based retrofit analysis. *Building and Environment* 2015;87:342–50.
- [15] Tian W, Yang S, Li Z, Wei S, Pan W, Liu Y. Identifying informative energy data in Bayesian calibration of building energy models. *Energy and Buildings* 2016;119:363–76.
- [16] Berger J, Orlando H, Mendes N, Guernouti S. Bayesian inference for estimating thermal properties of a historic building wall. *Building and Environment* 2016;106:327–39.
- [17] Kim YJ, Yoon SH, Park CS. Stochastic comparison between simplified energy calculation and dynamic simulation. *Energy and Buildings* 2013;64:332–42.
- [18] ASHRAE. Handbook of Fundamentals. Atlanta: American Society of Heating, Air-Conditioning and Refrigeration Engineers. In. Inc.; 2013.

- [19] Ebrahimi K, Jones GF, Fleischer AS. Thermo-economic analysis of steady state waste heat recovery in data centers using absorption refrigeration. *Applied Energy* 2015;139:384–97.
- [20] Rubio-Maya C, Pachero-Ibarra JJ, Belman-Flores JM, Galvan-Gonzalez SR, Mendoza-Covarrubias C. NLP model of a LiBr–H₂O absorption refrigeration system for the minimization of the annual operating cost. *Applied Thermal Engineering* 2012;37:10–8.
- [21] Sun DW. Thermodynamic design data and optimum design maps for absorption refrigeration systems. *Applied Thermal Engineering* 1997;17(3):211–21.
- [22] Lansing FL. Computer modeling of a single-stage lithium bromide/water absorption refrigeration unit. *JPL Deep Space Network Progress Report DSN 1976;42-32:247–57*.
- [23] Kaita Y. Thermodynamic properties of lithium bromide–water solutions at high temperatures. *International Journal of Refrigeration* 2001;24:374–90.
- [24] Lee RJ, DiGuilio RM, Jeter SM, Teja AS. Properties of lithium bromide–water solutions at high temperatures and concentration. II. Density and viscosity. *ASHRAE Trans* 1990;96(Pt. 1):709–28.
- [25] Aphornratana S, Eames IW. Thermodynamic analysis of absorption refrigeration cycle using the second law of thermodynamics method. *International Journal of Refrigeration* 1995;18(4):244–52.
- [26] Dudley RM. *Uniform Central Limit Theorems*. Cambridge University Press, ISBN-13:978-0521498845; 1999.
- [27] Gilks WR, Richardson S, Spiegelhalter D. *Markov Chain Monte Carlo in Practice*. Chapman and Hall; 1995.
- [28] Gamerman D, Lopes HF. *Markov Chain Monte Carlo-Stochastic Simulation for Bayesian Inference*. Chapman and Hall; 2006.
- [29] Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E. Equations of state calculations by fast computing machines. *Journal of Chemical Physics* 1953;21(6):1087–92.
- [30] Hastings WK. Monte Carlo sampling methods using Markov Chains and their applications. *Biometrika* 1970;57(1):97–109.
- [31] Deb K. Genetic algorithms for function optimisation. In: Ferrera F, Verdegay JL, editors. *Genetic algorithms and soft computing*. Heidelberg: Physica-Verlag; 1996. p. 3–29.
- [32] Mitchell M. *An introduction to genetic algorithm*. The MIT Press; 1997.