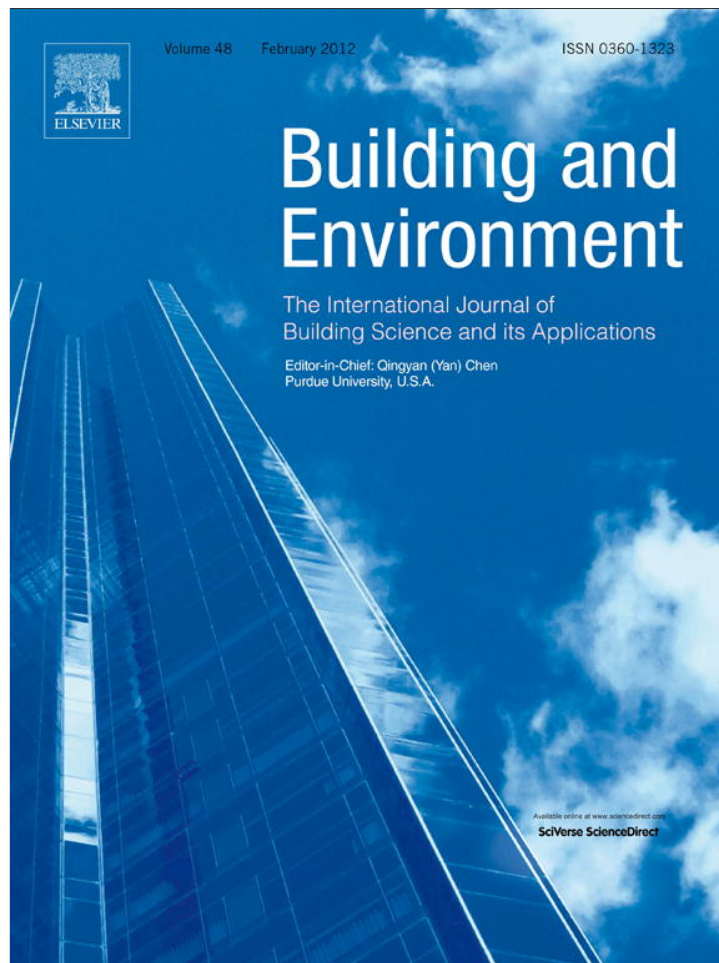


Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



(This is a sample cover image for this issue. The actual cover is not yet available at this time.)

This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Contents lists available at SciVerse ScienceDirect

Building and Environment

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

System parameter identification theory and uncertainty analysis methods for multi-zone building heat transfer and infiltration

Hiroyasu Okuyama*, Yoshinori Onishi¹

Institute of Technology, Shimizu Corporation, 4-17 Etchujima, 3-chome Koto-ku, Tokyo 135-8530, Japan

ARTICLE INFO

Article history:

Received 27 October 2011

Received in revised form

1 February 2012

Accepted 2 February 2012

Keywords:

Thermal network model

State space equation

System identification

Least squares

Discrepancy ratio

Optimum excitation waveform

ABSTRACT

Methods for on-site measurement of building thermal performance system parameters such as coefficient of heat loss, solar heat gain, effective thermal capacity, infiltration rate, and effective mixing volume are very important, yet a nontrivial task. Although these are steady-state parameters, on-site measurements are exposed to changing meteorological conditions and are affected by the thermal capacity of the building. In addition, these parameters should generally be estimated by using a multi-zone model such as inter-zone flow rates. In this regard, a state space equation model, referred to as a “thermal network model,” has been devised to generalize such multi-zone heat transfer system and tracer gas diffusion system measurements. This model is composed of three parameter types, and we have developed a system parameter identification theory and uncertainty analysis method using least squares, as well as actual measurement systems. In the present paper, we improve the least-squares regression equation, the uncertainty analysis method, and the reliability evaluation method. We investigate appropriate excitation waveforms and frequencies for heating and tracer gas release, as well as a low-pass filter for pre-processing measurement data. We verify these theories and methods using computer-simulated measurement.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

Parameters related to the energy efficiency of heating and cooling, as well as to the thermal comfort of the building environment, include the coefficient of external wall heat transmission, solar heat gain, and effective thermal capacity. In addition, parameters such as infiltration rate and effective mixing volume are related to healthy indoor air quality. In this regard, it is necessary to develop methods for on-site measurements of these system parameters, because buildings might not have been built in accordance with plans and from specified materials, components might have deformed, or material performance may have deteriorated.

In contrast to measurements performed on test buildings constructed in laboratories in an artificial climate, it is difficult to establish these system parameters through on-site measurements on actual buildings, because outdoor fluctuations in temperature and solar radiation and the influence of the tested building's thermal capacity make conditions unstable. Furthermore, it may be

necessary to apply a multi-zone model to estimate inter-zone flow rates and heat transmission coefficients.

In Japan, research into on-site measurement methods for assessing building performance began in the late 1970s with Matsuo et al. [1], who devised a method called Matsuo's digital filtering. Emura et al. improved upon Matsuo's method by using condensed orthogonal meteorological effects [2], and further developed an alternative estimation method for states and parameters [3]. Akashi et al. performed research on least-squares-based system identification methods for the static and dynamic thermal characteristics of office buildings, utilizing a transfer function model [4]. Hattori et al. [5] continue to research Matsuo's method. A recent version of Matsuo's method estimates the heat loss coefficient of a single-zone model by using step excitation from an electric heater. In the estimation process, time series response factors are calculated by the least-squares method. Most models are single-zone and require assumptions of invariability and linearity. These methods are insufficient for discerning simultaneous system identification of both thermal and infiltration systems. Excitations for this system identification are mainly researched using square wave or step functions with high-frequency sinusoidal components as the Fourier series expansion. Furthermore, these methods have unsolved problems related to estimating the parameters of solar

* Corresponding author. Tel.: +81 3 3820 6438; fax: +81 3 3820 5955.

E-mail addresses: okuyama@shimz.co.jp (H. Okuyama), ohnishi@shimz.co.jp (Y. Onishi).¹ Tel.: +81 3 3820 6439; fax: +81 3 3820 5955.

heat gain and effective thermal capacity, as well as with evaluating the statistical uncertainty of these estimated parameters.

The state of the art is similar in other countries, where copious research has been performed mainly on single-zone models using various estimation methods, such as the recursive least-squares method developed by Zaheer-uddin [6], the iterative descent technique by Dewson et al. [7] (which involves an evaluation function of the root mean square of the predicted and the measured temperature), the autoregressive moving average method by Norlen [8], and the multi-variable time series least-squares method by Crawford et al. [9]. Baudier et al. [10] adopted a state space equation model and applied Marquardt's algorithm [11] to obtain parameters. Wang et al. [12] studied the application of genetic algorithms [13] to a simple thermal network. Another example is the two-node model developed by Jimenez et al. [14,15], which utilizes the MATLAB software package.

Okuyama [16] gives another possibility, a method with fewer measurement points than nodes in the system identification model. That method, however, requires measurement results for multiplex frequency sinusoidal excitation and accurate higher-order time derivatives, and furthermore has problems related to theoretical generalization.

Because of the complexity of heat transfer in buildings, simulation of changes in room temperature are assumed to require complex models consisting of many nodes, including the building structure, with varying solar heat gain coefficients. We note, however, that the parameters to be estimated are for simplified steady-state models, meaning that time-averaged parameters are sufficient for certain terms.

In contrast to heat transfer systems, system identification models for a tracer gas diffusion system used in multi-zone infiltration measurements can trace concentration changes accurately, even under a single-node model for each zone, by using mixing fans in each zone. System identification then becomes relatively easier than in the case of a thermal system. Extensive research on measurement methods for multi-zone airflow rates has been carried out; for example, 47 papers were cited by Miller et al. in the Introduction section of their paper [17]. Among these, Sherman, Prior, and O'Neill are frequently cited. Papers written by Japanese researchers, including Enai, Honma, and Okuyama, have also been introduced.

Such multi-zone heat transfer and infiltration systems, characterized by heat and tracer gas diffusion systems, respectively, can be represented by a general mathematical model (a thermal network model), even though the diffusion variables are different. It is well known that the various dynamic systems are described by the state space equation model in modern control theory. Jiang [18] and Hong [19] researched the formulation of an analytical solution for an application to building air-conditioning. However, the analytical solution can be improved by using the spectral resolution of a projection on eigenspaces [20], and the universally applicable modeling concept [21] devised by Okuyama.

Okuyama [22] derived two types of least-squares methods for batch identification and recursive identification of the thermal network state space equation model, the latter using the matrix inversion lemma introduced by Woodbury [23]. That study also developed the system identification program SPID. There, the recursive least-squares methods seemed to have the ability to follow parameter time variations, but performance did not live up to expectations. It was also realized that estimation accuracy was not higher than that for batch identification. A moving batch identification term T by Δt was therefore investigated, and the method tested on actual buildings. Changes in infiltration rate in relation to the changes in indoor/outdoor temperature differences have been appropriately estimated and verified [24].

In recent years, Okuyama et al. developed a statistical data analysis method for steady-state multi-zone infiltration measurements using multiple perfluorocarbon tracers [25]. They introduced a discrepancy ratio β , which evaluates prerequisite validity for measurements and data analysis, and is also useful in heat transfer systems.

Our latest study [26], despite concerning a single-zone model with tracer gas pseudo-non-uniformity, revealed that low-frequency sinusoidal excitation enables satisfactory results, even in a rough system identification model of fewer nodes. We also found that a low-pass filter using moving term averages is necessary for pre-processing measurement data. We solved the problem of determining the optimum concentration decay term or, more generally, the optimum excitation stopping term. We furthermore found an appropriate calculation method for the uncertainty propagation equation, and the usefulness of both the coefficient of determination (COD) and the discrepancy ratio β for reliability evaluation.

The present paper improves the earlier theory [22], and confirms the aforementioned findings for the multi-zone system through use of a single-zone model [26]. Appropriate verification is impossible in case studies of actual buildings because the true system parameters are unknown. The first step in this study, therefore, is computer-simulated measurements, for which true or comparable reference values are known. When measuring the thermal performance of buildings, mechanical ventilation is usually stopped and heat loss by infiltration is included in the heat transmission of external walls. For the computer-simulated measurements in this research, the program NETS² [27] was used for combined simulation of heat, air, and tracer gas transfer. Simulated multi-zone infiltration measurements were also performed for reference.

2. System parameter estimation by double application of least-squares

2.1. Primary regression equation for diffusion system parameters

The framework of the spatial discretization model of a diffusion system can be written in the form of Equation (1), which is referred to as the nodal equation of a completely linked system. From this, the simultaneous ordinary differential Equation (2), also known as the state space equation, is constructed. Here, x_j , m_{ij} , c_{ij} , and r_{ij} are, respectively, diffusion variables such as the temperature of node j , the generalized capacitance of node i , the generalized conductance from node j toward node i , and the flux coefficient from the flux source j towards node i . Note that the flow direction is opposite the order of the subscripts i and j , in accordance with the linear algebra rule indicating element position in the matrix. In addition, n is the number of nodes with unknown values, no is the number of nodes with given values, and ng is the number of flux sources. Equation (2) is the entire equation for all nodes.

Generalized conductance c_{ij} such as infiltration or heat transfer by long wave radiation varies with temperature and time. Nevertheless the present methods are useful in cases where average results of the system identification term are sufficient. A linearizing approximation for long wave radiation is described in Appendix A.

$$\sum_{j=1}^n m_{ij} \dot{x}_j = \sum_{j=1}^{n+no} c_{ij} \cdot (x_j - x_i) + \sum_{j=1}^{ng} r_{ij} \cdot g_j \quad (1)$$

² NETS is authorized by the Ministry of Land, Infrastructure and Transport, Japan, as a calculation method for building annual heating and cooling load, as of 22 October 2002.

$$\mathbf{M} \cdot \dot{\mathbf{x}} = \mathbf{C} \cdot \mathbf{x} + \mathbf{C}_0 \cdot \mathbf{x}_0 + \mathbf{R} \cdot \mathbf{g} \quad (2)$$

Following the method described in [22], we transform Equation (2) into Equation (3), which will be the primary regression equation for system parameters. If we assume that there is at least one known parameter for a node, the product of that parameter with the variable x_j or g_j is transposed into \mathbf{y} on the left. Next, the vectors incorporating the unknown parameters m_{ij} , c_{ij} , and r_{ij} are denoted with \mathbf{m} , \mathbf{c} , and \mathbf{r} , and their corresponding matrices are defined as \mathbf{D} , \mathbf{X} , and \mathbf{G} . Combining the above, the matrix \mathbf{Z} ($n \times na$) is formed from the measurement values, and the vector \mathbf{a} (na) incorporates the system parameters.

$$\mathbf{y} = \mathbf{D}(\dot{x}_i) \cdot \mathbf{m} + \mathbf{X}(x_i) \cdot \mathbf{c} + \mathbf{G}(g_i) \cdot \mathbf{r} = [\mathbf{D} \quad \mathbf{X} \quad \mathbf{G}] \cdot \begin{bmatrix} \mathbf{m} \\ \mathbf{c} \\ \mathbf{r} \end{bmatrix} = \mathbf{Z} \cdot \mathbf{a} \quad (3)$$

The interval between measurements is Δt , the total number of measurement intervals is nt , and the measurement term is T . Below, \mathbf{y}_k and \mathbf{Z}_k are defined from the linear interpolation integral from $(k-1)\Delta t$ to $k\Delta t$.

$$\mathbf{y}_k = \int_{(k-1)\Delta t}^{k\Delta t} \mathbf{y} dt \quad (4)$$

$$\mathbf{Z}_k = \int_{(k-1)\Delta t}^{k\Delta t} \mathbf{Z} dt \quad (5)$$

The next Equation (6) is defined as the primary regression equation for the system parameters between measurement intervals $(k-1)\Delta t$ and $k\Delta t$.

$$\mathbf{y}_k = \mathbf{Z}_k \cdot \mathbf{a} \quad (6)$$

2.2. Two quadratic form equations for regression

In the early method for system identification [22], linear dependencies between the system parameters were prepared using the air flow rate balance and/or the symmetry of thermal conduction, and some system parameters were represented by others. These equations were embedded in the primary regression equation, reducing the number of parameters before performing the least-squares formulation. The algorithm for preparing dependent equations for system parameters was rather complex, however. In addition, when applying the non-negative least-squares method [28], the parameters hidden by these dependencies were not subject to the non-negative constraint. These inconveniences are resolved by adding the dependency constraint equations in order of the rows of the simultaneous regression equations of the least-squares method.

The formulation of the least-squares method in this paper differs from the commonly accepted one. The regression equations obtained from the measurement values are normally added in rows. This paper is the first to define the equation error of the primary regression Equation (6) through the following equation, where the error vector ${}_n\mathbf{e}_k$ is of dimension n .

$${}_n\mathbf{e}_k = \mathbf{y}_k - \mathbf{Z}_k \cdot \mathbf{a} \quad (7)$$

The sum of the products of this transposed error vector and the error vector for the measurement term is defined as the evaluation function for the least squares. Note that the t on the left shoulder of vector \mathbf{e} represents the transposition of a matrix or vector.

$$J_n = \sum_{k=1}^{nt} {}_n\mathbf{e}_k \cdot {}_n\mathbf{e}_k = \sum_{k=1}^{nt} (\mathbf{y}_k - \mathbf{Z}_k \cdot \mathbf{a}) \cdot (\mathbf{y}_k - \mathbf{Z}_k \cdot \mathbf{a}) \quad (8)$$

This evaluation function is differentiated by vector \mathbf{a} , composed of the system identification parameters, and the result is set equal to $\mathbf{0}$. Rearranging with respect to \mathbf{a} , we obtain the following equation. Although it seems that this equation can be solved for \mathbf{a} , in fact an appropriate solution cannot be obtained because constraint conditions such as the air flow balance are not considered. Equation (9) is called the quadratic form primary equation.

$$\frac{\partial J_n}{\partial \mathbf{a}} = - \sum_{k=1}^{nt} {}^t\mathbf{Z}_k \cdot \mathbf{y}_k + \sum_{k=1}^{nt} {}^t\mathbf{Z}_k \cdot \mathbf{Z}_k \cdot \mathbf{a} = 0 \quad (9)$$

In the composite regression equation, defined later, the quadratic form primary equation produces an equation error \mathbf{e}_a (na), as follows.

$$\mathbf{e}_a = \sum_{k=1}^{nt} {}^t\mathbf{Z}_k \cdot \mathbf{y}_k - \sum_{k=1}^{nt} {}^t\mathbf{Z}_k \cdot \mathbf{Z}_k \cdot \mathbf{a} \quad (10)$$

The air flow rate balance, the mass flow rate balance, the symmetry of thermal conductance, and a number of system parameters linearly dependent on a few physical characteristics, such as thermal conductivity, are included in the constraint equations. These equations, with a total number of ns , are referred to as first-degree constraint equations. When several system parameters are known, they can be included into the first-degree constraint equations, which can be expressed as in the following equation by using a matrix \mathbf{S} ($ns \times na$) and a vector \mathbf{d} (ns). Here, the elements of \mathbf{S} and \mathbf{d} can be 1, -1 , or 0, depending on the conditions of the air flow balance and the symmetry of the generalized conductance. When system parameter values are given, however, they are included in \mathbf{d} , and \mathbf{S} is composed only of 1s and 0s. We consider the first-degree constraint Equation (11) as satisfying the meaning of least squares, which minimizes the evaluation function J_d expressed by Equation (12), where the right side of Equation (11) is integrated along Δt , corresponding to the primary regression Equation (6).

$$\mathbf{e}_d = \Delta t \cdot (\mathbf{d} - \mathbf{S} \cdot \mathbf{a}) \quad (11)$$

$$J_d = {}^t\mathbf{e}_d \cdot \mathbf{e}_d = \Delta t^2 \cdot {}^t(\mathbf{d} - \mathbf{S} \cdot \mathbf{a}) \cdot (\mathbf{d} - \mathbf{S} \cdot \mathbf{a}) \quad (12)$$

The minimizing equation for (12) is obtained by differentiation of J_d by \mathbf{a} . This solution will have an equation error \mathbf{e}_s (ns) in the composite regression equation, and this quadratic form constraint equation is written as follows.

$$\mathbf{e}_s = \Delta t^2 \cdot {}^t\mathbf{S} \cdot \mathbf{d} - \Delta t^2 \cdot {}^t\mathbf{S} \cdot \mathbf{S} \cdot \mathbf{a} \quad (13)$$

2.3. Weighting matrices

The presence of large variation in the magnitude of matrix elements between the quadratic form primary equation and the quadratic form constraint equation, and between each row, exerts a negative influence on the estimation accuracy of the least squares. The weighting matrices \mathbf{W}_a and \mathbf{W}_s are therefore introduced to ensure an unbiased estimate.

For each row in the matrix $\mathbf{H}_a = \sum {}^t\mathbf{Z}_k \cdot \mathbf{Z}_k$ ($na \times na$), which is multiplied by \mathbf{a} in Equation (10), the maximal absolute value of the i -th row is taken, and its squared inverse value is substituted into the i -th diagonal position in \mathbf{W}_a . Here, if all elements of a row in \mathbf{H}_a are 0, then 1 is substituted. The square corresponds to the square of ${}^t\mathbf{e} \cdot \mathbf{e}$. As a result, the dimension of \mathbf{W}_a is ($na \times na$).

$$\mathbf{W}_a = \begin{bmatrix} \text{abs max}(1\text{st row in } \mathbf{H}_a)^{-2} & 0 & \dots & 0 \\ 0 & \text{abs max}(2\text{nd row in } \mathbf{H}_a)^{-2} & \vdots & \\ \vdots & \vdots & 0 & \\ 0 & \dots & 0 & \text{abs max}(na - \text{th row in } \mathbf{H}_a)^{-2} \end{bmatrix} \quad (14)$$

In multi-zone air flow measurements, if the tracer gas concentration of the outdoor air is 0, and if the air flow rate coming from outdoors to a certain indoor zone is the j -th element in the vector \mathbf{a} , then all elements of the j -th row in \mathbf{H}_a will be 0.

For the matrix $\mathbf{H}_S = \Delta t^2 \cdot {}^t\mathbf{S} \cdot \mathbf{S}$, ($na \times na$) in Equation (13), the maximal absolute value of each row in the matrix \mathbf{H}_S ($na \times na$) is taken, and its squared inverse value is substituted into the diagonal position in \mathbf{W}_S ($na \times na$). Here, if all elements of a row in \mathbf{H}_S are 0 then 1 is substituted.

$$\mathbf{W}_S = \begin{bmatrix} \text{abs max}(1\text{st row in } \mathbf{H}_S)^{-2} & 0 & \dots & 0 \\ 0 & \text{abs max}(2\text{nd row in } \mathbf{H}_S)^{-2} & \vdots & \\ \vdots & \vdots & 0 & \\ 0 & \dots & 0 & \text{abs max}(na - \text{th row in } \mathbf{H}_S)^{-2} \end{bmatrix} \quad (15)$$

2.4. Solution by double application of least-squares

The following equation defines the error vector \mathbf{e} ($2na$) of the composite regression equation by coupling the error vectors of the quadratic form constraint equation in Equation (13) with the quadratic form primary equation in Equation (10) in the direction of increasing rows.

$$\mathbf{e} = \begin{bmatrix} \mathbf{e}_a \\ \mathbf{e}_S \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^{nt} {}^t\mathbf{Z}_k \cdot \mathbf{y}_k \\ \Delta t^2 \cdot {}^t\mathbf{S} \cdot \mathbf{d} \end{bmatrix} - \begin{bmatrix} \sum_{k=1}^{nt} {}^t\mathbf{Z}_k \cdot \mathbf{Z}_k \\ \Delta t^2 \cdot {}^t\mathbf{S} \cdot \mathbf{S} \end{bmatrix} \cdot \mathbf{a} \quad (16)$$

The matrix \mathbf{F} ($2na \times na$) and the vector \mathbf{b} ($2na$) of the composite constraint equation are defined as follows.

$$\mathbf{F} = \begin{bmatrix} \sum_{k=1}^{nt} {}^t\mathbf{Z}_k \cdot \mathbf{Z}_k \\ \Delta t^2 \cdot {}^t\mathbf{S} \cdot \mathbf{S} \end{bmatrix} \quad (17)$$

$$\mathbf{b} = \begin{bmatrix} \sum_{k=1}^{nt} {}^t\mathbf{Z}_k \cdot \mathbf{y}_k \\ \Delta t^2 \cdot {}^t\mathbf{S} \cdot \mathbf{d} \end{bmatrix} \quad (18)$$

From these definitions, Equation (16) can be rewritten as follows.

$$\mathbf{e} = \begin{bmatrix} \mathbf{e}_a \\ \mathbf{e}_S \end{bmatrix} = \mathbf{b} - \mathbf{F} \cdot \mathbf{a} \quad (19)$$

The product of the transposed error vector and the error vector of the composite regression equation is defined as an evaluation function for the least squares. Weighting matrices \mathbf{W}_a and \mathbf{W}_S will be combined into \mathbf{W} , as in the following equation.

$$J = {}^t\mathbf{e} \cdot \begin{bmatrix} \mathbf{W}_a & 0 \\ 0 & \mathbf{W}_S \end{bmatrix} \cdot \mathbf{e} = {}^t\mathbf{e} \cdot \mathbf{W} \cdot \mathbf{e} = {}^t(\mathbf{b} - \mathbf{F} \cdot \mathbf{a}) \cdot \mathbf{W} \cdot (\mathbf{b} - \mathbf{F} \cdot \mathbf{a}) \quad (20)$$

Differentiating by the vector \mathbf{a} , we obtain the following equation.

$$\frac{\partial J}{\partial \mathbf{a}} = 2 \cdot {}^t\mathbf{F} \cdot \mathbf{W} \cdot \mathbf{F} \cdot \mathbf{a} - 2 \cdot {}^t\mathbf{F} \cdot \mathbf{W} \cdot \mathbf{b} = 0 \quad (21)$$

In this way, estimation of the vector \mathbf{a} is obtained through double application of the least-squares method.

$$\hat{\mathbf{a}} = ({}^t\mathbf{F} \cdot \mathbf{W} \cdot \mathbf{F})^{-1} \cdot ({}^t\mathbf{F} \cdot \mathbf{W} \cdot \mathbf{b}) \quad (22)$$

3. Uncertainty analysis and reliability evaluation methods

3.1. Variance-covariance matrix of estimated system parameters

Differences between the premises of the system identification model and the actual phenomena, combined with the uncertainty of measurement values, can result in negative parameters. We therefore recommend application of the non-negative least-squares method [28] to Equation (19). The following equation calculates the variance-covariance matrix of the uncertainty of the parameters. The dimension of this matrix is Λ_a ($na \times na$), and Appendix B describes its derivation.

$$\Lambda_a = ({}^t\mathbf{F} \cdot \mathbf{W} \cdot \mathbf{F})^{-1} \cdot ({}^t\mathbf{F} \cdot \mathbf{W} \cdot E(\mathbf{e} \cdot {}^t\mathbf{e}) \cdot \mathbf{W} \cdot \mathbf{F}) \cdot {}^t\{({}^t\mathbf{F} \cdot \mathbf{W} \cdot \mathbf{F})^{-1}\} \quad (23)$$

The expected error matrix of the composite regression equation $E(\mathbf{e} \cdot {}^t\mathbf{e})$ ($2na \times 2na$) is written as follows, since the covariance of the errors for the upper half (the quadratic form primary equation) and the lower half (the quadratic form constraint equation) can be regarded as 0.

$$E(\mathbf{e} \cdot {}^t\mathbf{e}) = E \left(\begin{bmatrix} \mathbf{e}_a \\ \mathbf{e}_S \end{bmatrix} \cdot {}^t \begin{bmatrix} \mathbf{e}_a \\ \mathbf{e}_S \end{bmatrix} \right) \quad (24)$$

$$= \begin{bmatrix} E(\mathbf{e}_a \cdot {}^t\mathbf{e}_a) & 0 \\ 0 & E(\mathbf{e}_S \cdot {}^t\mathbf{e}_S) \end{bmatrix}$$

There are two ways of defining the expected equation error matrices for the quadratic form primary and constraint equations; based on the equation residual, or based on measurement uncertainty. These methods are described in Sections 3.2 and 3.3. Below, we use r (residual) and m (measurement) as subscripts for the stochastic expectation operator E .

3.2. Propagation from equation residual and coefficient of determination

We now consider calculation of the expected error matrix for the composite regression equation of Equation (24) from the residuals. The origin of the residual is not only uncertainty of measurement, but also differences in the premises and the structure of the system identification model as compared to the actual phenomena. The residual vector \mathbf{v}_k (n) of the primary regression equation is calculated as follows.

$$\mathbf{v}_k = \mathbf{y}_k - \mathbf{Z}_k \cdot \hat{\mathbf{a}} \quad (25)$$

First, the expected matrix $E(\mathbf{v}_k \cdot {}^t\mathbf{v}_k)$ ($n \times n$) of the residual for the primary regression equation is calculated from Equation (26). Here, the covariance of off-diagonal elements of $\Sigma \mathbf{v}_k \cdot {}^t\mathbf{v}_k$ is regarded as 0; therefore the *diag* operation, which extracts only the elements on the diagonal, is used. Diagonalization also avoids asymmetries in the calculated matrix due to numerical computation errors in Equation (27).

$$E(\mathbf{v}_k \cdot {}^t\mathbf{v}_k) = \text{diag} \left(\frac{1}{(nt - na)} \sum_{k=1}^{nt} \mathbf{v}_k \cdot {}^t\mathbf{v}_k \right) \quad (26)$$

Using Equation (26) allows calculation of the expected error matrix $E_r(\mathbf{e}_a \cdot {}^t\mathbf{e}_a)$ ($na \times na$) for the quadratic form primary equation from the following Equation (27), which is derived in Appendix C.

$$E_r(\mathbf{e}_a \cdot {}^t\mathbf{e}_a) = \sum_{k=1}^{nt} {}^t\mathbf{Z}_k \cdot E(\mathbf{v}_k \cdot {}^t\mathbf{v}_k) \cdot \mathbf{Z}_k \quad (27)$$

The sum of residual squares of $s(\mathbf{a})$, which is necessary for calculating the coefficient of determination (COD), is now obtained from the following equation.

$$s(\hat{\mathbf{a}}) = \sum_{k=1}^{nt} {}^t\mathbf{v}_k \cdot \mathbf{v}_k = \sum_{k=1}^{nt} {}^t(\mathbf{y}_k - \mathbf{Z}_k \cdot \hat{\mathbf{a}}) \cdot (\mathbf{y}_k - \mathbf{Z}_k \cdot \hat{\mathbf{a}}) \quad (28)$$

Also, the total variation s_y is calculated as follows. Where the top bar of \mathbf{y}_k means average.

$$\begin{aligned} s_y &= \sum_{k=1}^{nt} {}^t(\mathbf{y}_k - \bar{\mathbf{y}}_k) \cdot (\mathbf{y}_k - \bar{\mathbf{y}}_k) \\ &= \sum_{k=1}^{nt} {}^t\mathbf{y}_k \cdot \mathbf{y}_k - \frac{1}{nt} \cdot \left(\sum_{k=1}^{nt} {}^t\mathbf{y}_k \right) \cdot \left(\sum_{k=1}^{nt} \mathbf{y}_k \right) \end{aligned} \quad (29)$$

COD is therefore calculated from $s(\mathbf{a})$ and s_y by the following equation.

$$\text{COD} = 1 - \frac{s(\hat{\mathbf{a}})}{s_y} \quad (30)$$

The expected error matrix $E_r(\mathbf{e}_s \cdot {}^t\mathbf{e}_s)$ ($na \times na$) for the quadratic form constraint equation is also calculated using the residual.

$$E_r(\mathbf{e}_s \cdot {}^t\mathbf{e}_s) = \Delta t^4 \cdot ({}^t\mathbf{S} \cdot \mathbf{d} - {}^t\mathbf{S} \cdot \hat{\mathbf{a}}) \cdot ({}^t\mathbf{S} \cdot \mathbf{d} - {}^t\mathbf{S} \cdot \hat{\mathbf{a}}) \quad (31)$$

The expected error matrix of the composite regression equation $E_r(\mathbf{e} \cdot {}^t\mathbf{e})$ ($2na \times 2na$) due to residuals is calculated from Equations (27) and (31), as follows.

$$E_r(\mathbf{e} \cdot {}^t\mathbf{e}) = \begin{bmatrix} E_r(\mathbf{e}_a \cdot {}^t\mathbf{e}_a) & 0 \\ 0 & E_r(\mathbf{e}_s \cdot {}^t\mathbf{e}_s) \end{bmatrix} \quad (32)$$

Next, the propagation of uncertainty towards the identification parameter from the residual of the composite regression equation is calculated.

$${}^r\Lambda_a = ({}^t\mathbf{F} \cdot \mathbf{W} \cdot \mathbf{F})^{-1} \cdot ({}^t\mathbf{F} \cdot \mathbf{W} \cdot E_r(\mathbf{e} \cdot {}^t\mathbf{e}) \cdot \mathbf{W} \cdot \mathbf{F}) \cdot {}^t \{ ({}^t\mathbf{F} \cdot \mathbf{W} \cdot \mathbf{F})^{-1} \} \quad (33)$$

Concerning Equation (33), the expected error matrix (32) must be calculated using residuals caused by the original measurement data, not by the moving average data.

3.3. Propagation from measurement uncertainty

Here, we discuss the uncertainty variance of identified system parameters resulting from uncertainty of measurement with respect to temperature, solar radiation, tracer gas release rate, and gas concentration. We assume that the measurement uncertainty of x_i and g_i have respective instantaneous variances of σ_x^2 and σ_g^2 . The uncertainty variance ${}_b\sigma_x^2$ of the incremental calculation results is obtained from the difference of the measurement values at the two ends of the interval Δt , and is calculated from Equation (34) based on the propagation of the uncertainty variance. Furthermore, x_i and g_i are integrated along the time interval Δt and the integration is subjected to trapezoidal approximation, so Equations (35) and (36) give the respective uncertainty variances ${}_s\sigma_x^2$ and ${}_s\sigma_g^2$.

$${}_b\sigma_{xi}^2 = 2 \cdot \sigma_{xi}^2 \quad (34)$$

$${}_s\sigma_{xi}^2 = (1/2) \cdot \Delta t^2 \cdot \sigma_{xi}^2 \quad (35)$$

$${}_s\sigma_{gi}^2 = (1/2) \cdot \Delta t^2 \cdot \sigma_{gi}^2 \quad (36)$$

Here, the vector corresponding to the measurement data and the standard deviation vector of the measurement uncertainty are defined as follows.

$${}_b\mathbf{x}_k = {}^t({}_b\mathbf{x}_{1,k}, \dots, {}_b\mathbf{x}_{n,k}) \quad (37)$$

$${}_b\boldsymbol{\sigma}_x = {}^t({}_b\sigma_{x1}, \dots, {}_b\sigma_{xn}) \quad (38)$$

$${}_s\mathbf{x}_k = {}^t({}_s\mathbf{x}_{1,k}, \dots, {}_s\mathbf{x}_{n,k}, \dots, {}_s\mathbf{x}_{n+no,k}) \quad (39)$$

$${}_s\boldsymbol{\sigma}_x = {}^t({}_s\sigma_{x1}, \dots, {}_s\sigma_{xn}, \dots, {}_s\sigma_{xn+no}) \quad (40)$$

$${}_s\mathbf{g}_k = {}^t({}_s\mathbf{g}_{1,k}, \dots, {}_s\mathbf{g}_{ng,k}) \quad (41)$$

$${}_s\boldsymbol{\sigma}_g = {}^t({}_s\sigma_{g1}, \dots, {}_s\sigma_{gng}) \quad (42)$$

Furthermore, ${}_b\mathbf{x}_k$, ${}_s\mathbf{x}_k$, and ${}_s\mathbf{g}_k$ are considered true values with added respective uncertainties ${}_b\mathbf{s}_{xk}$, ${}_s\mathbf{s}_{xk}$, and ${}_s\mathbf{s}_{gk}$. If we assume that the estimated uncertainty of system parameters is due to only the uncertainty of measurements of x_j and g_j , then the state space equation error for the true values x_j and g_j can be regarded as 0.

$$\begin{aligned} {}_n\mathbf{e}_k &= -\mathbf{M} \cdot {}_b\mathbf{x}_k + [\mathbf{C} \quad \mathbf{C}_0] \cdot {}_s\mathbf{x}_k + \mathbf{R} \cdot {}_s\mathbf{g}_k \\ &= -\mathbf{M} \cdot {}_b\mathbf{s}_{xk} + [\mathbf{C} \quad \mathbf{C}_0] \cdot {}_s\mathbf{s}_{xk} + \mathbf{R} \cdot {}_s\mathbf{s}_{gk} \end{aligned} \quad (43)$$

Furthermore, if we assume that the state space equation error originates only from the uncertainty of measurement of x_j and g_j , then the expected matrix error $E({}_n\mathbf{e}_k \cdot {}^t{}_n\mathbf{e}_k)$ ($n \times n$) of the equation errors ${}_n\mathbf{e}_k$ at time step k is calculated as follows.

$$\begin{aligned} E({}_n\mathbf{e}_k \cdot {}^t{}_n\mathbf{e}_k) &= \text{diag} \left(\mathbf{M} \cdot E({}_b\mathbf{s}_{xk} \cdot {}^t{}_b\mathbf{s}_{xk}) \cdot {}^t\mathbf{M} + [\mathbf{C} \quad \mathbf{C}_0] \cdot E({}_s\mathbf{s}_{xk} \cdot {}^t{}_s\mathbf{s}_{xk}) \cdot {}^t[\mathbf{C} \quad \mathbf{C}_0] + \mathbf{R} \cdot E({}_s\mathbf{s}_{gk} \cdot {}^t{}_s\mathbf{s}_{gk}) \cdot {}^t\mathbf{R} \right) \\ &= \text{diag} \left(\mathbf{M} \cdot \text{diag}({}_b\sigma_x \cdot {}^t{}_b\sigma_x) \cdot {}^t\mathbf{M} + [\mathbf{C} \quad \mathbf{C}_0] \cdot \text{diag}({}_s\sigma_x \cdot {}^t{}_s\sigma_x) \cdot {}^t[\mathbf{C} \quad \mathbf{C}_0] + \mathbf{R} \cdot \text{diag}({}_s\sigma_g \cdot {}^t{}_s\sigma_g) \cdot {}^t\mathbf{R} \right) \end{aligned} \quad (44)$$

Here, we use the assumption that the expected value of the covariance of the uncertainties ${}_b\mathbf{s}_{xk}$, ${}_s\mathbf{s}_{xk}$, and ${}_s\mathbf{s}_{gk}$ is 0, and that the expected value of the covariance of the elements of these three vectors is also 0. Thus the expected error matrix ($na \times na$) of the quadratic form primary equation is calculated as follows.

$$E_m(\mathbf{e}_a \cdot {}^t\mathbf{e}_a) = \sum_{k=1}^{nt} {}^t\mathbf{Z}_k \cdot E({}_n\mathbf{e}_k \cdot {}^t\mathbf{e}_k) \cdot \mathbf{Z}_k \quad (45)$$

Regarding the expected error matrix ($na \times na$) for the quadratic form constraint equation, if measurement uncertainty is taken as the sole origin of equation error, the expected error of the measurement uncertainty is 0, and the following equation holds true.

$$E_m(\mathbf{e}_s \cdot {}^t\mathbf{e}_s) = 0 \quad (46)$$

From this, the expected error matrix $\mathbf{e} \cdot {}^t\mathbf{e}$ ($2na \times 2na$) for the composite regression equation is calculated as follows.

$$E_m(\mathbf{e} \cdot {}^t\mathbf{e}) = \begin{bmatrix} E_m(\mathbf{e}_a \cdot {}^t\mathbf{e}_a) & 0 \\ 0 & 0 \end{bmatrix} \quad (47)$$

By using Equation (47) instead of Equation (32), we can use Equation (48) to calculate the variance-covariance matrix ${}_m\Lambda_a$ ($na \times na$) of the estimated parameters arising only from measurement uncertainty.

$${}_m\Lambda_a = ({}^t\mathbf{F} \cdot \mathbf{W} \cdot \mathbf{F})^{-1} \cdot ({}^t\mathbf{F} \cdot \mathbf{W} \cdot E_m(\mathbf{e} \cdot {}^t\mathbf{e}) \cdot \mathbf{W} \cdot \mathbf{F}) \cdot {}^t \left\{ ({}^t\mathbf{F} \cdot \mathbf{W} \cdot \mathbf{F})^{-1} \right\} \quad (48)$$

3.4. Discrepancy ratio for system identification model premises

Failure to meet the premises of the system identification model in the actual phenomena is a more significant cause of identification uncertainty than is measurement uncertainty. To address this, the uncertainty variance of estimated system parameters arising from the equation residual described in the preceding section can be utilized. If variance arising from the residual is larger than that from measurement uncertainty, the premises of the system identification model structure, and its invariability, uniformity, or linearity, may be insufficiently established.

We therefore introduce an index to evaluate these discrepancies. Here, the j -th element of the diagonal in ${}_m\Lambda_a$ is denoted as ${}_m\sigma_{\lambda_j}^2$, and that of ${}_r\Lambda_a$ as ${}_r\sigma_{\lambda_j}^2$. Next, using the square roots of these elements, we define the ratio β by the following equation, which is referred to as the discrepancy ratio of the model premises.

$$\beta_j = \frac{{}_r\sigma_{\lambda_j}}{{}_m\sigma_{\lambda_j}} \quad (49)$$

The averaged value of β_j for all diagonal elements can be used as an evaluating index. When the discrepancy ratio is considerably larger than 1, there might be insufficient establishment of measurement premises for an appropriate system identification model. In this case we should also consider the coefficient of determination. If neither index is satisfactory, we should reconsider the system identification model, the moving average term of the measurement data, and the measurement procedure, including the magnitude and frequency of the excitation wave.

4. Optimum excitation and low-pass filter and measurement time intervals

The majority of system identification error is caused by approximation of the spatial discretization of that model. We can reduce this error by using a longer period T_p of sinusoidal excitation

[26] and an appropriate term T_a of the moving average. However, in the case of multi-zone infiltration measurement, the discrepancy between the actual phenomenon and the model lessens if we use mixing fans.

In practice, simultaneous tracer gas release in, or electric heating of, all rooms is difficult. Additionally, producing greater temperature or concentration differences is preferable between adjacent rooms, and the gas release or heating should be intermittent in all rooms. Sinusoidal excitations are applied from trough-to-trough with period T_p and with a smooth pattern to produce less high-frequency noise.

The stopping term, T_s , of gas release in multi-zone infiltration measurement can be broadly set to be between one-half and twice the optimum decay term [26].

The maximum rate of electric heating or gas release should be determined so as not to cause excessively high room temperatures or gas concentrations outside of the measurement range. Solar radiation acts as an excitation, not as noise, and so the maximum electric heating rate should be close to that of the solar radiation. In an actual measurement system, these values are controlled by state feedback.

In the case that the measurement interval is 1 min, by using the moving average term T_a as a low-pass filter, about 10–60 min is considered an adequate duration to reduce the effects of measurement uncertainty resulting in high-frequency noise. However, several hours or more are required to transform the measurement data into a form that is suitable for the rough system identification model of the heat transfer system. Lacking prior experience, we can search for an optimum T_a that gives a relatively large effective capacity $m_{i,i}$ and the coefficient of determination COD by numerical experimentation, because the $m_{i,i}$ of a rougher model should have a larger value to represent spatially wider material.

The present method of system identification requires the time integration of state variables along Δt , and these are approximated by linear interpolation. A shorter Δt is therefore preferable to improve the approximation. Moreover, in order to improve the accuracy of the moving term average, we should obtain as many measurements as possible within T_a by using shorter Δt . Therefore, the measurement interval Δt should be shorter than 10 min and 1 min is likely sufficient.

5. Verification case study

5.1. Models for simulated measurements

We assumed a two-story, two-zone building, 10 m wide in the east-west direction, 10 m wide in the south-north direction, and 6 m tall. The total floor area is 200 m². We conducted simulation measurements for multi-zone air flow rates and thermal performance. Building thermal and air flow rate network models are superimposed on the building section in Fig. 1. The south direction is leftward and the depth direction is east to west. The east and west sides have walls and windows, and the model is three-dimensional. The transmitting heat flow paths through east (E) and west (W) walls and windows are superimposed on this section. Concerning the thermal network model, Table 1 shows the thermal properties of the walls and windows. Changes in outdoor temperature, solar radiation, and electric heater output drive the thermal model. We also constructed a massive building model, supposing a reinforced concrete building with wall node thermal capacitances multiplied by 2.5, but same thermal conductance as the wooden building.

Only infiltration is considered in the air flow rate network model, with the 400 cm² total equivalent gap area equally distributed to the flow paths over external walls, and a leakage exponent of 1.5. During

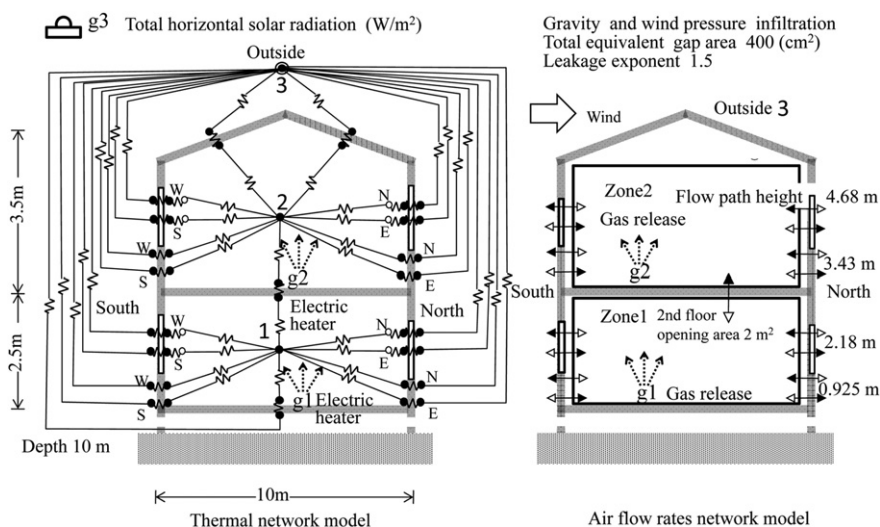


Fig. 1. Measurement simulation models.

simulated measurement, mechanical ventilation is stopped and influenced only by temperature differences and wind. Wind pressure coefficients are same as in the former study [25].

We used the computational simulation program NETS to obtain simulated measurement results, under standard meteorological data conditions for Tokyo with various electric heating and tracer gas release waveforms. According to our research at least one wave peak is needed for system identification. We also tested sinusoidal heating waves with a 72 h period, and therefore simulated 6 days of measurement data after a 3 day run starting on January 1. Fig. 2 shows changes in outdoor temperature and solar radiation for the 6 days, and Fig. 3 shows wind direction and velocity.

As described in Section 4, electric heater capacities were set at 8 kW for 6–24 h periods and 4 kW for 48–72 h periods. Tracer gas release capacities were set to 1 mg/s according to the assumed air exchange rate and the concentration measurement range for SF₆. Both of these were delivered as sinusoidal waves.

Time integration intervals for the heat transfer and tracer gas diffusion system were 30 s, and the measurement data are per-

minute results. We additionally added random error to the simulated measurement data, assuming measurement uncertainty standard deviations $m\sigma$ as 0.2 °C for temperature, 4 W for the electric heater, 0.2 mg/m³ for gas concentration, and 0.012 mg/s for gas release. Outdoor gas concentration was regarded as 0.

5.2. System identification model

Fig. 4 shows the system identification models for the heat transfer and tracer gas diffusion systems. For the gas diffusion system, we use a two-node model similar to the simulation measurement model. Here, the entire air flow rates c_{ij} are regarded as asymmetric, and the effective mixing volume $m_{i,i}$ is also identified. Known parameters are the gas release conversion coefficients r_{ij} of 3600.

Taking idealized measurements inside walls and surfaces is difficult when measuring temperature in actual buildings, so only room air-temperature measurements are assumed. The system identification model of heat transfer was therefore also taken as

Table 1
Thermal properties of building materials.

<p>Walls</p> <p>Out In</p>	<p>Roof</p> <p>Out In</p>
<p>1st floor</p>	<p>Glass and air layer</p> <p>Out In</p> <p>Window area of each floor 10 m²</p>
<p>2nd floor</p>	<p>Air layer of 2nd floor 5.81 W/m²K</p> <p>Air layer of walls and roof 11.11 W/m²K</p> <p>Air layer of windows 11.11 W/m²K</p> <p>equivalent conductance 14.49 W/m²K</p>
	<p>Plywood Glass wool Glass Air</p>
Density	550 kg/m ³ 10 kg/m ³ 2500 kg/m ³ 1.3 kg/m ³
Specific heat	1.30 kJ/kgK 0.84 kJ/kgK 0.75 kJ/kgK 1.0 kJ/kgK
Thermal conductivity	0.186 W/mK 0.06 W/mK 0.78 W/mK
<p>Solar thermal property: Building outer surface absorptance is 0.80, each glass transmittance is 0.87 and reflectance is 0.08 at normal incidence angle, and for Shoji screen 0.25 and 0.60 respectively.</p>	

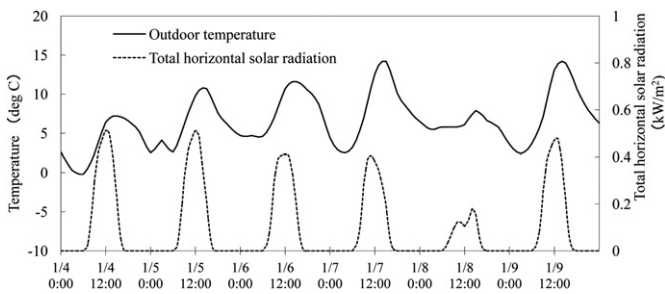


Fig. 2. Outdoor temperature and solar radiation.

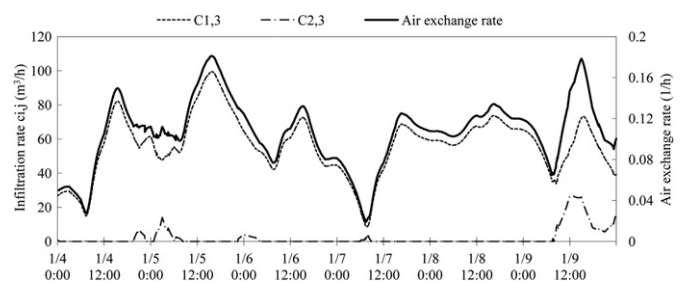


Fig. 5. Air exchange rate and infiltration rates in a wooden building.

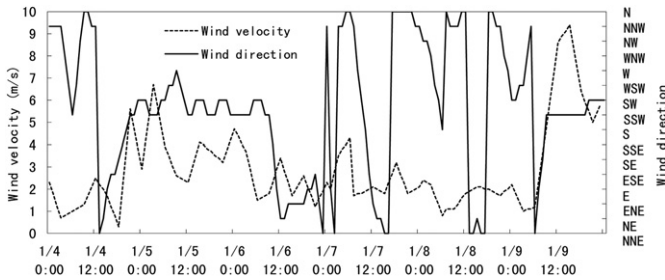


Fig. 3. Wind velocity and direction.

time interval for system identification was set to 1 min for both the heat transfer and gas diffusion systems.

5.3. True values or references of system identification

Although some measurement parameters such as air infiltration rate fluctuate to a certain extent, these values are estimated by the system identification as averages over the measurement term. These averages are thus regarded as references for comparison. For example, the changes in infiltration and air exchange rate in Fig. 5 show approximately $\pm 60\%$ fluctuations around the term averages.

The estimated generalized thermal conductance c_{ij} of the system identification can be considered as the sum of two types of c_{ij} , wall heat transmission and convection due to infiltration through the wall. Such overall thermal conductance c_{ij} explains the change in Fig. 6. The convection portion is rather small compared to transmission, and thus the overall change in conductance is also small. NETS has a heating and cooling load calculation function. By utilizing this function only to the thermal model and excluding the air infiltration model, all room temperatures are maintained at 1°C with the outdoor 0°C condition, so the thermal load through the exterior wall can be computed in steady state and consequently the conductances are calculated. Inter-zone wall transmission conductance can be computed by maintaining 1°C for the opposite room and 0°C for the other all rooms including 0°C for the outdoor condition. The convection portion is computed from the infiltration

a two-node model, in which the thermal conductance c_{ij} , the effective thermal capacity $m_{i,i}$, and the solar heat gain coefficient r_{ij} are identified. The solar heat gain coefficient is $r_{1,3}$ (m^2) for the first floor and $r_{2,3}$ (m^2) for the second floor, and the horizontal total solar radiation is defined as g_3 (W/m^2). Known parameters are the electricity to heat conversion coefficient r_{ij} .

Since the generalized conductance c_{ij} of infiltration-induced convection is asymmetric, c_{ij} should also ideally be identified as asymmetric in the heat transfer system. If we assume that the convection portion can be obtained from the infiltration measurement, it should be possible to calculate the heat transmission portion by subtraction. Actual attempts to do so, however, resulted in unreasonable results due to various errors, and thus we applied symmetry constraints to the thermal conductance c_{ij} . The sampling

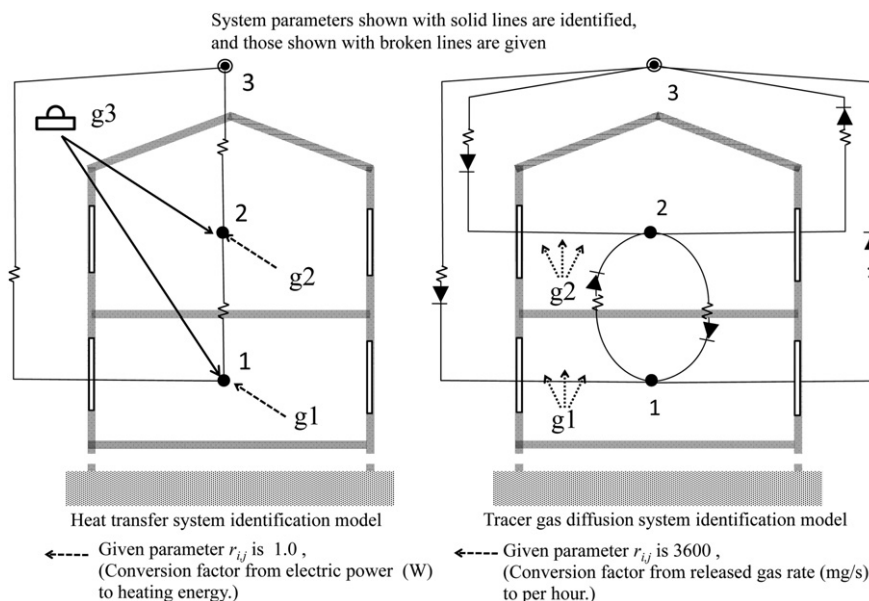


Fig. 4. System identification models.

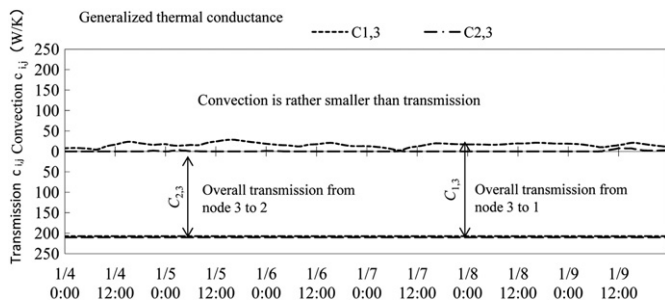


Fig. 6. Overall heat transfer generalized conductance c_{ij} in a wooden building.

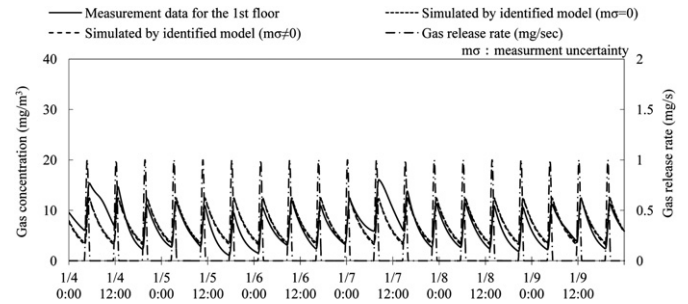


Fig. 7. Tracer gas release and concentration change for first floor in a wooden building.

rate average for the measurement term. These convection and transmission portions are summed as the overall conductance. The heat loss coefficient Q , often used in Japan, is calculated as the heating load necessary to maintain an indoor temperature of $1\text{ }^\circ\text{C}$ at a constant outdoor temperature of $0\text{ }^\circ\text{C}$ and dividing it by the total floor area. The heating load includes the actual air infiltration load.

The solar heat gain coefficient r_{ij} represents the conversion factor from the horizontal total solar radiation to the heating energy for the air in each room, but the actual coefficient varies, making it difficult to determine its reference value. Using NETS, maintaining both the indoor and outdoor temperatures at $0\text{ }^\circ\text{C}$ during the measurement term, we can compute the unsteady-state cooling load caused only by solar radiation. Taking the solar heat load averaged over the measurement term as the numerator and the averaged total horizontal solar radiation as the denominator, the ratio s_{gj} is considered as a reference for comparison.

Since it is assumed that mixing fans are used during measurement, the zone's geometrical volume can be regarded as a reference for the effective mixing volume $m_{i,i}$ of the gas diffusion system. However, it is difficult to determine a reference for the effective thermal capacity $m_{i,i}$ for the heat transfer system's zone nodes. We can consider that a certain portion of the thermal capacity of the walls is included in this effective thermal capacity $m_{i,i}$. For a massive building, the estimated value should be larger than that for a wooden building.

5.4. Verification of system identification for multi-zone air flow rate measurements

The structural discrepancy between the model and the actual phenomenon is small in this measurement. The major causes of system identification error are changes in infiltration over time and

measurement uncertainty. We tried four sinusoidal periods $T_p = 0.5, 1.5, 3,$ and 6 (h) for the gas release and five moving average terms $T_a = 10, 30, 60, 90,$ and 120 (min) under conditions of added measurement uncertainty. Thus $4 \times 5 = 20$ test calculations were performed. The gas release stopping term T_s was set at 6 h by a recent study [26]. The electric heating sinusoidal period was 72 h. We obtained optimum results of $COD = 0.952$ and $\beta = 1.21$ with $T_p = 1.5$ h and $T_a = 60$ min, so these conditions were used in the present case.

Table 2 shows system identification results for multi-zone air flow rate measurements. The system identification carries two conditions, that $m\sigma = 0$ with no measurement uncertainty, and $m\sigma \neq 0$ for non-zero measurement uncertainty.

The identified air flow rates c_{ij} are satisfactory in comparison with the term-averaged c_{ij} , and the results for the effective mixing volume $m_{i,i}$ are also satisfactory. Because the infiltration rates c_{ij} show some fluctuation in the measurement term, the discrepancy ratio β of the measurement premises are greater than 1 (about 1.21). As for the estimated standard deviations, ${}_r\sigma_\lambda$ of the identified parameters are appropriate in comparison with the actual deviations from the term-averaged parameters. Figs. 7 and 8 show the first and second floor, respectively, and each shows a graph comparing the concentration simulated by a model consisting of the identified system parameters with the concentration according to measurement. The variation in infiltration and air exchange rates over time is large, so the identified model with term-averaged parameters shows large discrepancies for some terms.

5.5. Verification of system identification in a heat transfer system

Most heat transfer system identification models have structural discrepancies in comparison with the actual phenomena, and it is

Table 2 System identification results for multi-zone infiltration rates measurement.

Tracer gas diffusion system parameter	Term ave.	System identification result		COD		Standard deviation, ${}_r\sigma_\lambda$ from the equation residual		Standard deviation $m\sigma_\lambda$ from the measurement uncertainty		Discrepancy ratio β of model premises	
		$m\sigma = 0$	$m\sigma \neq 0$	$m\sigma = 0$	$m\sigma \neq 0$	$m\sigma = 0$	$m\sigma \neq 0$	$m\sigma \neq 0$	$m\sigma \neq 0$	$m\sigma \neq 0$	$m\sigma \neq 0$
$c_{3,2}$	57.41	53.29	55.31	0.9641	0.9522	4.414	14.72	10.95		1.344	
$c_{3,1}$	1.323	0.000	0.000			7.919	26.44	19.81		1.335	
$c_{2,3}$	2.094	1.920	3.100			3.860	10.59	8.870		1.194	
$c_{1,3}$	56.56	51.37	52.21			2.513	8.161	7.063		1.155	
$c_{2,1}$	55.58	53.00	52.63			6.891	19.64	16.00		1.228	
$c_{1,2}$	0.000	1.630	0.420			2.396	8.645	6.647		1.301	
$m_{2,2}$	300.0	291.9	285.2			9.591	23.37	21.27		1.099	
$m_{1,1}$	250.0	245.4	241.3			5.396	16.02	15.16		1.056	
Air exchange rate(1/hour)	0.107	0.100	0.100			—	—	—		—	
Average value	—	—	—	—	—	—	—	—		1.214	

Note: c_{ij} : air flow rate(m^3/h) from zone j to zone i . Here, zone 3 corresponds to the outdoor, $m_{i,i}$: effective mixing volume of the i -th zone (m^3), $m\sigma$: measurement uncertainty standard deviation.

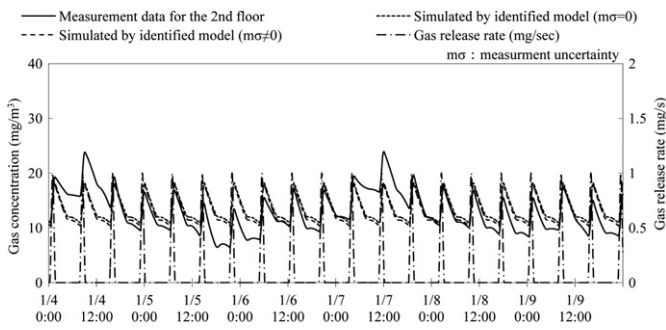


Fig. 8. Tracer gas release and concentration change for second floor in a wooden building.

usually difficult to obtain satisfactory identification. We therefore tested various periods of sinusoidal heating with electric heaters and moving average terms to eliminate the negative effect of these discrepancies.

To clarify the effect of these periods we adopted a condition of no measurement uncertainty. We tried sinusoidal excitation

periods of $T_p = 6, 12, 24, 48,$ and 72 h, and for the moving average terms of $T_a = 1, 3, 6, 8, 12, 18,$ and 24 h. We carried out numerical experiments for these $5 \times 7 = 35$ cases in the wooden building, and compared the resulting coefficients of determination COD and effective heat capacities $m_{i,j}$. However some unreasonable results were observed in 5 cases with $T_a = 24$, in addition to the 3 cases with $(T_p = 6, T_a = 12), (T_p = 6, T_a = 18)$ and $(T_p = 12, T_a = 18)$. We observed a general tendency of increasing COD with increases in T_p and T_a . For each T_p greater than 12 h, a relatively large $m_{i,j}$ occurred at $T_a = 8$ h. The conditions $T_p = 72$ h and $T_a = 8$ h gave the optimum results with $COD = 0.9940$, and these conditions were thus used in the present case.

We next carried out verification experiments for measurement uncertainty. Tables 3 and 4 show the system identification results for the wooden and massive building, respectively. Because of structural discrepancies of the system identification model, the index β is slightly larger than 1. We observe that the estimated Q value is slightly smaller than the true value with about 1.3% error for wooden and 4.5% for massive building. Almost all estimated parameters $c_{i,j}$ are reasonable.

The identified solar heat gain coefficient $r_{i,j}$ was slightly smaller than the solar heat load coefficient $S_{g,i}$. Nevertheless, these

Table 3
System identification results for thermal performance of a wooden building.

Wooden building system parameter	Term Ave. (transmission + convection)	System identification result		COD		Standard deviation, r, σ_λ from the equation residual		Standard deviation m, σ_λ from the measurement uncertainty		Discrepancy ratio β of model premises	
		$m\sigma = 0$	$m\sigma \neq 0$	$m\sigma = 0$	$m\sigma \neq 0$	$m\sigma = 0$	$m\sigma \neq 0$	$m\sigma \neq 0$	$m\sigma \neq 0$	$m\sigma \neq 0$	
$c_{3,2}$	229.5(210.3 + 19.23)	206.1	205.9	0.9940	0.9934	4.325	37.79	37.92		0.997	
$c_{3,1}$	207.7(207.3 + 0.443)	224.6	225.4			3.99	37.41	37.08		1.009	
$c_{2,3}$	211.0(210.3 + 0.701)	206.1	205.9			4.325	37.79	37.92		0.997	
$c_{1,3}$	226.2(207.3 + 18.94)	224.6	225.4			3.99	37.41	37.08		1.009	
$c_{2,1}$	193.0(174.4 + 18.61)	187.4	186.7			3.70	33.05	33.21		0.995	
$c_{1,2}$	174.4(174.4 + 0.000)	187.4	186.7			3.70	33.05	33.21		0.995	
$m_{1,1}$	mt	3307	3319			125.0	1168	1156.9		1.010	
$m_{2,2}$		3244	3283			135.3	1183	1187.5		0.996	
$r_{1,3}$	S_{g1}	9.871	9.935			0.335	3.135	3.107		1.009	
$r_{2,3}$	S_{g2}	8.954	8.932			0.3729	3.259	3.269		0.997	
Total $c_{i,j}$	437.2	430.7	431.3	–	–	–	–	–		–	
Q value	2.186	2.153	2.156	–	–	–	–	–		–	
Average value	–	–	–	–	–	–	–	–		1.001	

Note: $S_{g1} : 10.40, S_{g2} : 9.510$. mt : Total thermal capacity for all nodes 10080(kj/K), the identified value $c_{i,j}$ is the sum of the convection and the transmission. $c_{i,j}$: thermal conductance (W/K) from j to i . Here, imposing the symmetry constraint $c_{i,j} = c_{j,i}$. $m_{i,j}$: thermal capacity of node i (kj/K), $r_{i,j}$: solar heat gain coefficient(m²) $S_{g,j}$: Solar thermal load coefficient(m²), Total $c_{i,j}$: sum of $c_{2,3}$ and $c_{1,3}$ (W/K), Q value : (W/Km², per floor area 200 m²) m, σ : measurement uncertainty standard deviation.

Table 4
System identification results for thermal performance of a massive building.

Massive building system parameter	Term Ave. (transmission + convection)	System identification result		COD		Standard deviation r, σ_λ from the equation residual		Standard deviation m, σ_λ from the measurement uncertainty		Discrepancy ratio β of model premises	
		$m\sigma = 0$	$m\sigma \neq 0$	$m\sigma = 0$	$m\sigma \neq 0$	$m\sigma = 0$	$m\sigma \neq 0$	$m\sigma \neq 0$	$m\sigma \neq 0$	$m\sigma \neq 0$	
$c_{3,2}$	229.6(210.3 + 19.31)	199.7	198.1	0.9913	0.9889	17.06	75.26	73.17		1.029	
$c_{3,1}$	207.8(207.3 + 0.535)	221.1	219.7			17.43	80.37	78.79		1.020	
$c_{2,3}$	211.0(210.3 + 0.718)	199.7	198.1			17.06	75.26	73.17		1.029	
$c_{1,3}$	226.4(207.3 + 19.10)	221.1	219.7			17.43	80.37	78.79		1.020	
$c_{2,1}$	193.1(174.4 + 18.68)	191.0	191.6			15.53	71.48	69.60		1.027	
$c_{1,2}$	174.4(174.4 + 0.000)	191.0	191.6			15.53	71.48	69.60		1.027	
$m_{1,1}$	mt	7229	7122			1182	5423	5318		1.020	
$m_{2,2}$		6784	6659			1094	4781	4649		1.028	
$r_{1,3}$	S_{g1}	9.568	9.427			1.600	7.380	7.239		1.020	
$r_{2,3}$	S_{g2}	8.621	8.457			1.571	6.881	6.691		1.028	
Total $c_{i,j}$	437.4	420.8	417.8	–	–	–	–	–		–	
Q value	2.187	2.104	2.089	–	–	–	–	–		–	
Average value	–	–	–	–	–	–	–	–		1.025	

Note: $S_{g1} : 10.40, S_{g2} : 9.889$. mt : Total thermal capacity for all nodes 23016.5(kj/K). m, σ : measurement uncertainty standard deviation.

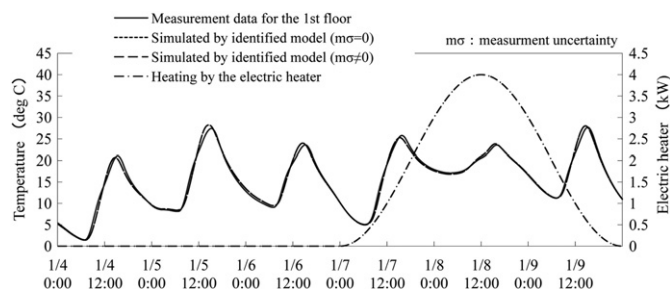


Fig. 9. Comparison of indoor temperature change for first floor in a wooden building.

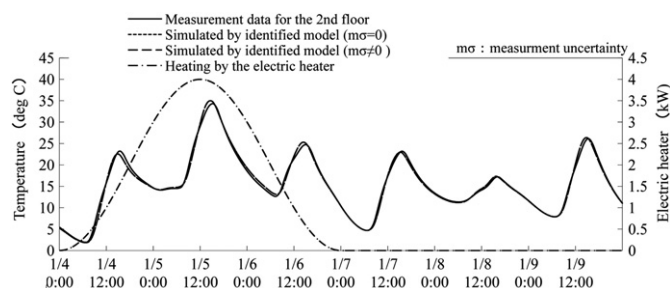


Fig. 10. Comparison of indoor temperature change for second floor in a wooden building.

constants $r_{i,j}$ enable closely tracing measurement room temperature changes. Furthermore, the total identified effective thermal capacity $m_{i,i}$ was about half the cumulative thermal capacity for all nodes. This result is appropriate if we consider that half of the wall thickness contributes to the air nodes in the room. The reliability index COD for the wooden building was larger than that for the massive building. Estimated uncertainty standard deviations of system parameters are satisfactory in predicting actual deviations from the true values.

Figs. 9 and 10 show comparisons between simulated room temperature changes using the identified model and the measurement temperature changes. Those figures respectively show first and second floor room temperature changes for the wooden building. Even though the model was constructed of only two nodes, the results are almost identical.

The optimum combination of the periods T_p and T_a can be numerically searched for relatively large COD and effective heat capacities $m_{i,i}$. However searches implementing some theoretical methodology are likely more practical. Longer sinusoidal excitation periods seem to give better results, though methods that reduce measurement time are also desirable.

6. Conclusions

A generalized system identification theory applicable to measurement of building heat transfer and infiltration systems, uncertainty analysis, and reliability evaluation theories was deduced and verified through computational case study. The main findings and results of this study are given below.

(a) Improvements of the estimation methods for system parameters of building heat transfer and infiltration systems.

We deduced a composite regression equation comprised of two types of quadratic form equation. One is a quadratic form primary equation derived from the state space equation, and the other is a quadratic form constraint equation. Each was obtained by least

squares. The estimation equation for system parameters was deduced by least squares for the composite regression equation. The final estimation equation was thus deduced through a double application of least squares. These formulations realize a consistency of physical units in the composite equation, and also an extensibility of the constraints. We introduced a weighting matrix \mathbf{W} for unbiased estimation, and the estimation accuracy for the system parameters and the uncertainty evaluation indices were improved.

(b) Uncertainty and reliability analysis methods

We deduced an error propagation equation from the expected error matrix of the composite regression equation to the uncertainty variances of the estimated parameters. Low-pass filtered measurement data should be used to estimate system parameters, but source data must be used for uncertainty analysis. We defined the discrepancy ratio β of the model premises. The β ratio is more sensitive than the coefficient of determination COD . Good judgment considering use of both indices is recommended.

(c) Excitation waveform suitable for system identification

A long-period sinusoidal excitation enabled us to obtain reasonable results for heat transfer system identification, even for a somewhat rough model with few nodes and constant solar heat gain coefficients. Despite such roughness, we estimated the heat loss coefficient Q with tolerable error. The optimum sinusoidal period T_p can be determined in preliminary experiments and the like with the coefficient of determination COD . The optimum decay or stopping term T_s of the recent study [26] is also helpful.

(d) Low-pass filter for the measurement data

A moving average of term T_a is required as a low-pass filter for pre-processing measurement data, for the purposes of reducing the adverse effect of high-frequency measurement noise and the spatial discretization discrepancy between actual phenomenon and the model. In the present case study, we numerically searched for an optimum T_a realizing the relatively large effective capacities $m_{i,i}$ and COD .

(e) Remaining problems

In relation to heat transfer systems, some problems remain, namely, a theoretical method is needed for finding the optimum period T_p of sinusoidal excitation and the moving average term T_a . A method that reduces the measurement term would furthermore improve practicality.

We conclude with a note about the expected scope of application of these methods. We presume that the present methods will be applied mainly to detached or multi-unit residential buildings. In situations where heating and cooling installations are in operation, or in cases where such installations serve as an excitation source, one should measure the supply heat flux and precisely control its sinusoidal wave.

Acknowledgments

We are grateful to Prof. Hiroshi Yoshino, Prof. Shinsuke Kato, Prof. Takashi Kurabuchi, and Prof. Toshio Yamanaka for giving us the research motivation and opportunity in the Committee of the Society for Heating, Air-Conditioning and Sanitary Engineering in Japan. In the relevant experimental measurements, we thank Dr. Takao Sawachi, Dr. Hiroataka Suzuki and Prof. Shintaro Yokoyama for giving us opportunities and assistance. In the committee of the

Architectural Institute of Japan, we are grateful to Prof. Yoh Matsuo, Prof. Mitsuhiro Udagawa, Prof. Harunori Yoshida, Prof. Akihiro Nagata, Prof. Tatsuo Nagai, Prof. Kazuo Emura, Prof. Masanori Shukuya and Prof. Ken-ichi Hasegawa for their valuable discussions and opportunities. We are grateful to Prof. Ken-ichi Kimura and Prof. Shin-ichi Tanabe for their advice on research directions. We thank Mr. Tomohisa Mashiko of Japan Industrial Testing, Ltd. for developing and adjusting the computational programs.

Nomenclature

m_{ij} generalized capacitance of node i
 c_{ij} generalized conductance from node j to node i
 r_{ij} generalized flux parameter from source j to node i
 x_i diffusion potential state of node i
 g_i flux from source i
 n total state dependent nodes
 no total state independent nodes
 ng total flux sources
 \mathbf{x} vector of state dependent nodes (n)
 \mathbf{x}_o input vector of state independent nodes (no)
 \mathbf{g} input flux vector (ng)
 \mathbf{M} capacitance matrix ($n \times n$)
 \mathbf{C} conductance matrix ($n \times n$)
 \mathbf{C}_o input conductance matrix ($n \times no$)
 \mathbf{R} input flux matrix ($n \times ng$)
 \mathbf{m} capacitance vector to be identified (nm)
 \mathbf{c} conductance vector to be identified (nc)
 \mathbf{r} flux parameter vector to be identified (nr)
 \mathbf{a} system parameter vector to be identified composed of ${}^t(\mathbf{m} \ \mathbf{c} \ \mathbf{r})$ ($nm + nc + nr$)
 $\hat{\mathbf{a}}$ estimated system parameter vector ($nm + nc + nr$)
 \mathbf{D} matrix containing time differential of state x ($n \times nm$)
 \mathbf{X} matrix containing state of x ($n \times nc$)
 \mathbf{G} matrix containing flux of g ($n \times ng$)
 \mathbf{Z} composite matrix of \mathbf{D} , \mathbf{X} and \mathbf{G} , [$\mathbf{D} \ \mathbf{X} \ \mathbf{G}$]
 \mathbf{y} observation vector containing products of given parameters and variables (n)
 k time step subscript
 Δt time increment of system identification
 T system identification term
 ${}^n\mathbf{e}_k$ primary regression equation error vector at time step k
 J_n least-squares evaluation function for primary regression equation
 nt total time steps of the system identification period
 na dimension of system parameters vector \mathbf{a} ($nm + nc + nr$)
 ns dimension of primary constraint equation
 \mathbf{e}_a quadratic form primary equation error vector (na) in composite regression equation
 \mathbf{S} first-degree constraint equation matrix ($ns \times na$)
 \mathbf{d} first-degree constraint equation vector (ns)
 \mathbf{e}_d first-degree constraint equation error (ns) for quadratic form constraint equation
 J_d least-squares evaluation function for quadratic form constraint equation
 \mathbf{e}_s quadratic form constraint equation error vector (na) in composite regression equation
 \mathbf{W}_a weighting matrix for matrix $\mathbf{H}_a = \sum {}^t\mathbf{Z}_k \cdot \mathbf{Z}_k$ in quadratic form primary equation ($na \times na$)
 \mathbf{W}_s weighting matrix for matrix $\mathbf{H}_s = \Delta t^2 \cdot {}^t\mathbf{S} \cdot \mathbf{S}$ in quadratic form constraint equation ($na \times na$)
 \mathbf{W} composite weighting matrix of \mathbf{W}_a and \mathbf{W}_s ($2na \times 2na$)
 \mathbf{e} composite error vector of ${}^t(\mathbf{e}_a, \mathbf{e}_s)$ ($2na$)
 \mathbf{F} composite matrix ($2na \times na$) $\sum {}^t\mathbf{Z}_k \cdot \mathbf{Z}_k$ and $\Delta t^2 \cdot {}^t\mathbf{S} \cdot \mathbf{S}$

\mathbf{b} composite vector ($2na$) $\sum {}^t\mathbf{Z}_k \cdot \mathbf{y}_k$ and $\Delta t^2 \cdot {}^t\mathbf{S} \cdot \mathbf{d}$
 J least-squares evaluation function for composite regression equation
 \mathbf{v}_k residual vector (n) of primary regression equation at time step k
 $E^{(*)}$ stochastic expectation operator for variables $*$, from the equation residual is $E_r^{(*)}$, and from the measurement uncertainty is $E_m^{(*)}$
 $s(\hat{\mathbf{a}})$ sum of residual square of primary regression equation
 s_y total variation of primary regression equation
 COD coefficient of determination
 Λ_a system parameters uncertainty variance and covariance matrix, from the composite regression equation residual is ${}_r\Lambda_a$, and from the measurement uncertainty is ${}_m\Lambda_a$
 σ_x measurement uncertainty standard deviation for x
 σ_g measurement uncertainty standard deviation for g
 ${}^n\mathbf{e}_k$ primary regression equation error vector at time step k from the measurement uncertainty
 ${}^r\sigma_{\lambda_j}$ estimated standard deviation of the system parameter λ_j corresponding to the j -th system parameter in the vector \mathbf{a} which uncertainty comes from the composite regression equation residual and the original variance is taken from the j -th diagonal element of matrix ${}_r\Lambda_a$
 ${}^m\sigma_{\lambda_j}$ estimated standard deviation of system parameter λ_j corresponding to the j -th system parameter in the vector \mathbf{a} which uncertainty comes from the measurement uncertainty and the original variance is taken from the j -th diagonal element of matrix ${}_m\Lambda_a$
 β_j system identification model premises discrepancy ratio of j -th system parameter in vector \mathbf{a}
 s_{g_i} solar thermal load coefficient(m^2), detail is described in 5.3.
 m_t total thermal capacity for all nodes(kJ/K)
 T_p sinusoidal excitation period
 T_s excitation stopping or decay term
 T_a moving average term

Appendix A

There exists a known method of linearizing approximation for long wave radiation heat transfer by the temperature difference between both surfaces. The equivalent heat transfer coefficient α_r for the radiation is defined by Equation (A.1):

$$\alpha_r = 0.04 \cdot \epsilon_i \cdot \epsilon_j \cdot c_b \cdot (x_m/100)^3 \tag{A.1}$$

where ϵ_i is the emissivity of the i -th surface, c_b is the constant of Stefan–Boltzmann law, and x_m is the average absolute temperature of both surfaces.

The generalized thermal conductance c_{ij} and c_{ji} can be defined by Equation (A.2) with area s_i of the i -th surface, and form factor $f_{i,j}$ viewing from the i -th surface to the j -th surface. The order of subscripts i and j follows the standard definition. These c_{ij} should be recalculated following changes of average temperatures of both surfaces.

$$c_{ij} = \alpha_r \cdot s_j \cdot f_{j,i} = \alpha_r \cdot s_i \cdot f_{i,j} = c_{j,i} \tag{A.2}$$

Infinite reflections between the inner surfaces of a space can be modeled by introducing an equivalent emissivity [29]. All heat transfer forms of conduction, convection, transmission, and this radiation are represented by the unified generalized conductance c_{ij} and realizes universal applicability with the complete linked nodal Equation (1).

Appendix B

The vector of the uncertainty of identification of system parameters is expressed with the following equation.

$$\begin{aligned} \hat{\mathbf{a}} - E(\hat{\mathbf{a}}) &= ({}^t\mathbf{F} \cdot \mathbf{W} \cdot \mathbf{F})^{-1} \cdot ({}^t\mathbf{F} \cdot \mathbf{W} \cdot \mathbf{b}) - E\left\{ ({}^t\mathbf{F} \cdot \mathbf{W} \cdot \mathbf{F})^{-1} \cdot [{}^t\mathbf{F} \cdot \mathbf{W} \cdot (\mathbf{F} \cdot \mathbf{a} + \mathbf{e})] \right\} \\ &= \hat{\mathbf{a}} - \hat{\mathbf{a}} - ({}^t\mathbf{F} \cdot \mathbf{W} \cdot \mathbf{F})^{-1} \cdot ({}^t\mathbf{F} \cdot \mathbf{W} \cdot \mathbf{e}) \\ &= - ({}^t\mathbf{F} \cdot \mathbf{W} \cdot \mathbf{F})^{-1} \cdot ({}^t\mathbf{F} \cdot \mathbf{W} \cdot \mathbf{e}) \end{aligned} \quad (\text{B.1})$$

From this, Equation (23) can be derived, which expresses the uncertainty variance-covariance matrix of the system parameters.

$$\begin{aligned} \Lambda_a &= E[(\hat{\mathbf{a}} - E(\hat{\mathbf{a}})) \cdot {}^t(\hat{\mathbf{a}} - E(\hat{\mathbf{a}}))] \\ &= E\left[({}^t\mathbf{F} \cdot \mathbf{W} \cdot \mathbf{F})^{-1} \cdot ({}^t\mathbf{F} \cdot \mathbf{W} \cdot \mathbf{e}) \cdot {}^t\left\{ ({}^t\mathbf{F} \cdot \mathbf{W} \cdot \mathbf{F})^{-1} \cdot ({}^t\mathbf{F} \cdot \mathbf{W} \cdot \mathbf{e}) \right\} \right] \\ &= E\left[({}^t\mathbf{F} \cdot \mathbf{W} \cdot \mathbf{F})^{-1} \cdot ({}^t\mathbf{F} \cdot \mathbf{W} \cdot \mathbf{e}) \cdot ({}^t\mathbf{e} \cdot \mathbf{W} \cdot \mathbf{F})^t \cdot \left\{ ({}^t\mathbf{F} \cdot \mathbf{W} \cdot \mathbf{F})^{-1} \right\} \right] \\ &= ({}^t\mathbf{F} \cdot \mathbf{W} \cdot \mathbf{F})^{-1} \cdot {}^t\mathbf{F} \cdot \mathbf{W} \cdot E(\mathbf{e} \cdot {}^t\mathbf{e}) \cdot \mathbf{W} \cdot \mathbf{F} \cdot {}^t\left\{ ({}^t\mathbf{F} \cdot \mathbf{W} \cdot \mathbf{F})^{-1} \right\} \end{aligned} \quad (\text{B.2})$$

Appendix C

Starting from Equation (10), the process for deriving Equation (27) is as follows.

$$\begin{aligned} E(\mathbf{e}_a \cdot {}^t\mathbf{e}_a) &= E\left(\left(\sum_{k=1}^{nt} {}^t\mathbf{z}_k \cdot \mathbf{y}_k - \sum_{k=1}^{nt} {}^t\mathbf{z}_k \cdot \mathbf{z}_k \cdot \mathbf{a} \right) \cdot {}^t \left(\sum_{k=1}^{nt} {}^t\mathbf{z}_k \cdot \mathbf{y}_k - \sum_{k=1}^{nt} {}^t\mathbf{z}_k \cdot \mathbf{z}_k \cdot \mathbf{a} \right) \right) \\ &= E\left(\left(\sum_{k=1}^{nt} {}^t\mathbf{z}_k \cdot \mathbf{y}_k - \sum_{k=1}^{nt} {}^t\mathbf{z}_k \cdot \mathbf{z}_k \cdot \mathbf{a} \right) \cdot \left(\sum_{k=1}^{nt} {}^t\mathbf{y}_k \cdot \mathbf{z}_k - \sum_{k=1}^{nt} {}^t\mathbf{a} \cdot {}^t\mathbf{z}_k \cdot \mathbf{z}_k \right) \right) \\ &= E\left(\sum_{k=1}^{nt} {}^t\mathbf{z}_k \cdot \mathbf{y}_k \cdot {}^t\mathbf{y}_k \cdot \mathbf{z}_k - \sum_{k=1}^{nt} {}^t\mathbf{z}_k \cdot \mathbf{y}_k \cdot {}^t\mathbf{a} \cdot {}^t\mathbf{z}_k \cdot \mathbf{z}_k - \sum_{k=1}^{nt} {}^t\mathbf{z}_k \cdot \mathbf{z}_k \cdot \mathbf{a} \cdot {}^t\mathbf{y}_k \cdot \mathbf{z}_k + \sum_{k=1}^{nt} {}^t\mathbf{z}_k \cdot \mathbf{z}_k \cdot \mathbf{a} \cdot {}^t\mathbf{a} \cdot {}^t\mathbf{z}_k \cdot \mathbf{z}_k \right) \\ &= E\left(\sum_{k=1}^{nt} {}^t\mathbf{z}_k \cdot (\mathbf{y}_k \cdot {}^t\mathbf{y}_k - \mathbf{y}_k \cdot {}^t\mathbf{a} \cdot {}^t\mathbf{z}_k - \mathbf{z}_k \cdot \mathbf{a} \cdot {}^t\mathbf{y}_k + \mathbf{z}_k \cdot \mathbf{a} \cdot {}^t\mathbf{a} \cdot {}^t\mathbf{z}_k) \cdot \mathbf{z}_k \right) \\ &= E\left(\sum_{k=1}^{nt} {}^t\mathbf{z}_k \cdot ((\mathbf{y}_k - \mathbf{z}_k \cdot \mathbf{a}) \cdot {}^t(\mathbf{y}_k - \mathbf{z}_k \cdot \mathbf{a})) \cdot \mathbf{z}_k \right) = \sum_{k=1}^{nt} {}^t\mathbf{z}_k \cdot E(\mathbf{v}_k \cdot {}^t\mathbf{v}_k) \cdot \mathbf{z}_k \end{aligned} \quad (\text{C.1})$$

References

- [1] Matsuo Y, Saito H. Thermal characteristics estimation of residential building based on on-site measurements. *AIJ J Build Environ Eng* April 1981;3:13–8.
- [2] Emura K, Saito H. A study on test method of the thermal insulation performance of room based on field measurement. *J Architecture, Plann Environ Eng (Transactions AIJ)* Aug.1985;354:32–41.
- [3] Emura K, Saito H. A study on test method of the thermal insulation performance of room based on field measurement (part 2). *Plann Environ Eng (Transactions AIJ)* Nov. 1986;369:1–11.
- [4] Akashi Y, Matsuo Y, Nagata A, Watanabe T. Building thermal load and thermal characteristic identification. *Plann Environ Eng (Transactions AIJ)* Aug. 1993; 450:19–27.
- [5] Hattori T, Sakamoto Y, Fukuda H. Developments and examples related to practical field measurements of the heat loss coefficients of detached houses. *AIJ J Technol Des* 2008;14:491–6.
- [6] Zaheer-uddin M. Combined energy balance and recursive least squares method for the identification of system parameters. *ASHRAE Trans* 1990;96: 239–44.
- [7] Dewson T, Day B, Irving AD. Least squares parameter estimation of a reduced order thermal model of an experimental building. *Build Environ* 1993;28: 127–37.
- [8] Norlen U. Estimating thermal parameters of outdoor test cells. *Build Environ* 1990;25:17–24.
- [9] Crawford RR, Woods JE. A method for deriving a dynamic system model from actual building performance data. *ASHRAE Trans* 1985;91:1859–74.
- [10] Baudier M, Marchio D. Dynamic model identification applied to the measuring of thermal static characteristics of buildings. *Energ Build* 1991;17:21–34.
- [11] Marquardt DW. An algorithm for least-squares estimation of non-linear parameters. *J Soc Indust Appl Math*; June 1983.
- [12] Wang S, Xu X. Simplified building model for transient thermal performance estimation using GA-based parameter identification. *Int J Thermal Sci* 2006; 45:419–32.
- [13] Mitchell M. An Introduction to genetic algorithm. Cambridge, MA: MIT Press; 1997.
- [14] Jimenez MJ, Heras MR. Application of different dynamic analysis approaches to estimate the U and G values of building components. *EPIC AIVC* 2006;2: 427–34.
- [15] Jimenez MJ, Madsen H, Andersen KK. Identification of the main thermal characteristics of building components using MATLAB. *Build Environ* 2008;43: 170–80.
- [16] Okuyama H. Development of a thermal network model system parameter identification theory for original higher order state equation by using observations of lower order equation system with higher temporal differentiated state variables. *Proceedings of the Annual Meeting of the Society of Heating, Air-conditioning and Sanitary Engineers of Japan (SHASE-J)*, Nagoya, Japan; September 2004. F-9;pp. 437–40.
- [17] Miller SL, Leiserson K, Nazaroff WW. Nonlinear least squares minimization applied to tracer gas decay for determining airflow rates in a two-zone building. *Indoor Air* 1997;7:64–75.
- [18] Jiang Yi. State-Space method for the calculation of air-conditioning loads and the simulation of thermal Behavior of the room. *ASHRAE Tans* 1982;Pt2: 122–41.
- [19] Hong Tianzhen, Jiang Yi. A New Multizone model for the simulation of building thermal performance. *Build Environ* 1997;32(2):123–8.
- [20] Okuyama H, Kimura K. Temperature transition matrix and numerical time integration scheme by spectral resolution for analysis of building thermal network system. *Transaction of AIJ* July 1978;269:127–37.
- [21] Okuyama H. State space approach to building environmental analysis using thermal network Concepts. *Shimizu Tech Res Bull Mar.*1985;4:45–51.
- [22] Okuyama H. System identification theory of the thermal network model and an application for multi-chamber airflow measurement. *Build Environ* 1990; 25:349–63.
- [23] Bierman GJ. Factorization methods for discrete sequential estimation. Wal-tham, MA: Academic Press; 1977.

- [24] Okuyama H. Recent progress on the multi-chamber airflow measurement system. International Symposium on Room Air Convection and Ventilation Effectiveness (ISRACVE), Tokyo, Japan; July 1992. pp. 351–56.
- [25] Okuyama H, Onishi Y, Tanabe S, Kashiwara S. Statistical data analysis method for multi-zonal airflow measurement using multiple kinds of perfluorocarbon tracer gas. *Build Environ* 2009;44:546–57.
- [26] Okuyama H, Onishi Y. Uncertainty analysis and optimum concentration decay term for air exchange rate measurements: estimation methods for effective volume and infiltration rate. *Build Environ* March 2012;49:182–92. published online:19-OCT-2011.
- [27] Okuyama H. Thermal and airflow network simulation program NETS. Proceedings of the 6th International IBPSA Conference (Building Simulation '99), Kyoto, Japan; September 1999. pp. 1237–44.
- [28] Charles L. Lawson of the US Jet Propulsion Laboratory and Richard J. Hanson of Visual Numerics, Inc.. Solving least squares problems. Philadelphia: Society for Industrial and Applied Mathematics, ISBN 0-89871-356-0; 1974.
- [29] Okuyama H. Thermal network model with space inner surface long wave radiation exchange by Infinite Reflection. Proceeding of Architectural Institute Japan Annual Meeting; 2009. Presentation number of the paper 41010, pp. 19–20.