Building and Environment 115 (2017) 54–66

Contents lists available at ScienceDirect

Building and Environment

journal homepage: www.elsevier.com/locate/buildenv

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

A R T I C L E I N F O

Article history:
Received 22 November 2016
Received in revised form
11 January 2017
Accepted 12 January 2017
Available online 16 January 2017

Keywords:
Sensor calibration
Virtual in-situ calibration
Bayesian MCMC
Genetic algorithm
Absorption refrigeration
Building systems

A B S T R A C T

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-H_2O 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 Ref. [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

* Corresponding author.
E-mail address: yuebinyu@gmail.com (Y. Yu).

http://dx.doi.org/10.1016/j.buildenv.2017.01.013
0360-1323/

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.

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

$$Y_c = g(M, x_1, x_2, \ldots, x_k) \tag{2}$$

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

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