Development of Data-driven Models for Thermal Dynamic Analysis of Buildings

by

Zequn Wang

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Civil (Cross-disciplinary)

Department of Civil and Environmental Engineering
University of Alberta

© Zequn Wang, 2019

Abstract

Data-driven modelling has been widely applied in building operation optimization, energy management, ongoing commissioning, and so on. This thesis presents a comprehensive study of data-driven modelling for analysis of building thermal dynamics. First, three types of data-driven models, namely, transfer-function based models (TF models), resistor-capacitor based models (RC models), and artificial-intelligence based models (AI models) are critically reviewed, including their formulations, interpretability of physical meanings, and prediction accuracy. Fundamental concepts and common techniques for model training and selection are also presented. By applying the data-driven approach to a low-energy house using on-site monitored data, features of different data-driven models are further demonstrated. It is found that, in general, RC models are the most suitable for physical interpretation.

Then, a simple yet effective methodology is proposed to obtain reliable RC models for building thermal dynamic analysis. In this methodology, complex preliminary model structures are first created based on physical principles and then simplified by progressively removing non-identifiable parameters. Two important techniques are adopted in the simplification process: 1) a genetic algorithm is employed during model training to ensure the satisfactory fitting ability of large model structures; 2) asymptotic confidence intervals are calculated for parameter estimates and used to define parameter non-identifiability. The methodology is illustrated using a case study of the low energy house. This case study shows that the obtained RC model can predict room temperatures with satisfactory accuracy, and the estimated parameters are physically interpretable.

Finally, the estimated RC model parameters are related to building configurations, and the model is used to evaluate the design of the low energy house, with a focus on the adequacy and effectiveness of its thermal energy storage (TES) system. The influences of different parameter values on energy consumption and temperature fluctuations are compared. The findings show that the current TES system (i.e., the concrete wall and slabs) is designed with sufficient thickness and surface area. Decreasing its thickness or surface area will result in considerably more fluctuations of indoor air temperatures. Regarding energy consumption for space heating, varying the design would not result in significant improvement. The RC model is also used in evaluating other aspects of the thermal performance of the house, such as the overall thermal transmission of the envelop, the equivalent solar aperture, and the relative significance of different energy flow paths.

Preface

Chapter 2 of this thesis has been submitted for publication as Zequn Wang, Yuxiang Chen, "Data-Driven Modeling of Building Thermal Dynamics: Methodology and State of the Art" to *Building and Environment*. I was responsible for the literature review, data analysis, as well as the manuscript composition. Dr. Yuxiang Chen was the supervisory author and assisted with data collection and manuscript edits.

Chapter 3 of this thesis has been submitted for publication as Zequn Wang, Yuxiang Chen, Yong Li, "Development of RC Model for Thermal Dynamic Analysis of Buildings through Model Structure Simplification" to *Energy and Buildings*. I was responsible for the model development, data analysis, and manuscript composition. Dr. Yuxiang Chen was the supervisory author and assisted with data collection. Both Dr. Yuxiang Chen and Dr. Yong Li contributed to concept formation and manuscript edits.

Chapter 4 of this thesis will be submitted for publication and presentation as Zequn Wang, Yuxiang Chen, "Evaluation of Building Thermal Performance using Data-driven RC Models" to an academic conference. I was responsible for the design of evaluation methods, data analysis, and manuscript composition. Dr. Yuxiang Chen was the supervisory author and involved in concept formation and manuscript edits.

Acknowledgments

I would like to express my deepest and sincerest gratitude to my supervisor, Dr. Yuxiang Chen, for his consistent support and constructive advice in the entire research process. Without his insightful guidance, completion of this study would have been improbable.

I wish to acknowledge Dr. Yong Li. His valuable comments and suggestions were critical to the writing of this thesis.

With a special mention to Robert Pawluk, Jiacheng Yao, Regina Celi Dias Ferreira, Marko Spasojevic, Anni Wei, Maysoun Ismaiel, and Yue Shao, I genuinely appreciate the delightful moments we have shared in our office.

Many thanks to Dr. Andreas K. Athienitis and his research team for providing the data. The research work presented in the thesis is financially supported by the University of Alberta.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

RC	Resistor-capacitor
TF	Transfer function
AI	Artificial intelligence
ANN	Artificial neural network
SVM	Support vector machine
ARX	Autoregressive with exogenous input
ARMAX	Autoregressive moving average with exogenous input
BJ	Box-Jenkins
OE	Output error
SS	State space
RNN	Recurrent neural network
RBFNN	Radial basis function neural network
GRNN	General regression neural network
PEM	Prediction error method
MLE	Maximum likelihood estimation
MSE	Mean square error
RMSE	Root mean square error
MAE	Mean absolute error
RMSE	Root mean square error
MIMO	Multiple-input multiple-output

Nomenclature

$\boldsymbol{u}$	Vector of inputs
y, $\boldsymbol{y}$	Output, the vector of outputs
$\boldsymbol{\theta}$	Vector of parameters
$\boldsymbol{x}$	The vector of states (before discretization)
$\mathbb{x}$	The vector of states (after discretization)
$\widetilde{\boldsymbol{A}}$, $\widetilde{\boldsymbol{B}}$, $\widetilde{\boldsymbol{H}}$	Matrices in continuous-time state-space representation
$\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{H}$	Matrices in discrete-time state-space representation
$\boldsymbol{K}$	Optimal Kalman gain matrix
$\boldsymbol{e}$	Vector or errors
$\boldsymbol{\omega}$	Process noise
$\boldsymbol{v}$	Measurement noise
G, H	Transfer functions
$\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$, $\mathcal{D}$	Polynomial functions
z	Forward-shift operator
na, nb, nc, nd	Orders of polynomials A, B, C, D
nu, ny	Number of inputs, number of outputs
ϕ	Radiation flux (kW/m^2)
U	Thermal transmittance
τ	Time constant

λ	Poles or characteristic frequencies
r	The root of $\mathcal{A}(z)$ or eigenvalue of $\boldsymbol{A}$
m	Net input signal
f	Activation function
Fit	The goodness of fit
nl	Number of neurons within the hidden layer
w, $\boldsymbol{w}$	Weight, the vector of weights
b, $\boldsymbol{b}$	Bias, the vector of biases
φ	Nonlinear mapping
V	Objective function
L	Likelihood function
N	Number of samples
h	Prediction horizon
$\boldsymbol{\Sigma}$	Conditional covariance
$\boldsymbol{\psi}$	The gradient of $\boldsymbol{y}$ with respect to $\boldsymbol{\theta}$
κ	Adaptation gain
$\mathcal{Q}$	The function of κ, $\boldsymbol{\psi}$, and $\boldsymbol{e}$
ud, yd	Input delay, output delay
T	Temperature (℃)
C	Thermal capacitance ($kWh/$℃)
R	Thermal resistance (℃$/kW$)
F, p, α	Evaluating factors
Q	Heat input (kW)

$\hat{\boldsymbol{y}}$	The vector of simulated outputs
V	Objective function
$\hat{\boldsymbol{\theta}}$	The vector of estimated parameters
$\hat{\boldsymbol{x}}, \hat{\mathbb{x}}$	Vectors of estimated states
$\mathcal{N}$	Normal distribution
$\boldsymbol{P_\theta}$	Asymptotic covariance matrix for model parameters
$\boldsymbol{V_\theta}, \boldsymbol{V_{\theta\theta}}$	Gradient and Hessian of V with respect to $\boldsymbol{\theta}$
$\mathbb{E}$	Expectation
δ	Significance level
$t_{\delta/2}$	t-statistics given δ
$\hat{\boldsymbol{\sigma}}_{\boldsymbol{\theta}}$	The estimated vector of standard errors for $\hat{\boldsymbol{\theta}}$
μ, ϵ	Threshold values
Subscripts	
$\tilde{t}$, t	Time, the t^{th} time step
o	Outdoor air
s	Solar
h	Heating (only used with Q)
$elec$	Gross electricity
i	Indoor air
g	Ground
vcs	Ventilation concrete slab
hp	Heat pump

$0b$, $1f$, $2f$	Basement, first-floor, second-floor
m, mm, n	Associated with concrete floor or slab
e	Associated with Building envelop
Superscript	
$'$	Transpose operation
j	Counter

Chapter 1. Introduction

1.1. Background

Buildings represent the largest energy-consuming sector in the world, with over one-third of all final energy and half of global electricity consumed there [1]. If no action is taken to improve energy efficiency in the buildings sector, this consumption is expected to rise by 50% by the year 2050 [1]. In addition, almost half the energy consumed in buildings in developed countries is used by heating, ventilation, and air conditioning systems [2]. Given this trend and trait of energy demand, a great deal of research efforts has been devoted to enhancing the energy efficiency of buildings such as operation optimization, energy management and ongoing commissioning.

Among such practices, data analysis and modelling techniques have demonstrated critical importance [3]. For example, it is believed that process modelling and identification is the most time-consuming and challenging part of predictive system control [4, 5]. Among a variety of data analysis and modelling techniques employed in building applications, data-driven modelling of building thermal dynamics has received particular interest. Data-driven modelling has shown promise in reducing operational energy consumption, shifting and shaving peak demand, and performance monitoring. These are achieved through model-based control of space heating and cooling [6-8], fault detection of mechanical systems [9, 10], retrofit evaluation [11], etc.

Data-driven modelling of building thermal dynamics (i.e., the data-driven approach) consists of three phases (see Figure 1.1): modelling, training and selecting. In the modelling phase, a mathematical model with unknown parameters is formulated which can predict system outputs using measured inputs. For example, the model may predict indoor air temperature using measured weather information, heating or cooling supply, and occupant activities. The unknown parameters are then estimated in the training phase by tuning the predicted outputs to the measured outputs. Finally, in the selecting phase, the best model is chosen by schematically comparing the well-trained model candidates.

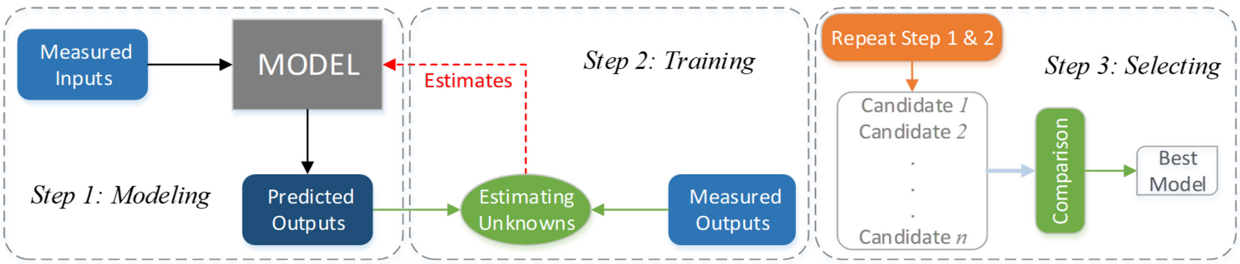


Figure 1.1 A general procedure of the data-driven approach

The data-driven approach is also known as an inverse approach, in contrast to the forward approach [12]. The forward approach uses engineering principles and prior-known design information (e.g., building geometry and thermal characteristics) to model building responses subjected to specified inputs. A severe drawback of the forward approach is that, in the rather complicated modelling process, unexpected interactions can occur between systems or between various modes of heat transfer [12]. On the contrary, the data-driven models are often simplified, easy to formulate and require less parameterization and computation time. The models tend to be highly adaptive to the actual performance data and give more accurate predictions of building responses. Moreover, multiple model candidates can be easily formulated and compared during model selection to derive a more representative model.

Data-driven dynamic modelling is relevant for prediction purpose (e.g., predicting room temperatures into the future) and/or explanatory purpose (e.g., inferring and verifying important thermal characteristics). Dynamic data-driven models can make hourly or sub-hourly predictions that are especially useful for predictive control of buildings with high thermal inertia. Dynamic data-driven models can also be used for describing the actual thermal behaviours and deriving as-built thermal properties of buildings. Moreover, by taking transient heat transfer into account, dynamic models can deliver better thermal characterization compared to steady-state models [13, 14]. As such, the dynamic data-driven models have broad applicability to buildings due to their prediction and explanatory abilities.

1.2. Objective & Scope

The primary objective of this research is to develop data-driven models that are simple and reliable, which can be used for the thermal dynamic analysis of buildings and are suitable for evaluation of the actual building thermal performance. Realization of this objective consists of addressing the followings:

i. Investigate different types of data-driven models used for the thermal dynamic analysis of buildings and compare their structure complexity, computation demand, as well as the capability of predicting and characterizing building thermal behaviours.
ii. Identify the type of data-driven models (e.g., RC models) that are the most suitable for explanatory purposes and have satisfactory prediction accuracy.
iii. Build up a methodology that can simplify complex model structures by removing unnecessary parameters to obtain a simpler and physically more reliable model.

iv. Apply the developed methodology to the modelling of a low energy house using on-site data and apply the obtained model to assessing the house's thermal performance and analyzing its design.

The thesis focuses on the physical interpretability of data-driven models. Starting from investigation of different model types, to buildup of model simplification methodology, to application of the obtained model on assessments of a low energy house's thermal performance, the entire process is to develop and benefit from a high-quality model that is readily physically interpretable. Another concentration of the research is to apply the developed data-driven models to evaluating the actual performance of existing designs and identify potential design improvements. The study has also incorporated various techniques of model training and selection along with their applicability in different situations.

1.3. Thesis Outline

The remainder of this thesis is organized as follows. Chapter 2 presents a literature review on the data-driven approach including variables for measuring, model formulation, training, and selection, and a small case study exemplifying the whole data-driven modelling process. Structure complexity, physical interpretability, and computational efficiency are discussed for different types of models. Chapter 3 proposes a methodology to simplify preliminary complex model structures by progressively removing non-identifiable parameters to obtain the most reliable model structures for explanatory purposes. The proposed methodology is illustrated through the case study of a low energy house for which a three-zone model is developed and justified. Chapter 4 demonstrates the physical interpretation of the simplified model structure through the design and modelling of a simple single-zone space. The simplified model structure is applied to evaluate the house's energy storage system and investigate the capability of internal thermal mass in reducing the fluctuations of indoor air temperatures. Chapter 5 concludes the thesis with a summary of the research, key contributions, and recommendations for future work.

Chapter 2. Literature Review

2.1. Introduction

The objective of this chapter is to conduct a literature review on the data-driven approach for the thermal dynamics of buildings. Section 2.2. "Categories of Data-driven Models and Their Fundamentals" demonstrates the formulations, constraints, and relations of three types of data-driven models. Section 2.3. "Training & Selection" summarizes input/output variables that are commonly employed for model training, and introduces parameter estimation methods, validation criteria, and model selection techniques. In Section 2.4. "Case study" a single-zone house is modelled using different models with on-site data. Section 2.5. "Discussion" explains the major differences between the reviewed models.

2.2. Categories of Data-driven Models and Their Fundamentals

Data-driven models are core components of the data-driven approach. In the literature, there are three main categories of data-driven models for building thermal dynamics, i.e., models based on resistor-capacitor networks (RC models), models based on discrete-time transfer functions (TF models), and models based on artificial intelligence techniques (AI models).

Building thermal dynamic process can be nonlinear with respect to the input variables, e.g., convective heat transfer on surfaces is nonlinearly dependent on the surface-air temperature difference. Besides, some thermal properties may vary over time due to changes in environment and operation (e.g., ventilation rate continually changes). Complex model structures can be developed to account for the nonlinearity and time-varying behaviours of building systems. However, most researchers adopt linear time-invariant versions of the data-driven models (except for AI models that are naturally nonlinear) for simplicity. Their results have shown that linear time-invariant models can characterize building thermal dynamics with promising performance [15-21].

This section will focus on describing linear time-invariant TF and RC models as well as time-invariant AI models. These models are either for prediction purposes (e.g., predicting room temperatures into the future) or for explanatory purposes (e.g., deriving specific thermal properties). They will be investigated regarding their structure formulations, prediction abilities, and physical interpretations.

2.2.1. RC models

The RC models capture building thermal dynamics by a network of temperature nodes, thermal resistors, and thermal capacitors. Resistances, capacitances, and other necessary parameters in the network are referred to as equivalent thermal parameters [22-24]. In other words, their estimates only enable the RC network to imitate building thermal dynamics and do not match precisely with the apparent quantities. Measured or prior-known node temperatures (e.g., zone temperatures) are often called system inputs or outputs, while temperatures linked to thermal capacitors are system states. The number of system states determines a model's order (or the number of ordinary differential equations as defined in the following context). Typically, models of lower orders are preferred since they require less parameterization and often have satisfactory prediction accuracy compared to larger ones [8, 13, 24-29].

Given an RC network and assuming one-dimensional heat transfer, thermal dynamics of each temperature node in the network are governed by the following ordinary differential equation (take the k^{th} node for example):

$$C_k \frac{dT_k}{dt} = \sum\nolimits_{\hbar} \frac{T_{\hbar} - T_k}{R_{\hbar,k}} + \sum\nolimits_j F_j Q_j \qquad (2.1)$$

where,

T_k represents the mperature of the k^{th} node;

C_k is the thermal capacity attached to the k^{th} node;

$T_{\hbar}$ represents the mperature of a neighbor of the k^{th} node.

$R_{\hbar,k}$ is the thermal resistance between the k^{th}node and its neighbor;

Q_j denotes the j^{th} heat input to the k^{th} node (e.g., solar radiation); and

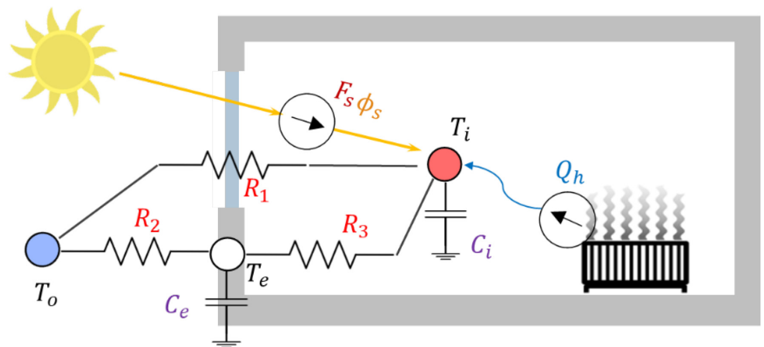
F_j is a factor (e.g., solar aperture) that evaluates the heat input Q_j.

Eqn. (2.1) is essentially a heat balance equation. On its right-hand side, $\sum_{\hbar} (T_{\hbar} - T_k)/R_{\hbar,k}$ describes the sum of heat fluxes from all the neighbor nodes and $\sum_j F_j Q_j$ for all the direct heat inputs. The left-hand side, $C_k\, dT_k/dt$, describes how the incoming energy is stored in the node. Such ordinary differential equations are set up for all the temperature nodes and rearranged into a state-space representation where model inputs, outputs, and parameters are clearly defined.

RC models are often used for zone air temperature prediction [15, 23, 24, 30, 31], as well as heating or cooling load prediction [7, 32-35]. Table 2.1 shows the formulation of a simple second-order RC model for temperature prediction.

In addition to temperature or load prediction, the RC models can also be used for inferring important building thermal characteristics [29, 36-38]. For example, if T_i, T_e, and T_o respectively, are zone air temperature, the average temperature of the building envelop, and outdoor air temperature, the overall thermal transmittance of the studied building can be calculated as $U = (R_2 + R_3)^{-1} + {R_1}^{-1}$; if ϕ_s is global solar radiation on the south façade, F_s can be interpreted as solar aperture and $F_s\phi_s$ becomes the transmitted effective solar gains.

Table 2.1 An example of RC model formulation

For a single-zone house	T_i: indoor air temperature T_e: the average temperature of building envelop T_o: outdoor air temperature Q_h: heating power ϕ_s: solar radiation on south façade
Model structure	
Ordinary differential equations	$C_e\frac{dT_e}{d\tilde{t}}=\frac{T_o-T_e}{R_2}+\frac{T_i-T_e}{R_3}$ $C_i\frac{dT_i}{d\tilde{t}}=\frac{T_o-T_i}{R_1}+\frac{T_e-T_i}{R_3}+F_s\phi_s+F_hQ_h\ (F_h=1)$
State-space representation in continuous-time	$\frac{d}{d\tilde{t}}\begin{bmatrix}T_e\\T_i\end{bmatrix}=\underbrace{\begin{bmatrix}-\frac{1}{C_eR_2}-\frac{1}{C_eR_3} & \frac{1}{C_eR_3}\\ \frac{1}{C_iR_3} & -\frac{1}{C_iR_1}-\frac{1}{C_iR_3}\end{bmatrix}}_{\tilde{\boldsymbol{A}}}\begin{bmatrix}T_e\\T_i\end{bmatrix}+\underbrace{\begin{bmatrix}\frac{1}{C_eR_2} & 0 & 0\\ \frac{1}{C_iR_1} & \frac{F_s}{C_i} & \frac{1}{C_i}\end{bmatrix}}_{\tilde{\boldsymbol{B}}}\begin{bmatrix}T_o\\ \phi_s\\ Q_h\end{bmatrix}$ $T_i=\underbrace{[0\quad 1]}_{\tilde{\boldsymbol{C}}}\begin{bmatrix}T_e\\T_i\end{bmatrix}+\underbrace{[0\quad 0\quad 0]}_{\tilde{\boldsymbol{D}}}\begin{bmatrix}T_o\\ \phi_s\\ Q_h\end{bmatrix}$
Variables	Input: $\boldsymbol{u}=[T_o\quad \phi_s\quad Q_h]'$; Output: $\boldsymbol{y}=T_i$; State: $\boldsymbol{x}=[T_e\quad T_i]'$; Parameter: $\boldsymbol{\theta}=[R_1\quad R_2\quad R_3\quad C_e\quad C_i\quad F_1\quad F_2]'$

The state-space representation obtained from differential equations is in continuous-time. For a multiple-input multiple-output (MIMO) system, the continuous-time state-space representation can be written as

$$\frac{d\boldsymbol{x}}{d\tilde{t}} = \widetilde{\boldsymbol{A}}\boldsymbol{x} + \widetilde{\boldsymbol{B}}\boldsymbol{u} \quad (2.2a)$$

$$\boldsymbol{y} = \widetilde{\boldsymbol{C}}\boldsymbol{x} + \widetilde{\boldsymbol{D}}\boldsymbol{u} \quad (2.2b)$$

where,

$\tilde{t}$ represents time;

$\boldsymbol{x}$, $\boldsymbol{u}$, and $\boldsymbol{y}$ are vectors of model states, inputs, and outputs, respectively; and

$\widetilde{\boldsymbol{A}}$, $\widetilde{\boldsymbol{B}}$, $\widetilde{\boldsymbol{C}}$, and $\widetilde{\boldsymbol{D}}$ are matrices determined from model parameters e.g., (R_1; R_2; R_3; C_e; C_i; F_s).

Eqn. (2.2) must be converted to a finite difference form in order to make use of the measured data for model training. The conversion of state-space representation is known as discretization and easily accessible in relevant studies [23, 39, 40]. Besides, environmental disturbances and measurement imperfections can result in noise-corrupted data. Thus, taking noise term into account, a discretized state-space representation is expressed by

$$\mathbb{x}_{t+1} = \boldsymbol{A}\mathbb{x}_t + \boldsymbol{B}\boldsymbol{u}_t + \boldsymbol{K}\boldsymbol{e}_t \quad (2.3a)$$

$$\boldsymbol{y}_t = \boldsymbol{C}\mathbb{x}_t + \boldsymbol{D}\boldsymbol{u}_t + \boldsymbol{e}_t \quad (2.3b)$$

where,

t represents the t^{th} time step;

$\mathbb{x}_t$ is a vector of states converted from $\boldsymbol{x}_t$;

$\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, and $\boldsymbol{D}$ are matrices calculated from $\widetilde{\boldsymbol{A}}$, $\widetilde{\boldsymbol{B}}$, $\widetilde{\boldsymbol{C}}$, and $\widetilde{\boldsymbol{D}}$;

$\boldsymbol{K}$ is the optimal Kalman gain matrix; and

$\boldsymbol{e}_t$ is a vector of stochastic processes, often assumed as Gaussian white noises [17, 21, 23, 41, 42].

Two types of interpolation are recommended for the discretization: piecewise constant interpolation and piecewise linear interpolation. They can be found in, e.g., Chapter 6 of [43]. For both interpolations. the matrix and state are converted as shown in Table 2.2.

Table 2.2 Discretization of continuous-time state-space representation

Piecewise linear interpolation	Piecewise constant interpolation
$\boldsymbol{A} = \exp(\widetilde{\boldsymbol{A}} T_s)$	$\boldsymbol{A} = \exp(\widetilde{\boldsymbol{A}} T_s)$
$\boldsymbol{B} = \boldsymbol{A}\boldsymbol{\Gamma}_2 + \boldsymbol{\Gamma}_1 - \boldsymbol{\Gamma}_2$	$\boldsymbol{B} = \widetilde{\boldsymbol{A}}^{-1}(\boldsymbol{A} - \boldsymbol{I})\widetilde{\boldsymbol{B}}$
$\boldsymbol{C} = \widetilde{\boldsymbol{C}}$	$\boldsymbol{C} = \widetilde{\boldsymbol{C}}$
$\boldsymbol{D} = \widetilde{\boldsymbol{C}}\boldsymbol{\Gamma}_2 + \widetilde{\boldsymbol{D}}$	$\boldsymbol{D} = \widetilde{\boldsymbol{D}}$
$\mathbb{x}_t = \boldsymbol{x}_t - \boldsymbol{\Gamma}_2 \boldsymbol{u}_t$	$\mathbb{x}_t = \boldsymbol{x}_t$
$\boldsymbol{\Gamma}_1 = \widetilde{\boldsymbol{A}}^{-1}(\boldsymbol{A} - \boldsymbol{I})\widetilde{\boldsymbol{B}}$	T_s: sampling interval
$\boldsymbol{\Gamma}_2 = \widetilde{\boldsymbol{A}}^{-1}\left({}^{\boldsymbol{\Gamma}_1}/_{T_s} - \widetilde{\boldsymbol{B}}\right)$	$\boldsymbol{I}$: identity matrix

Eqn. (2.3) is called the innovation form of state-space representation, compared to its stochastic form where two noise terms (process noise $\boldsymbol{\omega}$ and measurement noise $\boldsymbol{v}$) are adopted (i.e., the equations become $\mathbb{x}_{t+1} = \boldsymbol{A}\mathbb{x}_t + \boldsymbol{B}\boldsymbol{u}_t + \boldsymbol{\omega}_t$ and $\boldsymbol{y}_t = \boldsymbol{C}\mathbb{x}_t + \boldsymbol{D}\boldsymbol{u}_t + \boldsymbol{v}_t$). Since the Kalman gain matrix $\boldsymbol{K}$ can be derived from covariances of $\boldsymbol{\omega}$ and $\boldsymbol{v}$ through the Algebraic Riccati Equation [44], these two forms of state-space representation are mathematically equivalent [17, 45].

2.2.2. TF models

The outputs of physical systems (e.g., indoor air temperature in a building) depend on current and past inputs [46]. Thus, for a linear system, the current output can be related to the history of inputs by the convolution sum: $y_t = \sum_{j=-\infty}^{t} \theta_{t-j} u_j$ or $y_t = \left(\sum_{j=-\infty}^{t} \theta_{t-j} z^{-j}\right) u_t$, where θ's are constant parameters and z is a forward-shift operator: $u_{t-1} = z^{-1} u_t$. Here, $\sum_{j=-\infty}^{t} \theta_{t-j} z^{-j}$ can be regarded as a transfer function (TF) in discrete-time. Such TFs can be used to characterize the input-output relationships and thermal dynamics of buildings. In general, a linear time-invariant TF model can be written as [46]:

$$\boldsymbol{y}_t = G(\boldsymbol{\theta}, z)\boldsymbol{u}_t + H(\boldsymbol{\theta}, z)\boldsymbol{e}_t \tag{2.4}$$

where,

z is a forward-shift operator: $\boldsymbol{u}_{t+1} = z\boldsymbol{u}_t$ or $\boldsymbol{u}_t = z^{-1}\boldsymbol{u}_{t+1}$;

$G(\boldsymbol{\theta}, z)$ is the transfer function for inputs; and

$H(\boldsymbol{\theta}, z)$ is the transfer function for system noise.

$G(\boldsymbol{\theta}, z)$ and $H(\boldsymbol{\theta}, z)$ are often expressed in rational function forms, e.g., $G(\boldsymbol{\theta}, z) = \mathcal{B}(z)/\mathcal{A}(z)$ and $H(\boldsymbol{\theta}, z) = \mathcal{C}(z)/\mathcal{D}(z)$, where $\mathcal{A}(z)$, $\mathcal{B}(z)$, $\mathcal{C}(z)$, and $\mathcal{D}(z)$ are polynomial functions:

$$\mathcal{A}(z) = a_0 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_{na} z^{-na}$$

$$\mathcal{B}(z) = b_0 + b_1 z^{-1} + b_2 z^{-2} + \cdots + b_{nb} z^{-nb}$$

$$\mathcal{C}(z) = c_0 + c_1 z^{-1} + c_2 z^{-2} + \cdots + c_{nc} z^{-nc}$$

$$\mathcal{D}(z) = d_0 + d_1 z^{-1} + d_2 z^{-2} + \cdots + d_{nd} z^{-nd}$$

Thus, the vector of model parameters $\boldsymbol{\theta}$ is composed of a_j, b_j, c_j, and d_j, and the model order is defined by na, nb, nc, and nd. For a MIMO system with nu inputs and ny outputs, $G(\boldsymbol{\theta}, z)$ and $H(\boldsymbol{\theta}, z)$ are respectively, $ny \times nu$ and $ny \times ny$ matrices of transfer functions.

Depending on the use of polynomial functions, TF models can be categorized into Box-Jenkins (BJ) model, autoregressive moving average with exogenous input (ARMAX) model, autoregressive with exogenous input (ARX) model, output error (OE) model, etc. [47, 48]. The BJ models employ different polynomials for $\mathcal{A}(z)$, $\mathcal{B}(z)$, $\mathcal{C}(z)$, and $\mathcal{D}(z)$. When $\mathcal{D}(z) = \mathcal{A}(z)$, the BJ models become ARMAX models, and if further $\mathcal{C}(z) = 1$, they are simplified to ARX models (i.e., $\mathcal{A}(z)\boldsymbol{y}_t = \mathcal{B}(z)\boldsymbol{u}_t + \boldsymbol{e}_t$, and $a_0 = 1$). These models are different merely by the structure of $H(\boldsymbol{\theta}, z)$, namely, how the noise is modelled. An $H(\boldsymbol{\theta}, z)$ with more freedom (such as that in a BJ model) allows more disturbances beyond the white noise $\boldsymbol{e}_t$ to be described. Hence, the BJ models may outperform the ARMAX and ARX models in terms of prediction accuracy [47]. However, the ARMAX and ARX models are the most popular among researchers because they are simple and easily implementable in building applications [18, 49-56]. Furthermore, the deterministic structure of ARMAX and ARX models (i.e., $A(z)y_t = B(z)u_t$) has already been derived in the forward approach, known as comprehensive room transfer function [57]. Such models in nature, tend to align with the fundamental physical laws.

To be additional, RC and TF models are closely related through the state-space representation. By applying $\boldsymbol{x}_{t+1} = z\boldsymbol{x}_t$ to Eqn. (2.3a), we have $(z\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x}_t = \boldsymbol{B}\boldsymbol{u}_t + \boldsymbol{K}\boldsymbol{e}_t$ where $\boldsymbol{I}$ is an identity matrix. Substitute this to Eqn. (2.3b), we get

$$\boldsymbol{y}_t = [\boldsymbol{C}(z\boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{B} + \boldsymbol{D}]\boldsymbol{u}_t + [\boldsymbol{C}(z\boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{K} + \boldsymbol{I}]\boldsymbol{e}_t \qquad (2.5)$$

In Eqn. (2.5), $[\boldsymbol{C}(z\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{B}+\boldsymbol{D}]$ and $[\boldsymbol{C}(z\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{K}+\boldsymbol{I}]$ can be explained as transfer functions. When the matrices $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{D}$ and $\boldsymbol{K}$ are directly parameterized with constant entries instead of using equivalent thermal parameters, state-space representations should also be categorized as TF models (referred to as SS models). This is because an identified SS model can always be mapped to a unique transfer function model through Eqn. (2.5) [58], but it is practically impossible to derive all equivalent thermal parameters from the SS model's matrices. After being mapped to transfer function forms, the SS models have similar structures as ARX models and are also commonly used for characterizing building thermal dynamics [17, 21, 45, 59, 60]

Table 2.3 gives an example of TF model formulation using ARX structure. Compared to the RC model example in Table 2.1, this ARX model is of lower order but has a larger set of parameters to be estimated, and the parameters are not as physically interpretable as in the RC models. However, developing a TF model is more straightforward, i.e., there is no need for creating a thermal network or discretizing continuous-time differential equations. As such, the obtained TF models are mostly for prediction purposes: predicting room temperatures [60, 61], humidity level [18, 47, 62], and heating or cooling load [63-66]. Furthermore, TF models are pure statistical models. Being identified from data that contains disturbances, they may violate conservation of energy, exhibit resonant behaviour, or be unstable or non-casual [51]. Therefore, certain constraints or supervisory rules should be applied to the models.

The purpose of constraints is to obtain physically plausible and stable models. If the constraints are satisfied, a TF model can also be used for deriving important thermal characteristics of the building, such as heat loss coefficient through the building envelop, solar aperture, and time constants [17, 48, 67].

Table 2.3 An example of TF model (ARX) formulation

For a single-zone house	T_i: indoor air temperature T_o: outdoor air temperature Q_h: heating power ϕ_s: solar radiation on south façade
Model structure	ARX (na, nb) $nb = [nb_1 \quad nb_2 \quad nb_3]$
Transfer function form	$(1 + a_1 z^{-1} + \cdots + a_{na} z^{-na}) T_{i,t}$ $= \begin{bmatrix} b_{1,0} + b_{1,1} z^{-1} + \cdots + b_{1,nb_1} z^{-nb_1} \\ b_{2,0} + b_{2,1} z^{-1} + \cdots + b_{2,nb_2} z^{-nb_2} \\ b_{3,0} + b_{3,1} z^{-1} + \cdots + b_{3,nb_3} z^{-nb_3} \end{bmatrix} [T_{o,t} \quad \phi_{s,t} \quad Q_{h,t}] + e_t$
Polynomial form	e.g., $na = 1, nb = [2 \quad 2 \quad 1]$ $T_{i,t} = -a_1 T_{1,t-1} + b_{1,0} T_{o,t} + b_{1,1} T_{o,t-1} + b_{2,0} \phi_{s,t} + b_{2,1} \phi_{s,t-1}$ $+ b_{3,0} Q_{h,t} + e_t$
Variables	Input: $\boldsymbol{u} = [T_o \quad \phi_s \quad Q_h]'$; Output: $\boldsymbol{y} = T_i$; Parameter: $\boldsymbol{\theta} = [a_1 \quad b_{1,0} \quad b_{1,1} \quad b_{2,0} \quad b_{2,1} \quad b_{3,0}]'$

The first and the most important constraint is the steady-state constraint, i.e., the identified TF models must hold under steady-state conditions. Take an unoccupied single-zone building for example, the steady-state constraint is given by Eqn. (2.6) [51, 67, 68],

$$Q_h + F_s \phi_s = U(T_i - T_o) \tag{2.6}$$

where,

$F_s \phi_s$ is the effective solar heat gain (F_s is the solar aperture and ϕ_s is the global solar radiation);

Q_h is the supplied heating power;

T_i and T_o are respectively, the indoor and outdoor air temperatures; and

U is the overall thermal transmittance of the single-zone building.

Under steady-state conditions, all input and output variables are constant (i.e., $z = 1$). By substituting $z = 1$ to the ARX model in Table 2.3 (polynomial form), we have $(1 + a_1)T_i = (b_{1,0} + b_{1,1})T_o + (b_{2,0} + b_{2,1})\phi_s + b_{3,0}Q_h$. In order to align this with Eqn. (2.6), the ARX model must satisfy $b_{1,0} + b_{1,1} = 1 + a_1$ so that $U = (b_{1,0} + b_{1,1})/b_{3,0}$ or $U = (1 + a_1)/b_{3,0}$. The solar aperture can also be derived by $F_s = (b_{2,0} + b_{2,1})/b_{3,0}$.

In addition to the steady-state constraint, Armstrong et al. [51] proposed the pole constraint, i.e., for diffusion processes (which building thermal dynamics are simplified as), the characteristic frequencies or poles, λ_j, must be real and positive. This corresponds to saying that roots of $A(z)$ in Eqn. (2.4) or eigenvalues of matrix $\boldsymbol{A}$ in Eqn. (2.3a) are real and less than one. The pole constraint assures the model to be stable. Furthermore, Chen [69] and Chen et al. [70] proposed to supervise the parameter estimator with certain supervisory rules. Specifically, the rules require that the sum of any first coefficients with respect to each input and output, as well as the common ratio of response factors of output to any input, should follow the same features as in theoretically derived room transfer functions. The violation of these features will lead to a physically meaningless or unstable model.

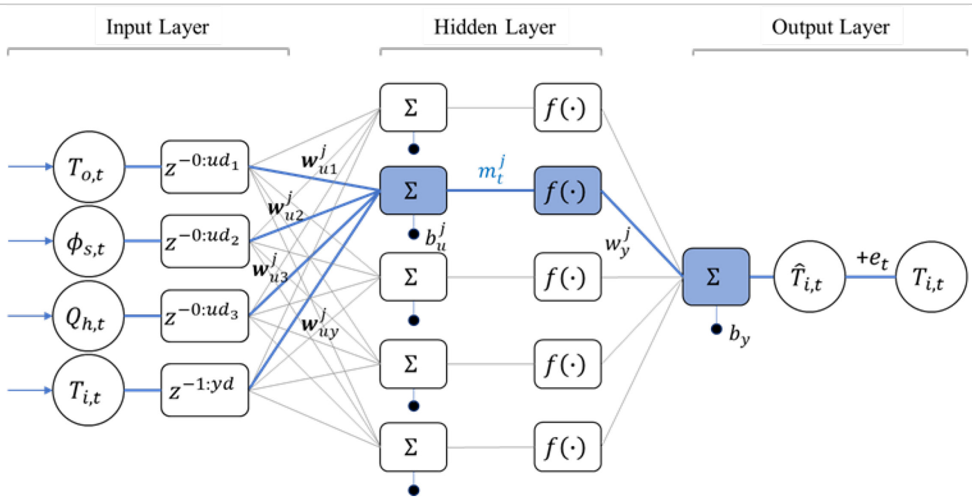
While applying the pole constraint, the time constants can be calculated by $\tau_j = 1/\lambda_j$ and $\lambda_j = \ln r_j$, where r_j is the j^{th} root of the polynomial $\mathcal{A}(z)$ or eigenvalue of the matrix $\boldsymbol{A}$ in SS models.

2.2.3. AI models

Artificial intelligence (AI) is “the science and engineering of making intelligent machines, especially intelligent computer programs” as defined by John McCarthy in 1956. Such intelligent machines (e.g., neural networks) have been extensively adopted for building energy use prediction since the 1990s. A representative example is “The Great Energy Predictor Shootout” competitions held by American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE), where neural networks provided the most accurate model of a building`s energy use [71, 72]. Among a variety of machine learning techniques, artificial neural networks (ANN) and support vector machines (SVM) (for regression) are the most commonly used for modelling of building thermal dynamics [73-75]. Other AI based modelling methods are also adopted such as artificial immune systems [76] and random forests [77], but they have received relatively less attention than ANN and SVM models.

Analogous to the human brain, ANN models map the input-output relationship by creating a large internal structure of artificial neurons. Typical ANN models for building thermal dynamics contain an input layer, one hidden layer, and an output layer [78-81]. Table 2.4 shows a simple AI model structured by a three-layered ANN, where one of the artificial neurons is highlighted. In each neuron, the input signals are weighted, summed and added by a bias. Then the net input signal m_t^j is assigned to an activation function $f(\cdot)$ and the activation is passed to the output function (often a linear function) to calculate the output. The example in Table 2.4 is a feedforward neural network, i.e., the input signals flow forward with no feedbacks. Its counterpart is a recurrent neural network (RNN), where the information can travel in loops from layer to layer [78, 82-84]. For example, in Table 2.4, if $T_{i,t}$ in the input layer is not measured but based on the model prediction $\hat{T}_{i,t}$ from the output layer, the feedforward neural network will turn into an RNN.

Table 2.4 An example of AI model (ANN) formulation

<table>
<tr><td>For a single-zone house</td><td>T_i: indoor air temperature
T_o: outdoor air temperature
Q_h: heating power
ϕ_s: solar radiation on south façade</td></tr>
<tr><td>Model Structure</td><td>
</td></tr>
<tr><td>Net input signal</td><td>For the j^{th} neuron:
$$m_t^j = z^{-1:yd}T_{i,t}\boldsymbol{w}_{uy}^j + z^{-0:ud_1}T_{o,t}\boldsymbol{w}_{u1}^j + z^{-0:ud_2}\phi_{s,t}\boldsymbol{w}_{u2}^j + z^{-0:ud_3}Q_{h,t}\boldsymbol{w}_{u3}^j + b_u^j$$</td></tr>
<tr><td>Activation & output functions</td><td>$$T_{i,t} = \sum\nolimits_j w_y^j f\left(m_t^j\right) + b_y + e_t$$
e.g., $f\left(m_t^j\right) = \tanh\left(m_t^j\right)$</td></tr>
<tr><td>Variables</td><td>Input: $\boldsymbol{u} = [T_o \quad \phi_s \quad Q_h]'$; Output: $\boldsymbol{y} = T_i$;
Parameter: $\boldsymbol{\theta} = \{\boldsymbol{w}_{u1}^{1:nl} \quad \boldsymbol{w}_{u2}^{1:nl} \quad \boldsymbol{w}_{u3}^{1:nl} \quad \boldsymbol{w}_{uy}^{1:nl} \quad b_u^{1:nl} \quad w_y^{1:nl} \quad b_y\}$;
e.g., $w_y^{1:nl} = [w_y^1, w_y^2, \ldots, w_y^{nl}]$</td></tr>
</table>

In Table 2.4,

nl is the number of neurons ($nl = 5$ in this example);

ud_1, ud_2, and ud_3 are input delays (e.g. $z^{-0:ud_1} = [1, z^{-1}, z^{-2}, \ldots, z^{-ud_1}]$);

yd is output delay (e.g., $z^{-1:yd} = [z^{-1}, z^{-2}, \ldots, z^{-yd}]$);

$\boldsymbol{w}_{uy}^{j}$, $\boldsymbol{w}_{u1}^{j}$, $\boldsymbol{w}_{u2}^{j}$, and $\boldsymbol{w}_{u3}^{j}$ are weight vectors for the j^{th} neuron in the input layer;

b_u^j is a bias for the j^{th} neuron in the input layer;

w_y^j is a weight for the j^{th} neuron in the output layer; and

b_y is a bias in the output layer.

The choice of activation function can affect the ANN performance. In general, the activation function introduces a degree of nonlinearity that is valuable for most ANNs [85], but there is no established rule defined for selecting activation functions to produce better network outputs [73]. Common activation functions include logistic-sigmoid functions [84, 86], hyperbolic tangent functions [78, 80, 81, 87], radial basis functions [88-91], etc. Particularly, ANNs that use radial basis functions are referred to as radial basis function neural networks (RBFNN). Such networks are said to have fast online learning ability, strong tolerance to noisy input data, good generalization, and easy design implementation [90]. A variation of RBFNN is the general regression neural network (GRNN), which also uses radial basis functions for activation and is even more suitable for online identification [92, 93].

The ANN model's performance is also impacted by the number of neurons within the hidden layer (nl). Too many neurons will cause the network to be overfitted and not generalize well beyond the training data. Too few neurons will weaken the network's ability to learn from the measurements. However, there is no strict rule for determining the right number of hidden neurons. Some researchers use empirical equations to calculate nl from the number of inputs nu and/or the number of outputs ny [84, 86, 94, 95], for example, $nl = 2nu + 1$. Other researchers consider nl as an indicator of model complexity and reduce it during model training or selecting without jeopardizing the prediction accuracy [11, 62, 80, 88]. Model training and selection will be further discussed in Section 2.3.

In addition to ANNs, the SVMs are also used increasingly in the modelling of building thermal dynamics [96-98]. The basic idea behind the SVMs is to map the input space into a high dimensional feature space through some nonlinear mapping, and then perform a linear regression in this feature space [99], namely,

$$\hat{\boldsymbol{y}} = \boldsymbol{w} \cdot \varphi(\boldsymbol{u}) + \boldsymbol{b} \tag{2.7}$$

where,

$\hat{\boldsymbol{y}}$ represents the predicted output vector;

$\varphi(\cdot)$ denotes the nonlinear mapping; and

$\boldsymbol{w}$ is a weight vector; and

$\boldsymbol{b}$ is a bias vector.

The nonlinear mapping in Eqn. (2.7) extract nonlinear features from inputs. For example, when $\boldsymbol{u} = [T_o, \phi_s, Q_h]$, the mapping can be $\varphi(\boldsymbol{u}) = \left[T_o, \phi_s, Q_h, T_o\phi_s, {Q_h}^2\right]$. Thus, a three-dimensional input space is mapped into a five-dimensional input space by including two nonlinear features: $T_o\phi_s$ and ${Q_h}^2$. In practice, the realization of $\varphi(\cdot)$ is often implicitly defined via kernels [96, 100]. By using kernels, all necessary computations can be performed directly in the input space $\boldsymbol{u}$ without having to compute the mapping $\varphi(\cdot)$ [101]. A detailed description of applying kernels in SVMs can be found in e.g., [102]. Furthermore, a major advantage of the SVM models is that they employ the structural risk minimization principle, which seeks to minimize an upper bound of the generalization error consisting of the sum of the training error and a confidence level [91, 97, 98, 101]. Owing to this feature, the SVM models can have fewer free parameters, and achieve better accuracy and generalization than conventional ANN and RBFNN models in e.g., predicting hourly cooling load of buildings [91].

For thermal dynamic problems, the AI models (ANN and SVM models in particular) can be regarded as nonlinear regressions, where the system inputs are regressors, and the outputs are dependent variables. In general, the AI models nonlinearly and implicitly relates outputs to inputs. They can also account for complex interactions between inputs through, e.g., an intertwined network in the input layer in ANN models or a nonlinear mapping of the input space in SVM models. Unlike RC or some TF models, the AI models are not physically interpretable, and cannot serve for any explanatory purposes. However, they tend to have higher prediction accuracy than the linear models [62, 79]. For that reason, they have also been extensively applied just like the RC or TF models, in model predictive control [89], fault detection [93, 95], retrofit evaluation [11], etc.

2.2.4. Enhanced models

The basic structures of RC, TF, and AI models can be combined or modified to create hybrid models. One possibility is to combine several different models by assigning linear weights to their outputs, so the combined model can take advantage of each model and gives higher prediction accuracy [100, 103, 104]. It is also possible to combine the model with techniques such as fuzzy logic [105, 106] and wavelet transform [106, 107] to improve the model's performance. Furthermore, a model can be modified with respect to weather conditions (e.g., outdoor air temperature) or system changes (e.g., opening of windows) to include a series of weather or system dependent models [63, 108]. This is equivalent to vary a model's structure with weather conditions and system changes. Its goal is to enhance the original model's ability to learn from the available data.

Although these hybrid models can exhibit better performances, they are not entirely new models. As such, this review only focuses on the basic structures of data-driven models.

2.3. Training & Selection

The data-driven models are formulated with unknown parameters. When training the models to learn buildings' thermal behaviours from the measured data, unknown parameters are estimated. After being trained, the models are evaluated by performing residual analysis or testing their generalization on new datasets. Then, the evaluation results are used as an indicator for comparing different models in the selecting phase. This section will focus on commonly adopted methods for training and selection of data-driven models along with typical input and output variables for measuring in building thermal dynamic studies.

2.3.1. Input/output variables

Inputs and outputs to be used for data-driven models are measured as a time series. This time series is usually sampled at a fixed sampling interval over a long duration (varies from several weeks to several years depending on the specific model and the quality of the data). The outputs should be an easily observable response, and the inputs should have significant influences on the system. Generally, input/output variables required for a thermal zone fall into the following types: zone air temperature, ambient air temperature, solar radiation, heating or cooling power supply, and internal gains. These variables are discussed in Table 2.5.

Table 2.5 Types of input/output variables for a thermal zone

Input/output variables	**Discussion**
Zone air temperature	Zone air temperature (or indoor air temperature) can be regarded as either input or output. When it is input, heating or cooling load is often output and vice versa. Some researchers measure temperatures from multiple locations in the zone and take average [17, 30, 32].
Ambient air temperature	Ambient air temperature is the outdoor air temperature in most cases. It impacts zone air temperatures, and heating/cooling patterns by heat transfer through building envelop, natural ventilation, fresh air intake, etc.
Solar radiation	This variable often comes available as global solar radiation. Some researchers directly use global horizontal or vertical radiation as input [29, 61, 109]. Other researchers employ mathematical methods to split the global radiation into direct normal and diffuse components to account for their distinctive solar effects [64, 66]. It is also possible to combine solar radiation and outdoor air temperature to a single input, i.e., sol-air temperature [34, 51, 52].
Heating or cooling power supply	Heating or cooling power supply, in general, is derived from other measured variables depending on the type of heating or cooling system. For examples, in a forced-air heating system, the heating power can be approximated using measured flow rate and temperature of the supply air from air handling units [7, 47]; if an electric heater provides space heating,

	the heating power can be approximated as the heater's electricity demand [23, 42]; if radiators are the heating source, they emit heat to the surroundings through both convection and radiation, which should be considered separately [30, 41].
Internal gains	Internal gains consist of heat gains from occupants, lighting, appliances, equipment, etc. For residential houses, some researchers use constant periodic values to approximate internal heat gains instead of measuring them [18, 64, 66]. For commercial buildings, many researchers relate internal gains to gross electricity demand excluding that for heating or cooling [17, 25, 109]. This is because occupants' behaviours can be highly correlated to the equipment (e.g., lights, printers, and computers) being used [110]. Besides, internal gains are often split into convective and radiative parts for separate consideration [32, 34].

Other variables to be measured include ground temperature, wind speed, wind direction, relative humidity, mechanical ventilation, etc. Since ground temperature varies negligibly compared to other inputs, it is often assumed to be constant [15, 17]. Wind speed and direction can change the conductive or convective heat transfer coefficient associated with building envelop over time [29, 64, 66]. Outdoor relative humidity can also influence zone air temperatures and heating/cooling loads by humification or dehumidification process. However, it is indicated that wind effects and humidity level are not as relevant as the ambient temperature and solar radiation [18, 80, 88]. As for the mechanical ventilation, the fresh air intake can be considered as heat loss to (or heat gain from) outdoor environment which can be integrated as part of the heating/cooling power supply [29]. Usually, the mechanical ventilation rate should be measured.

The measured variables are prepared for model training. It is generally necessary to acquire data with significant and persistent variations in order to train models that can provide accurate predictions. Furthermore, reliable parameter estimates can be obtained only when the measurements contain sufficient magnitude variance among every input and output [111]. If a system is not significantly excited, the data will not be as dynamically informative as needed for robust parameter estimation, e.g., certain parameters could be non-identifiable (cannot be uniquely identified) [112]. Therefore, ensuring the quality of on-site measurements is as equally important as developing a good model structure.

2.3.2. Model training

Model training is an optimization process that estimates the unknown parameters by minimizing the value of an objective function. If the parameter estimate is denoted by $\widehat{\boldsymbol{\theta}}$, then

$$\widehat{\boldsymbol{\theta}} = \min_{\boldsymbol{\theta}} V(\boldsymbol{\theta}) \tag{2.8}$$

$V(\boldsymbol{\theta})$ in Eqn. (2.8) is the parameter-dependent objective function (or loss function) often defined by prediction error method (PEM) or maximum likelihood estimation method (MLE).

Prediction error method (PEM)

In PEM, the loss function for minimization is a function of prediction errors (usually defined in the quadratic form) [113], for example,

$$V(\boldsymbol{\theta}) = \frac{1}{N}\sum_{t=1}^{N} \boldsymbol{e}_t(\boldsymbol{\theta})' \boldsymbol{e}_t(\boldsymbol{\theta}) \tag{2.9}$$

where,

N is the number of samples in training dataset;

$\boldsymbol{e}_t(\boldsymbol{\theta})$ is the prediction error at time t: $\boldsymbol{e}_t(\boldsymbol{\theta}) = \boldsymbol{y}_t - \widehat{\boldsymbol{y}}_{t|t-h}(\boldsymbol{\theta})$;

h is the prediction horizon;

$\widehat{\boldsymbol{y}}_{t|t-h}(\boldsymbol{\theta})$ is the predicted output at t given $\boldsymbol{\theta}$ and $\boldsymbol{Y}_{t-h}$; and

$\boldsymbol{Y}_{t-h}$ contains the measured outputs up to $t-h$: $\boldsymbol{Y}_{t-h} = [\boldsymbol{y}_{t-h}, \boldsymbol{y}_{t-h-1}, \dots, \boldsymbol{y}_1]$.

When $h = 1$, the error is referred to as one-step ahead prediction error, and the model parameters that give the smallest loss function value will give the best one-step ahead prediction performance. When $h > 1$, the error becomes multi-step ahead prediction error. A loss function with $h > 1$ is especially suitable for model predictive controllers since they often require models that can provide good predictions over a finite-time horizon [114-118]. When $h = \infty$, the error is known as simulation error. Using simulation errors for the loss function corresponds to, e.g., $H(\boldsymbol{\theta}, z) = 1$ for TF models in Eqn. (2.4). Such TF models are so-called output error (OE) models whose identification has been a long-standing topic studied in various respects [119-121]. By taking $H(\boldsymbol{\theta}, z) = 1$, the OE models focus on the dynamics of inputs and not the disturbance properties of noise. The prediction horizon influences the estimation results. Choice of h should be based on the purpose: whether the model is for short-term or long-term prediction.

Maximum likelihood estimation (MLE)

The MLE method estimates the parameters that make the observations (i.e. the measured outputs) most likely to be within the predicted outputs. In other words, the joint probability density of all the observations - the likelihood function - should be maximized. Maximizing the likelihood function is to minimize the loss function:

$$V(\boldsymbol{\theta}) = -\log L(\boldsymbol{\theta};\ \boldsymbol{Y}_N) \qquad (2.10)$$

where,

$L(\cdot)$ is the likelihood function: $L(\boldsymbol{\theta};\ \boldsymbol{Y}_N) = \prod_{t=1}^{N} p(\boldsymbol{y}_t | \boldsymbol{Y}_{t-h}, \boldsymbol{\theta})$;

$\boldsymbol{Y}_t$ contains the observations up to time t: $\boldsymbol{Y}_t = [\boldsymbol{y}_t, \boldsymbol{y}_{t-1}, \dots, \boldsymbol{y}_1]$; and

$p(\cdot \mid \cdot)$ is the conditional density function, often assumed to be Gaussian.

The Gaussian densities (assume $h = 1$) are determined by the conditional mean $\widehat{\boldsymbol{y}}_{t|t-1}$ and the conditional covariance $\boldsymbol{\Sigma}_{t|t-1}$, i.e.,

$$p(\boldsymbol{y}_t | \boldsymbol{Y}_{t-1}, \boldsymbol{\theta}) = \exp\left[-\frac{1}{2}\left(\boldsymbol{y}_t - \widehat{\boldsymbol{y}}_{t|t-1}\right)' \boldsymbol{\Sigma}_{t|t-1}\left(\boldsymbol{y}_t - \widehat{\boldsymbol{y}}_{t|t-1}\right)\right] \Big/ \sqrt{(2\pi)^{ny} \left|\boldsymbol{\Sigma}_{t|t-1}\right|}$$

In the stochastic state-space representation, $\widehat{\boldsymbol{y}}_{t|t-1}$ and $\boldsymbol{\Sigma}_{t|t-1}$ can be calculated recursively by using a Kalman filter [23, 122]. If the conditional covariance is assumed to be a time -invariant constant (i.e., $\boldsymbol{\Sigma}_{t|t-1} = \boldsymbol{\Sigma}$), it can be found by minimizing the loss function with respect to $\boldsymbol{\Sigma}$ independently from the other parameters [113, 123]. Under the special conditions of Gaussian densities and $\boldsymbol{\Sigma}$, the MLE method is equivalent to the one-step ahead PEM in its least-squares form [123]. As such, a major advantage of the PEM is that no probabilistic assumptions must be made.

Optimization algorithms

Either PEM or MLE defines a loss function to be minimized. For specific models, the value of the loss function can be minimized by linear regression techniques, e.g., linear least squares for the ARX models [124, 125] and subspace identification for the state space (SS) models [126, 127]. When linear regression techniques are allowed, a global minimum is always guaranteed.

For other models in general, this minimization process can be realized by iterative search algorithms such as Gauss-Newton algorithm [19, 25] and Levenberg-Marquardt algorithm [17, 32, 62, 88, 89, 95]. Typically, these algorithms evaluate gradients or Hessians for a searching direction that leads to the optimal solution. Specifically, backpropagation is a method used in ANN models for calculating gradients [11, 62, 78, 84]. However, the iterative algorithms require a proper initial guess for the parameter estimation. Without a good starting point, the searching guided by gradients or Hessians may lead to a local minimum. Therefore, some authors have proposed to employ global optimization routines, such as the genetic algorithm [34], the multi-start searching [32], the modal trimming method [81], the differential evolution algorithm [98], etc., to approach the global optimal.

Online algorithms

The iterative or linear algorithms are off-line. Namely, the parameters are estimated with fixed datasets. On-line algorithms, on the other hand, update the estimates as new data become available for the next time step [26, 54, 55]. Thus, estimation using on-line algorithms is also referred to as real-time estimation [55]. It can be expressed by

$$\widehat{\boldsymbol{\theta}}_{t+1} = \widehat{\boldsymbol{\theta}}_t + Q\left(\kappa_t, \boldsymbol{\psi}_t, \hat{\boldsymbol{e}}_{t+1|t}\right) \tag{2.11}$$

where,

$\widehat{\boldsymbol{\theta}}_t$ is the estimated parameter at time t;

$Q(\cdot)$ is a correction term determined by κ_t, $\boldsymbol{\psi}_t$, and $\hat{\boldsymbol{e}}_{t+1|t}$;

$\boldsymbol{\psi}_t$ is the gradient $\partial\hat{\boldsymbol{y}}_t/\partial\widehat{\boldsymbol{\theta}}_t$;

$\hat{\boldsymbol{e}}_{t+1|t}$ is the one-step ahead prediction error $\boldsymbol{y}_{t+1} - \boldsymbol{\psi}_t\widehat{\boldsymbol{\theta}}_t$; and

κ_t is an adaptation gain.

The adaption gain can be interpreted by forgetting factor or Kalman filter: the forgetting factor introduces increasingly weaker weighting on the old data while the Kalman filter reflects the evolution of the covariance of the parameter error [128]. Thus, the real-time estimation is to recursively minimize the same loss function as in the off-line estimation but modified by the forgetting factor or the parameter covariance. Here, the definition of "real-time" should be distinguished from “adaptive”. For example, accumulative training or sliding window training [86, 88] adapts to new data but uses off-line (non-recursive) algorithms.

Depending on the underlying loss functions, Eqn. (2.11) represents a family of recursive algorithms: the recursive least squares, the recursive instrumental variables, the recursive maximum likelihood, etc. [129]. It is shown that these different algorithms have substantially the same structure and can be unified to a general description known as the recursive prediction error algorithm [130]. It is also shown that this on-line algorithm has the same convergence properties as its off-line counterparts [130].

2.3.3. Model selection

Models trained by off-line algorithms are candidates for model selection, a systematic routine through which these candidates are validated, tested, and compared. The purpose of model selection is to find the most suitable model that achieves favourable prediction accuracy with the least possible complexity. In other words, the selected model should be able to characterize the principal thermal dynamics using fewest input variables and model parameters (number of the model parameters is highly dependent on the model order, e.g., the number of ordinary differential equations in RC models, the number of time lags in ARX models or the number of hidden neurons in ANN models).

Forward/backward selection

Two types of model selection are shown in Figure 2.1. One is the forward selection which starts from the simplest model and progressively extends the model until the model’s performance can no longer be improved in a significant sense [17, 20, 42, 66, 131]. The other is the backward selection that consecutively creates sub-models from a larger model until a decrease of the model’s performance allows for no further model reduction [87, 104, 132]. To perform either selection process, the candidate models' performances must be evaluated by validation or/and testing (here, validation refers to any analyses conducted on the dataset used for training while testing refers to any analyses conducted on an independent dataset from training).

(a) forward selection

(b) backward selection

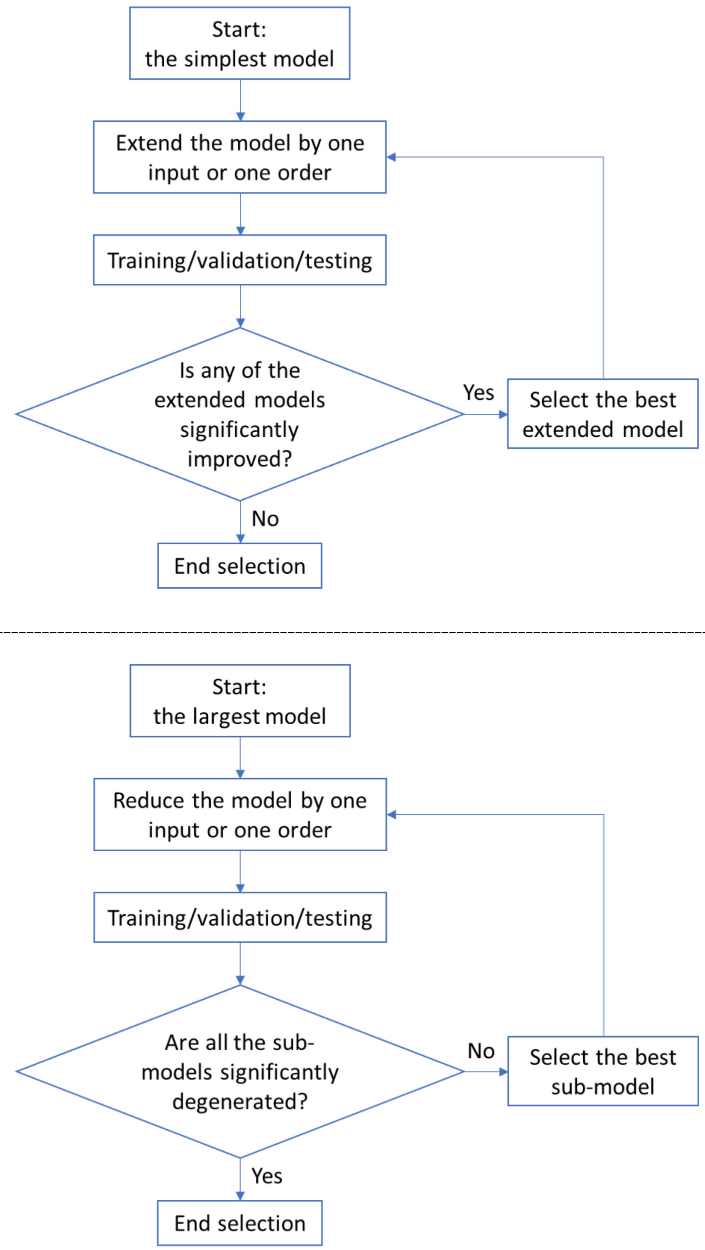


Figure 2.1 Model selection procedures: (a) forward, and (b) backward

Model evaluation/comparison

The trained models can be evaluated by quality criteria including mean square error (MSE), root mean square error (RMSE), mean absolute error (MAE), the goodness of fit (*Fit*) (see Eqn. (2.13)), the coefficient of determination (R^2-value), the coefficient of variation, mean bias error and so on [20, 41, 47, 61, 62, 80, 87, 104, 131, 132]. Choosing which criterion to use depends on how the residuals (i.e., after-training errors) should be penalized and interpreted. For examples, the MSE takes the average of the squared residuals and emphasizes larger errors; the MAE takes the average of absolute residuals and equally treats errors of different magnitude; the R^2-value and the *Fit* are expressed in percentages and suitable to assess how much variance of the output variable is explained by the model. These quality criteria can be used for either validating a model's training performance or testing its generalization (or reproductivity) on new observations. Essentially, they measure a model's success at fitting the available data.

Another category of methods for evaluating the trained models is to perform residual analyses based on autocorrelation function, cumulative periodogram, and cross-correlation function [17, 20, 29, 42, 66, 131]. The autocorrelation function and the cumulative periodogram are used to check for whiteness of the residuals. The cross-correlation function is used to check for independence of the residuals and the input variables. Given a pre-specified level of significance, if the hypothesis of whiteness or independence is rejected, there could be significant dynamics remaining unexplained in the residuals, which indicates the need for a larger model and more input variables. Such statistical analyses are mostly conducted on the training dataset.

Additionally, the likelihood ratio test is a useful tool for comparing two nested models, i.e., one model is the sub-model of the other [20, 42]. If the likelihood of the sub-model is L_0 and the likelihood of the larger model is L, the test statistic can be expressed by $-2\log L_0/L$, which converges to a χ^2 distribution as the number of samples goes to infinity. For large values of the test statistic, it can be concluded that reducing the larger model can cause a significant decrease in the model performance.

Frequency response analysis can also be used for model comparison [24, 133]. The basic idea is that for two models of different orders, the one of lower order can be used if their frequency responses resemble each other (especially for frequencies corresponding to large amplitudes), leading to statistically negligible difference between their outputs.

Model pruning

Finally, excessive model comparisons can be caused when there are a lot of possibilities of reducing or extending a model. This is true for complex models (e.g., ANN models) with large sets of input variables and model parameters. To make the selection more efficient, techniques like model pruning [62, 80, 87] and recursive feature elimination [104, 132] can be integrated into the backward selection procedure. Such techniques identify unnecessary parameters or unimportant input variables right after model training. Then the unnecessary parameters or unimportant inputs are removed until a decrease of the prediction accuracy is no longer tolerated. Model pruning or recursive feature elimination can effectively guide model reduction. Moreover, model pruning is a solution to the potential overfitting problem of the ANN models.

In practice, model training and model selection are often realized on various toolboxes such as System Identification Toolbox™ [134], Neural Network Toolbox™ [135], CTSM-R [136], CAPTAIN [137], CONTSID [138], etc. These toolboxes have provided a convenient platform for data-driven modelling.

2.4. Case study

This section illustrates the data-driven approach through a case study of a single-detached low-energy house. The house has large glazing areas facing south and a significant amount of thermal mass from concrete floor/slab/walls. Its wood-frame building envelop is well-insulated and air-tight. A geothermal heat pump mainly provides space heating through forced hot air.

2.4.1. Measured data

The house is simplified as one thermal zone. Relevant input and output variables are summarized in Table 2.6. Outdoor air temperature (T_o) is the driving input. Global radiation on the south façade (ϕ_s) and gross electricity demand (Q_{elec}) are used to approximate effective solar heat gains and internal heat gains, respectively [109]. Heating power provided by the geothermal heat pump (Q_h) is calculated based on measured air flowrate and temperature difference between supply air and return air. Influence of the ground temperature is neglected since there is a 51 mm (RSI 1.8) insulation between the ground and the slab. The desired model output is indoor air temperature (T_i) which is an average of room temperatures weighted by their floor areas.

Table 2.6 Summary of input and output variables

Variables	**Unit**	**Description**
T_o	℃	Outdoor air temperature
ϕ_s	kW/m^2	Global irradiation on the south façade
Q_h	kW	Heating power provided by the geothermal heat pump
Q_{elec}	kW	Gross electricity demand
T_i	℃	Indoor air temperature (average)

This house was monitored for two months (January and February in 2011). Raw data obtained during monitoring was preprocessed (formatting, synchronization, deleting outliers, etc.) into measurements that can be used for model training. All measurements are sampled every 0.5 hour. The measurements are then divided into one training dataset and one testing dataset (see Figure A.1). The training dataset includes 30-day data points from January 2011 (number of samples N = 1440) while the testing dataset consists of 29-day data points from February 2011.

2.4.2. Training and testing criteria

Three data-driven models (i.e., RC, ARX, and ANN models) are developed. The PEM is adopted for model training, where the objective function is defined based on one-step ahead prediction errors:

$$V(\boldsymbol{\theta}) = \frac{1}{N}\sum_{t=1}^{N}\left[e_{t|t-1}(\boldsymbol{\theta})\right]^2 \qquad (2.12)$$

Minimizing $V(\boldsymbol{\theta})$ gives parameter estimate $\widehat{\boldsymbol{\theta}}$. For RC and AI models, the minimization of $V(\boldsymbol{\theta})$ is realized by the Levenberg–Marquardt algorithm. For ARX models, the parameters are estimated by linear least squares.

The obtained models are often adopted for one-day ahead forecasting, especially in applications like model predictive control. To examine the models' forecasting ability, each model is demanded to forecast the indoor air temperature 48 steps into the future (i.e., 24-hour ahead) given the previous one-week data. Then, the forecasted temperatures are compared with the measured temperatures using the following criteria:

$$Fit_j = \left(1 - \frac{\sqrt{\sum_{t=1}^{48}\left(y_{j,t} - \hat{y}_{j,t}\right)^2}}{\sqrt{\sum_{t=1}^{48}\left(y_{j,t} - \bar{y}_j\right)^2}}\right) \cdot 100\% \qquad (2.13)$$

where,

Fit_j is the goodness of fit for the j^{th} day;

$y_{j,t}$ is the measured output at time step t in j^{th} day;

$\hat{y}_{j,t}$ is the forecasted output at time step t in j^{th} day; and

$\bar{y}_j$ is the average of $y_{j,t}$, i.e., $\bar{y}_j = \frac{1}{48}\sum_{t=1}^{48} y_{j,t}$.

The examination by Eqn. (2.13) is performed on the testing data, which yields $Fit's$ for 29 days. The goodness of fit in percentage informs how closer the data are to the fitted curve compared to a straight line (i.e., $\bar{y}_j$). The larger the Fit, the more accurately a model can fit the measurements. Here, instead of displaying all 29 values of Fit, its average and standard deviation are used:

$$\overline{Fit} = \frac{1}{29}\sum_{j=1}^{29} Fit_j \tag{2.14}$$

and

$$Fit\_sd = \sqrt{\frac{1}{28}\sum_{j=1}^{29}\left(Fit_j - \overline{Fit}\right)} \tag{2.15}$$

A good model should give a high $\overline{Fit}$ and a low Fit_sd as much as possible. These two criteria will be used for the following evaluation and selection of model structures.

2.4.3. Model development

The full RC model considers four inputs ($T_o \phi_s Q_h Q_{elec}$) and three states ($T_i T_m T_e$). Its structure can be found in Figure 2.2 (e). T_m and T_e represent temperatures of internal concrete mass and building envelop, respectively. C_m and C_e represent their thermal capacitances. C_i denotes thermal capacitance of the indoor air, furniture, etc. The outdoor air temperature (T_o) affects the indoor air temperature (T_i) through a fast response path (R_1) and a slow response path

(R_2-C_e-R_3). The fast response path captures the thermal impact through windows and ventilation while the slow response path is mainly to characterize the transient conduction through thick walls, roof, and ceilings. F_s and F_c are respectively, the solar gain factor (i.e., solar aperture) and internal gain factor that evaluate effective solar gains (from ϕ_s) and effective internal gains (from Q_{elec}) to the indoor.

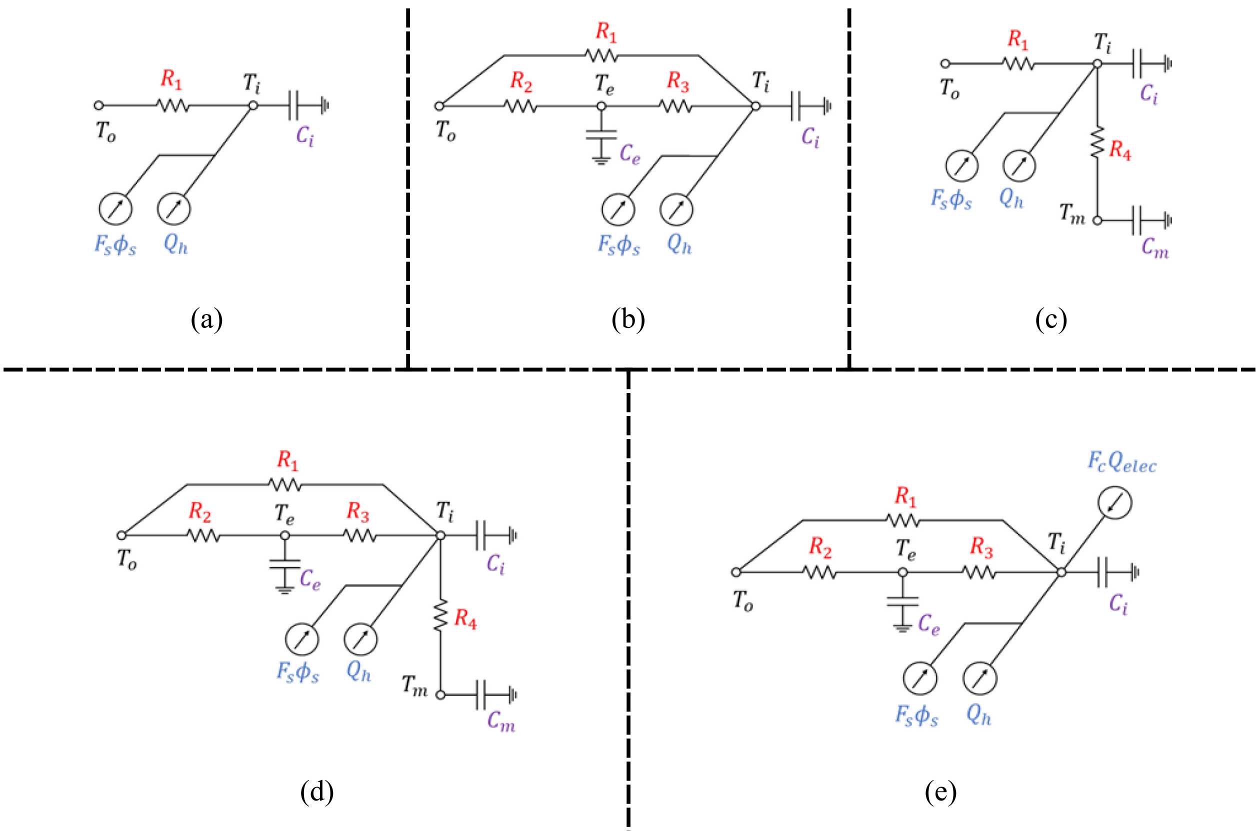


Figure 2.2 RC model structures: (a) $T_o\phi_sQ_h$-T_i*; (b)* $T_o\phi_sQ_h$-T_iT_e*; (c)* $T_o\phi_sQ_h$-T_iT_m*; (d)* $T_o\phi_sQ_h$-$T_iT_mT_e$*; (e)* $T_o\phi_sQ_hQ_{elec}$-T_iT_e

To facilitate explanation, an RC model is labelled in the format of RC (inputs-states). The most suitable model structure is selected by the forward selection technique starting from RC ($T_o\phi_sQ_h$-T_i). T_o, ϕ_s, and Q_h are considered initial inputs and Q_{elec} is tested as a model extension.

A list of model structures is given in Figure 2.2. The selection process is shown in Table 2.7 where the selected model structure is highlighted. RC ($T_o\phi_sQ_h$-$T_iT_mT_e$) is not selected though it has slightly better $\overline{Fit}$ and Fit_sd than RC ($T_o\phi_sQ_h$-T_iT_m). This is because that the focus of RC models is to obtain physically interpretable parameters. For RC ($T_o\phi_sQ_h$-$T_iT_mT_e$), some parameters are estimated to have unreasonable values and considerable uncertainties, hence are not interpretable.

Table 2.7 indicates that extending the RC model by Q_{elec} makes a negligible improvement to either its training or testing performance. Thus, Q_{elec} is no longer considered for the following development of ARX and ANN models.

Table 2.7 Selection of a suitable RC model

RC model structures		Training	Testing	
Inputs	States	V	$\overline{Fit}$ %	Fit_sd %
$T_o\phi_sQ_h$	T_i	0.3186	-29.64	59.62
	T_iT_e	0.0098	72.60	12.12
	T_iT_m	0.0098	68.90	13.24
	$T_iT_mT_e$	0.0099	73.57	10.82
$T_o\phi_sQ_hQ_{elec}$	T_iT_e	0.0100	72.18	12.32

An ARX model is labelled in the format of ARX (na, nb) with $nb = [nb_1 \quad nb_2 \quad nb_3]$. na and nb are output and input delays, respectively. The backward selection routine is employed for selecting the most suitable model structure. The selected structure ARX (10, [1 10 10]), as highlighted in Table 2.8, has $\overline{Fit}$ and Fit_sd less than 1% different from the largest model structure. Further reducing its order will cause an apparent decrease in the testing performance.

Table 2.8 Selection of a suitable ARX model

ARX model structures			Training	Testing	
Inputs	na	nb	V	$\overline{Fit}$ %	Fit_sd %
$T_o \phi_s Q_h$	12	[12 12 12]	0.0044	76.13	12.42
		[1 10 10]	0.0045	76.25	13.03
		[1 10 8]	0.0046	75.52	14.08
	10	[1 10 10]	0.0045	76.25	13.02
	8	[1 10 10]	0.0046	75.64	14.15

A feedforward ANN model is labelled in the format of ANN (nl, 0: ud, 1: yd). Here, all input delays are set the same (i.e., $ud_1 = ud_2 = ud_3 = ud$) which leaves the training algorithm to adjust relative importance (i.e., weight) of each delay. Hyperbolic tangent function (i.e., tanh) is adopted as the activation function. Before training, the weights and biases are initialized with random small values. Since a lot of weights and biases are being used, there may exist multiple local minima. To overcome this problem, each model structure is initialized and trained for 20 times from which the one with the best testing performance is screened out. Moreover, to avoid overfitting, early stopping is adopted and 20% of the training data is used for cross-validation. Based on the above settings, model selection is performed forwardly starting from ANN (4, 0:2, 1:2). The selected model structure is ANN (4, 0:2, 1:3) which gives the highest $\overline{Fit}$ and the lowest Fit_sd (see Table 2.9).

Table 2.9 Selection of a suitable ANN model

ANN model structures				Training	Testing	
Inputs	nl	$0{:}\,ud$	$1{:}\,yd$	V	$\overline{Fit}$ %	Fit_sd %
$T_o\phi_sQ_h$	4	0:2	1:2	0.0024	77.85	11.38
		0:3	1:2	0.0016	77.56	10.98
		0:2	1:3	0.0023	78.36	7.00
		0:2	1:4	0.0017	77.71	10.56
	5	0:2	1:3	0.0016	77.02	10.80

2.4.4. Analysis of results

The analysis of results focuses on examining the physical interpretations and prediction accuracy of the selected models.

<u>Physical interpretations</u>

Parameter estimates for RC ($T_o\phi_sQ_h$-T_iT_e) are listed in Table 2.10 together with the corresponding approximate standard errors. Methods for calculation of the standard errors can be found in [139]. The standard errors are small compared to the estimates, suggesting a relatively low model uncertainty. Besides, C_e (9.145 $kWh/$℃) should not only account for the thermal mass of the wood-framed building envelop but also that of the concrete floor/walls/slabs. Correspondingly, T_e should be regarded as the equivalent temperature of the building fabric.

Table 2.10 Parameter estimates and uncertainty for RC ($T_o\phi_sQ_h$ - T_iT_e)

Parameters	Estimates	Standard errors
$R_1(°C/kW)$	14.85	4.8747
$R_2(°C/kW)$	16.33	6.3594
$R_3(°C/kW)$	0.4818	0.0134
$C_i(kWh/°C)$	4.601	0.0622
$C_e(kWh/°C)$	9.145	0.3671
$F_s(m^2)$	10.15	0.3147

The selected RC model has demonstrated strong physical interpretability for the house's thermal behaviours. To be compared, the selected ARX (10, [1 10 10]) exhibits a more favourable testing performance. However, without applying any supervisory rules, it tends to have parameter estimates with large standard errors and poles that fail the pole constraint. Even though, respective summations of the estimates of a_j, $b_{1,j}$, $b_{2,j}$, and $b_{3,j}$ show low uncertainties (see Table 2.11). This suggests the model's potential for deriving important thermal properties.

Table 2.11 Parameter estimates and uncertainty for ARX (10, [1 10 10])

ARX (10, [1 10 10])	Estimates	Standard errors
$\sum_j a_j$	0.0034	0.0004
$\sum_j b_{1,j}$	0.0038	0.0004
$\sum_j b_{2,j}$	0.2775	0.0300
$\sum_j b_{3,j}$	0.0278	0.0026

As shown in Table 2.11Table 2.11, ARX (10, [1 10 10]) has automatically satisfied the steady-state constraint: $\sum_j a_j \approx \sum_j b_{1,j}$. By taking the average, the overall thermal transmittance of the house can be approximated by

$$U_{ARX} = \frac{\sum_j a_j + \sum_j b_{1,j}}{2 \cdot \sum_j b_{3,j}} = 0.1295\ kW/℃ \tag{2.16}$$

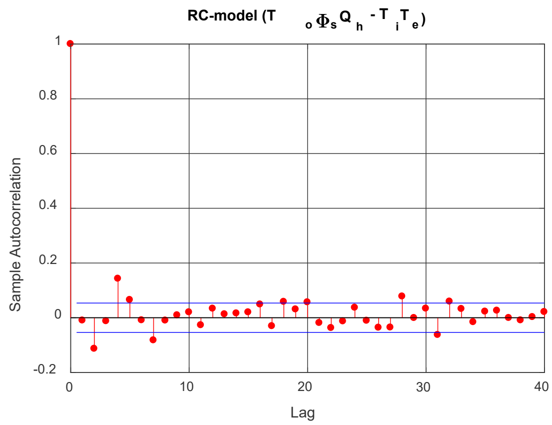
The equivalent solar aperture can be approximated by

$$F_{s,ARX} = \frac{\sum_j b_{2,j}}{\sum_j b_{3,j}} = 9.982\ m^2 \tag{2.17}$$

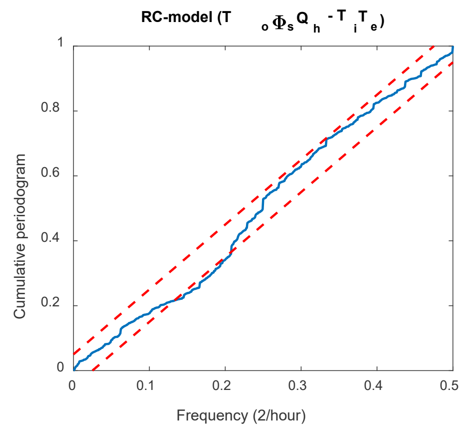
U_{ARX} and $F_{s,ARX}$ are very close to U_{RC} and $F_{s,RC}$, indicating consistent estimations of the overall thermal transmittance and solar aperture by RC and ARX models. However, the ARX model can only give thermal properties under steady-state conditions. Deriving properties that are associated with thermal dynamics (e.g., thermal capacitances) will require a proper RC model. In addition, the ANN model, as previously mentioned in Section 2.2.3. "AI models", cannot be used to infer any thermal properties.

Prediction accuracy

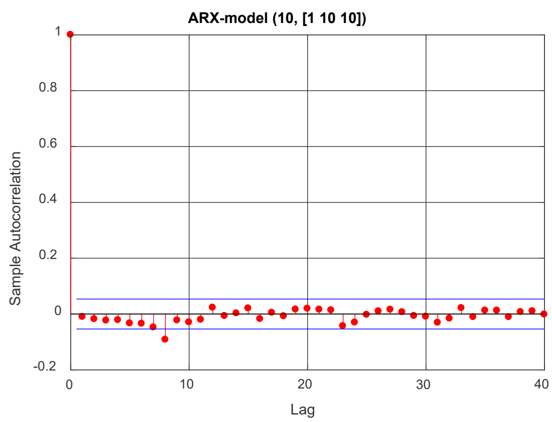
Autocorrelations and cumulative periodograms of the training residuals ($y - \hat{y}$) are plotted for the selected RC, ARX and ANN models in Figure 2.3. For each plot, the 95% confidence interval under the null hypothesis that the residuals are white noise is also shown (by parallel lines). Although RC ($T_o \phi_s Q_h$-$T_i T_e$) exhibits autocorrelations and cumulative periodogram outside the confidence region slightly more than ARX (10, [1 10 10]) and ANN (4, 0:2, 1:3), there is a clear low dependency of autocorrelations on lags and periodograms on frequencies for all three models. It is reasonable to accept the hypothesis that the residuals are white noise, indicating the thermal dynamics are well modeled.



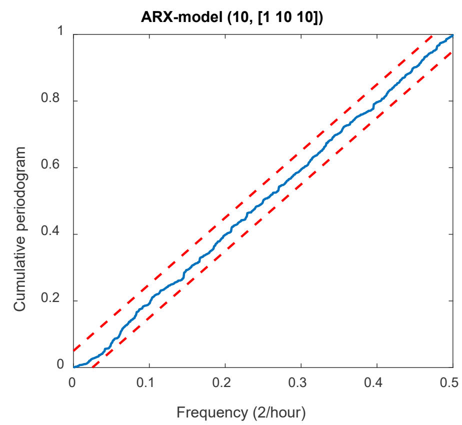
(a)



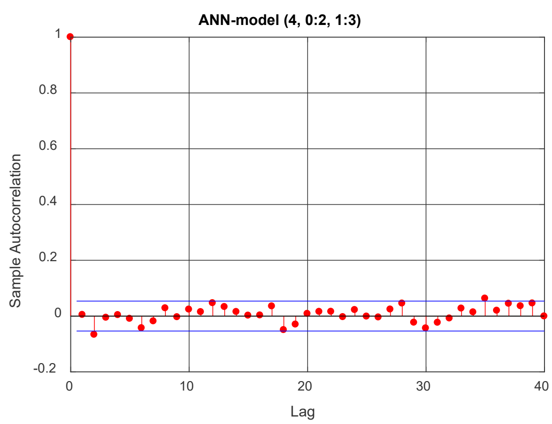
(b)



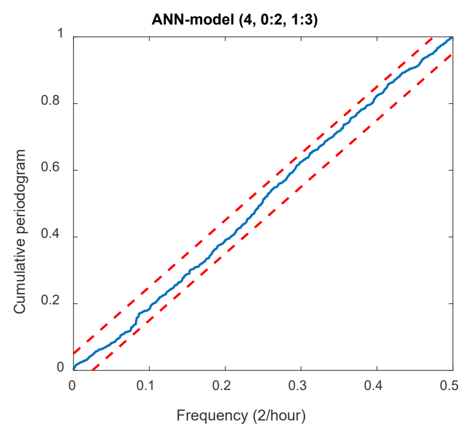
(c)



(d)



(e)



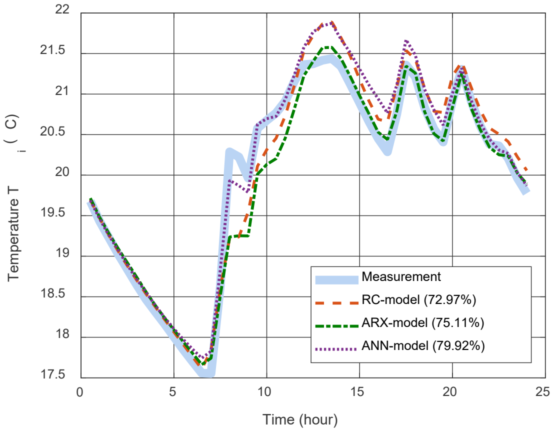
(f)

Figure 2.3 Autocorrelation and cumulative periodogram of residuals (with 95% confidence intervals)

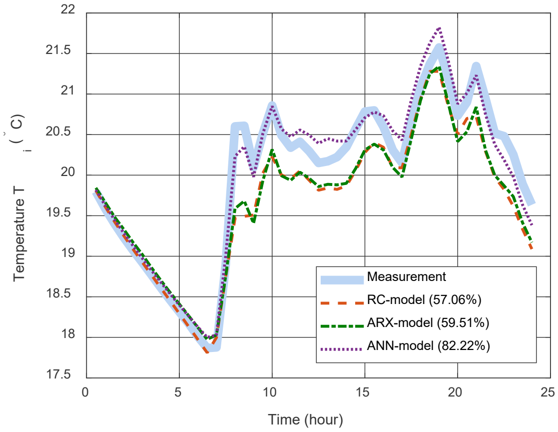
The testing performance of the selected models is summarized in Table 2.12. The ANN model, with the highest $\overline{Fit}\%$ and the lowest $Fit\_sd\%$, demonstrates strong ability and reliability of one-day ahead forecasting. The forecasted indoor air temperatures by the selected models are displayed against the measurements for four representative days in the testing period (Figure 2.4). For most days, the ANN model has higher forecasting accuracy and can capture dynamics overlooked by the RC and ARX models (e.g., Figure 2.4 (a, b)). On worse days for the RC and ARX models (e.g., Figure 2.4 (b, d)), though the forecasting accuracy is unfavorable, the forecasting errors at most of the time are constrained within 0.5 ℃; the maximum error is no more than 1 ℃ and only occurs occasionally. So, the RC and ARX models' forecasting ability is still acceptable.

Table 2.12 Summary of the testing performance of selected models

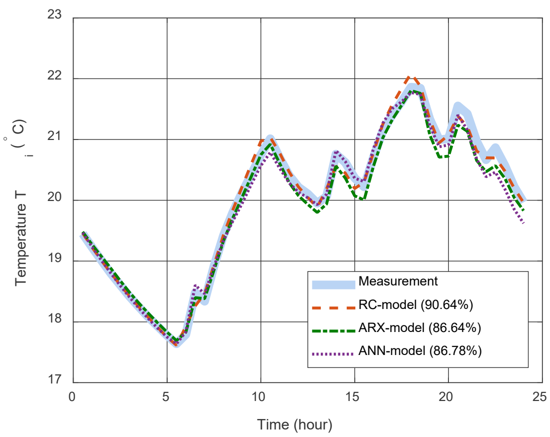
Selected models	$\overline{Fit}$ %	Fit_sd %
RC ($T_o\phi_sQ_h$ - T_iT_e)	72.60	12.12
ARX (10, [1 10 10])	76.25	13.02
ANN (4, 0:2, 1:3)	78.36	7.00



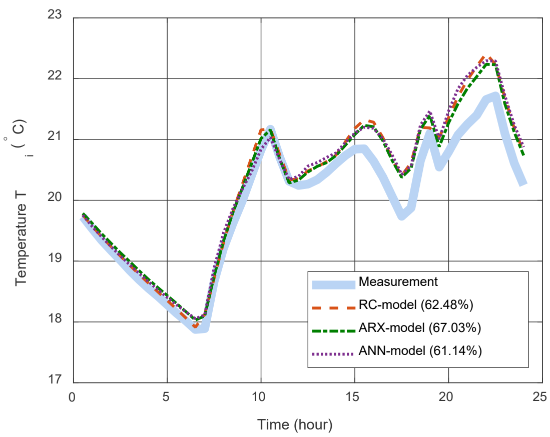
(a)



(b)



(c)



(d)

Figure 2.4 Testing results on representative days: (a) model performance around $\overline{Fit}$; (b) ANN model outperforming RC and ARX models; (c) model performance above average $\overline{Fit}$; (d) model performance below $\overline{Fit}$. In parenthesis is the Fit of each model for that day.

Although the ANN model outperforms the RC and ARX model in terms of forecasting accuracy, it is hardly possible to infer any thermal properties from the estimated weights and biases. Due to the large size of parameters, the ANN model may subject to multiple local minima during model training and demand more computation power. Nevertheless, the ARX model is easy to formulate and allows for linear least squares for model training so that the loss function is guaranteed to be globally optimized. The RC model, on the other hand, is the most physically plausible, requires the least parameterization, thus tend to be more robust under errand input signals.

2.5. Discussion

The RC, TF, and AI models are manifestly different by their structure formulations. The RC models are constructed based on a series of ordinary differential equations that characterize thermal dynamics in buildings. The appropriate design of RC models requires a thorough understanding of the studied thermal system. Moreover, since the ordinary differential equations are set up in continuous-time, discretization is required for the models to be trained with measured data. In contrast, the TF models are more straightforward to develop. They simply use rational functions or polynomials to incorporate the input and output variables. Since the rational functions or polynomials are already in discrete-time, discretization is not needed. Compared to the linear RC and TF models, the AI models can be regarded as nonlinear regression models that have complex inner structures built upon machine learning techniques. Due to their inherent implicit nonlinearity, the AI models allow for raw measurements (e.g., supply air temperature and flow rate instead of the heating power) to be directly used as inputs.

Another difference between these three types of models arises from their physical interpretability. Both RC and TF models have their counterparts in the forward approach, e.g., thermal networks and comprehensive room transfer functions. They are mathematically connected through state-space representations. In the RC models, the equivalent thermal parameters are positive and have intuitive physical meanings. Hence they are suitable for interpretation or explanatory purposes. However, the parameters in the TF models are not directly physically interpretable. Without applying proper constraints on the parameters, physically implausible or unstable models may be obtained. For AI models, they are impossible for any physical meaning interpretation. This is because the models are in no accordance with the forward approach and their structure are too complicated to be physically understood.

In terms of training, TF models (ARX and SS models in particular) are computationally more efficient than RC models. This is because that in the RC models, the ordinary differential equations and their discretization can make the outputs highly nonlinear with respect to the unknown parameters. Due to the same reason, more local minimums may occur in the RC models' training. As for AI models, other concerns like overfitting can be more crucial than computational efficiency.

Most ANN models are trained using backpropagation-based searching algorithms. Since there are often a lot of parameters (e.g., weights and biases) to be estimated, overfitting can become a serious problem that influences the model's generalization. A conventional technique to avoid overfitting is early stopping which uses a portion of the training dataset for cross-validation and ends training when the cross-validation error starts to increase [78]. The problem can also be overcome by including the model complexity (e.g., number of neurons) as an objective to minimize [88]. However, not all the AI models are subjected to overfitting. For example, the GRNN models can be trained in one pass through the data with no need for any iterative algorithm [92, 93]. An essential advantage of the GRNN models is fast-learning. Besides, the SVM models are trained with the structural risk minimization principle which defines a trade-off between the fitting quality and model complexity [99]. This intrinsic feature of SVMs also prevents the model from being overfitted.

2.6. Summary

Data-driven modelling of building thermal dynamics consists of three phases: modelling, training, and selecting. TF, RC and AI models are three main categories of data-driven models. The RC models are the most suitable for physical interpretation, the TF models are the easiest to formulate, and the AI models can conveniently manage nonlinearity and complex interactions between inputs. In the training phase, the prediction error method (PEM) and maximum likelihood estimation (MLE) are two popular methods to build up loss functions. Unknown parameters in the data-driven models are estimated by minimizing the values of the loss function, which is accomplished by either linear or nonlinear search algorithms. In the selecting phase, the most suitable model structure is selected through a forward or backward selecting procedure. It is a trade-off process that balances the model's prediction accuracy against its complexity. Quality criteria, residual analyses, etc. are used to validate, test, and compare model candidates with different inputs and structures.

The whole data-driven approach is illustrated by the case study of a single-zone house. Three models, i.e., an RC model, an ARX model, and an ANN model, are developed for the thermal dynamics of the house. After training and selecting, all models exhibit favourable forecasting ability, with the ANN model generally outperforming the RC and ARX models. On the other hand, both the ARX and RC models can be used to derive important thermal properties of the house, but the RC model serves better for explanatory purposes. Finally, the ARX model has the advantage of being trained by linear least squares.

Chapter 3. RC Model Development

3.1. Introduction

The RC model captures building thermal dynamics using a network of thermal resistors and capacitors. Thermal resistance, capacitances, and other necessary parameters in the network are estimated by tuning the model outputs to measured outputs. Unlike pure statistical models (e.g., autoregressive models), parameters of the RC models tend to be physically interpretable [23, 25]. For example, the estimated R's value of a composite wall can represent the wall's effective thermal resistance. Some existing studies in the literature took advantage of this intrinsic feature of RC models to derive important thermal properties of building components [36-38]. Some other studies focused more on obtaining a physically plausible and statistically well-performing model for characterizing building thermal dynamics [29, 140] as well as predicting transient building load or indoor temperatures [32, 141].

Nevertheless, parameter estimates of the RC models are not guaranteed to be physically interpretable: when model parameters are non-identifiable (i.e., cannot be uniquely determined), their physical meanings will be ambiguous. The concept of non-identifiability encompasses two notions: structural non-identifiability and practical non-identifiability [142, 143]. The structural non-identifiability arises from parameter redundancy where the model parameterization is not unique regardless of measurement patterns. The practical non-identifiability, on the other hand, indicates a parameter estimate is not confident using available data. In other words, non-identifiability has two main causes: the model structure is overcomplicated (causing structural and practical non-identifiability), or the measurements are inadequate (causing practical non-identifiability).

Given that the measured data is sufficient and in sound quality, a complex model structure can still lead to non-identifiability. On the contrary, an oversimplified model structure is incapable of explaining the thermal dynamics. Model selection is a standard process in finding a suitable RC model structure. In model selection, several potential model structures are constructed as candidates and then compared through statistical tests [42], validation criteria [41], or frequency responses [24]. Even a relatively suitable model structure is often obtained in the end, constructing all potential candidates and selecting the best one is rather time-consuming. This chapter proposes an alternative approach which first creates one complex preliminary model structure and then removes non-identifiable parameters (instead of comparing candidates). This approach is relatively robust as the resulted model structures will have the advantage of being both physically interpretable and computationally efficient.

The objective of this chapter is to propose a methodology for obtaining reliable RC model structures for thermal dynamic analysis of houses. The methodology is presented in Section 3.2. “Methodology” followed by a case study illustrating this methodology in Section 3.3. “Case Study”. Then relevant discussions, such as the physical interpretability of the simplified model structure, will be discussed in Section 3.4. “Discussion”.

3.2. Methodology

This section describes a methodology for developing reliable and simple RC models for thermal dynamic analysis of houses. The methodology consists of model formulation, model training, and model simplification. In the model formulation phase, the necessity and formulation of a multi-zone model are investigated, as well as the associated assumptions. In the model training phase, criteria for model training and testing are introduced, and model parameters are estimated when using the methodology. In the model simplification phase, a model simplification procedure is proposed as an alternative to model selection. The basic idea behind the proposed methodology is to start from a complex model structure and simplify it to a more suitable one.

3.2.1. Model formulation

Due to uneven heating or cooling, houses can exhibit considerable thermal stratifications across floors, with a floor being 1 to 2 ℃ higher than the one below it. Furthermore, solar gains may cause the equator-oriented rooms warmer than those that are not. Under such cases, single-zone models are often incompetent. To reflect the room temperature difference, a house can be divided into multiple thermal zones and modelled accordingly. However, having more zones means impractically more measurement points and higher complexity in the model structures and hence causes potential non-identifiability of parameters. For a typical house, a reasonable way is to treat each floor as one thermal zone [35]. It is also possible to further divide each floor into south and north zones to consider the local nonuniformity of solar energy. The methodology being presented aims to accommodate multi-zone RC models for houses.

In modelling building thermal dynamics, zonal temperatures are considered as outputs while outdoor temperature, solar radiation, heating power, etc. are selected as inputs. An RC network maps the input-output relationship. Some important assumptions for constructing the RC model include:

i. Air in each zone is well mixed;
ii. Heat transfer coefficients for conduction, convection, and radiation are constant;
iii. Natural ventilation rate is constant;

iv. Thermal conduction is one-dimensional within each building component; and

v. Heating power is uniformly distributed across the whole house.

Based on the above assumptions, the thermal dynamics of each temperature node in the RC network are governed by the ordinary differential equation expressed in Eqn. (2.1). Similar ordinary differential equations are established for all the temperature nodes and rearranged into a state-space representation where model inputs, outputs, and parameters are clearly defined. An example of the RC model formulation can be found in Table 2.1.

The acquired state-space representation is in continuous-time and needs discretization. This methodology assumes piecewise linear interpolation for discretization (see Table 2.2). The final model expression will be in the form of Eqn. (2.3), where the matrices $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, and $\boldsymbol{D}$ contain the model parameters to be estimated while training the model with measured inputs and outputs.

3.2.2. Model training

In this methodology, the well-known Prediction Error Method (PEM) is adopted for model training [134]. A corresponding algorithm is provided in the System Identification Toolbox™ in MATLAB. In short, the PEM estimates unknown model parameters by minimizing the value of an objective function of prediction or simulation errors. The model training here focuses on simulation (i.e., infinite step ahead prediction or $h = \infty$). When the simulation errors are assumed to be jointly Gaussian with zero mean and time-invariant unknown covariances, we have:

$$\boldsymbol{\theta} = \min_{\boldsymbol{\theta}} V(\boldsymbol{\theta}) \tag{3.1a}$$

$$V(\theta) = \det\left\{\frac{1}{N}\sum_{t=1}^{N}[\boldsymbol{y}_t - \widehat{\boldsymbol{y}}_t(\boldsymbol{\theta})]\,[\boldsymbol{y}_t - \widehat{\boldsymbol{y}}_t(\boldsymbol{\theta})]'\right\} \tag{3.1b}$$

where,

det is the determinant operator;

$V(\cdot)$ is the objective function in a maximum likelihood sense [123, 134];

N is the number of data samples;

$\boldsymbol{\theta}$ is the vector of model parameters to be estimated;

$\boldsymbol{y}_t$ is the measured output (i.e., a column vector of zonal temperatures) at t^{th} time step; and

$\hat{\boldsymbol{y}}_t$ is the simulated model output at t^{th} time step.

The trained model is then validated with new data (i.e., testing data). RMSE is used as a criterion to evaluate the performance of the model for both training and testing. It is defined (for the j^{th} output) as follows:

$$RMSE_j = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left[y_{j,t} - \hat{y}_{j,t}\right]^2} \qquad (3.2)$$

where $y_{j,t}$ the j^{th} entry of the output vector $\boldsymbol{y}_t$.

Another testing criterion *Fit* is also used. *Fit* can be seen as a normalized RMSE term expressed as a percentage. It is defined (for the j^{th} output) as follows:

$$Fit_j = \left(1 - \frac{\sqrt{\sum_{t=1}^{N}\left(y_{j,t} - \hat{y}_{j,t}\right)^2}}{\sqrt{\sum_{t=1}^{N}\left(y_{j,t} - \bar{y}_j\right)^2}}\right) \cdot 100\% \qquad (3.3)$$

where $\bar{y}_j$ is the sample mean of y_j.

Comparing these two criteria, both RMSE and *Fit* emphasize larger errors by taking the square of the residuals. RMSE expresses the average output error in degree Celsius. The smallest RMSE indicates the best training or testing performance. *Fit*, on the other hand, informs how closer the data are to the fitted curve compared to a straight line (i.e., $\bar{y}_j$). The larger the *Fit* is, the more accurately a model can fit the measurements.

3.2.3. Model simplification

A good RC model should have both a simple structure and satisfactory accuracy but seeking such a model is often not straightforward. Previous studies [20, 42] employed a trial-and-error method to search for the best model structure from a group of candidates based on specific criteria. Such a selection process can be quite laborious especially when a model has a lot of possible structures (e.g., a multi-zone model). Alternatively, we can start from a complex preliminary RC model structure that can be created following physical principles and then progressively simplify it by removing non-identifiable parameters. A complex RC model structure includes many parameters to reflect the detailed thermal interactions between different building components.

This simplification process involves mainly two challenges. First, the RC model needs to be trained to have a satisfactory fitting at training and testing data. Unsatisfactory fitting often originates from over-parameterization or unreliable initial guesses of model parameters. Local searching algorithms tend to give unfavourable training results, especially for a complex model structure when initial guesses for model parameters are far away from the true values. Second, non-identifiable parameters must be precisely detected. Parameters that are not identifiable have negligible or no influence on the model predictions. It is, therefore, reasonable to remove those non-influential parameters. Whereas, non-identifiability is not straightforward in some cases. In linear regression, the significance test [144] is sufficient for detecting non-influential parameters. However, for such ordinary differential equations as in RC models, extra efforts are needed and will be presented below.

To overcome the first challenge, a model structure can be trained first by a global search algorithm (e.g., genetic algorithm) for a rough search and followed by a local search algorithm (e.g., the Levenberg-Marquardt algorithm) for a refined search. Global search algorithms have been applied to RC models [32, 34] and proven to give favourable estimates. Genetic algorithm (GA) is adopted in the model development methodology. GA imitates the process of biological evolution. It starts with a population of randomly generated individuals (i.e., initial estimates of θ) within assumed bounds. All individuals are measured by a fitness function (i.e., the objective $V(\theta)$) and those of small fitness function values are selected as elites. A new generation is produced from the elites through operations like mutation and crossover. The GA solver terminates when either a fixed number of generations is reached, or the difference between the best fitness values of two consecutive generations are less than a small value (e.g., 1e-6). The initial population is set to 200 for parameter size larger than 5, the elites take up 5% of the population, and the maximum number of generations is 30.

The second challenge is to detect non-identifiable parameters and remove them. A rigorous method for non-identifiability analysis is to examine the profile likelihood [142, 143] near the parameter estimates. This method requests a series of re-optimization of the model for each parameter and is rather computationally demanding. Another common method is approximate but simpler, e.g., using Hessian or asymptotic covariance [145-147] to detect non-identifiable parameters. Under typical excitation conditions, non-identifiable parameters detected by the asymptotic-covariance-based method tend to align with those by the profile-likelihood-based method [112]. Thus, the asymptotic-covariance-based method is adopted in the methodology and explained as follows.

Since the prediction error estimator is asymptotically normally distributed [123, 148], for large enough sample size N, we have,

$$\sqrt{N}\left(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right) \xrightarrow{\mathrm{d}} \mathcal{N}(0, \boldsymbol{P}_{\boldsymbol{\theta}}) \tag{3.4a}$$

$$\boldsymbol{P_\theta} = (\mathbb{E}\boldsymbol{V_{\theta\theta}})^{-1}(\mathbb{E}\boldsymbol{V_\theta V_\theta'})(\mathbb{E}\boldsymbol{V_{\theta\theta}})^{-1} \tag{3.4b}$$

where,

$\hat{\boldsymbol{\theta}}$ is the estimate of the model parameter vector $\boldsymbol{\theta}$;

$\boldsymbol{P_\theta}$ is the asymptotic covariance matrix [148]; and

$\boldsymbol{V_\theta}$ is the gradient vector $\partial V/\partial\boldsymbol{\theta}$ and $\boldsymbol{V_{\theta\theta}}$ is the hessian matrix $\partial^2 V/\partial\boldsymbol{\theta}^2$.

Estimation for $\boldsymbol{P_\theta}$ can be obtained during model training. Standard errors of the parameter estimates are merely the square root of diagonal elements in the asymptotic covariance matrix, i.e., $\hat{\boldsymbol{\sigma}}_{\boldsymbol{\theta}} = \sqrt{diag(\boldsymbol{P_\theta})/N}$. The more a parameter affects the model outputs, the easier it will be to determine its value, and the less uncertainty the parameter estimate will have. It can be seen from Eqn. (3.4) that large standard errors mean large uncertainty and thus indicates non-identifiability. In addition, parameters in the RC models are strictly positive but the normal distribution in Eqn. (3.3) is unbounded. To take the constraint of positivity into account, it is advantageous to reparametrize the parameters by log-transform [147]. According to the delta-method [149], the log-transformed estimator is also asymptotically normally distributed, namely, $\left(\log\hat{\theta} - \log\theta\right) \to \mathcal{N}\left(0, \hat{\sigma}_\theta/\hat{\theta}\right)$ for any parameter θ in the parameter vector $\boldsymbol{\theta}$. Here, a parameter is defined to be identifiable only when $t_{\delta/2}\,\hat{\sigma}_\theta/\hat{\theta} < 25\%$, where $t_{\delta/2}$ is the t-statistic given significance level δ. This corresponds to saying with $(1-\delta)\times 100\%$ confidence that $\theta \in \left(\hat{\theta}e^{-0.25}, \hat{\theta}e^{0.25}\right)$. By taking $\delta = 0.05$, we can say with 95% confidence that true values of the identifiable parameters fall between 78% ($\approx e^{-0.25}$) and 128% ($\approx e^{0.25}$) of their estimates.

The model structure simplification procedure is summarized below (Step 1 ~ 3). The flowchart in Figure 3.1 schematically shows the simplification process (together with model formulation and model training). This simplification approach has a definite advantage: there is no need to construct a lot of candidate models to choose the best from. Since the simplification is achieved by directly removing non-identifiable parameters, it requires fewer rounds of model formulations and comparisons so that it is more efficient than the trial-and-error method.

- Step 1

First, train the complex model using a genetic algorithm to obtain initial guesses for model parameters. Then, train the model with Levenberg–Marquardt algorithm based on the initial guesses for a refined search. The refined search yields the parameter estimate $\widehat{\boldsymbol{\theta}}$, the standard error estimate $\widehat{\boldsymbol{\sigma}}_{\boldsymbol{\theta}}$, and the optimized objective function value $V(\widehat{\boldsymbol{\theta}})$. To assure that the model has an acceptable fitting ability, $V(\widehat{\boldsymbol{\theta}})$ must not exceed a small number μ (e.g., $\mu = 10E - 5$). Otherwise, the model is not well structured, and the simplification procedure should be ended.

- Step 2

Remove non-identifiable parameters whose standard error to parameter estimate ratio ($\hat{\sigma}_\theta/\hat{\theta}$) surpasses a threshold ϵ). The threshold is defined differently in every round of the simplification. Its value should have the same order of magnitude as the largest $\hat{\sigma}_\theta/\hat{\theta}$ in that round (see the thresholds in Table B.1 for example). Only a small number of parameters are removed in each round. When a parameter is removed, any component or input that depends on this parameter should also be removed. For example, in Table 2.1, if $R_{2,3}$ is removed from the network, T_3 is also discarded.

- Step 3

Repeat Step 1 and 2 to train the reconstructed model until there are no non-identifiable parameters ($t_{\delta/2}\,\hat{\sigma}_\theta/\hat{\theta} < 25\%$ for all θ) or the optimized objective function value becomes too large ($V(\widehat{\boldsymbol{\theta}}) > \mu$). Whichever condition is met, the simplification procedure ends, and the results are exported for analysis. When $V(\widehat{\boldsymbol{\theta}}) > \mu$, the current model is no longer successful at fitting the training data, so the last model should be kept. Thus, this step delivers a model that not only has a simple structure but also fits the measured data with adequate accuracy.

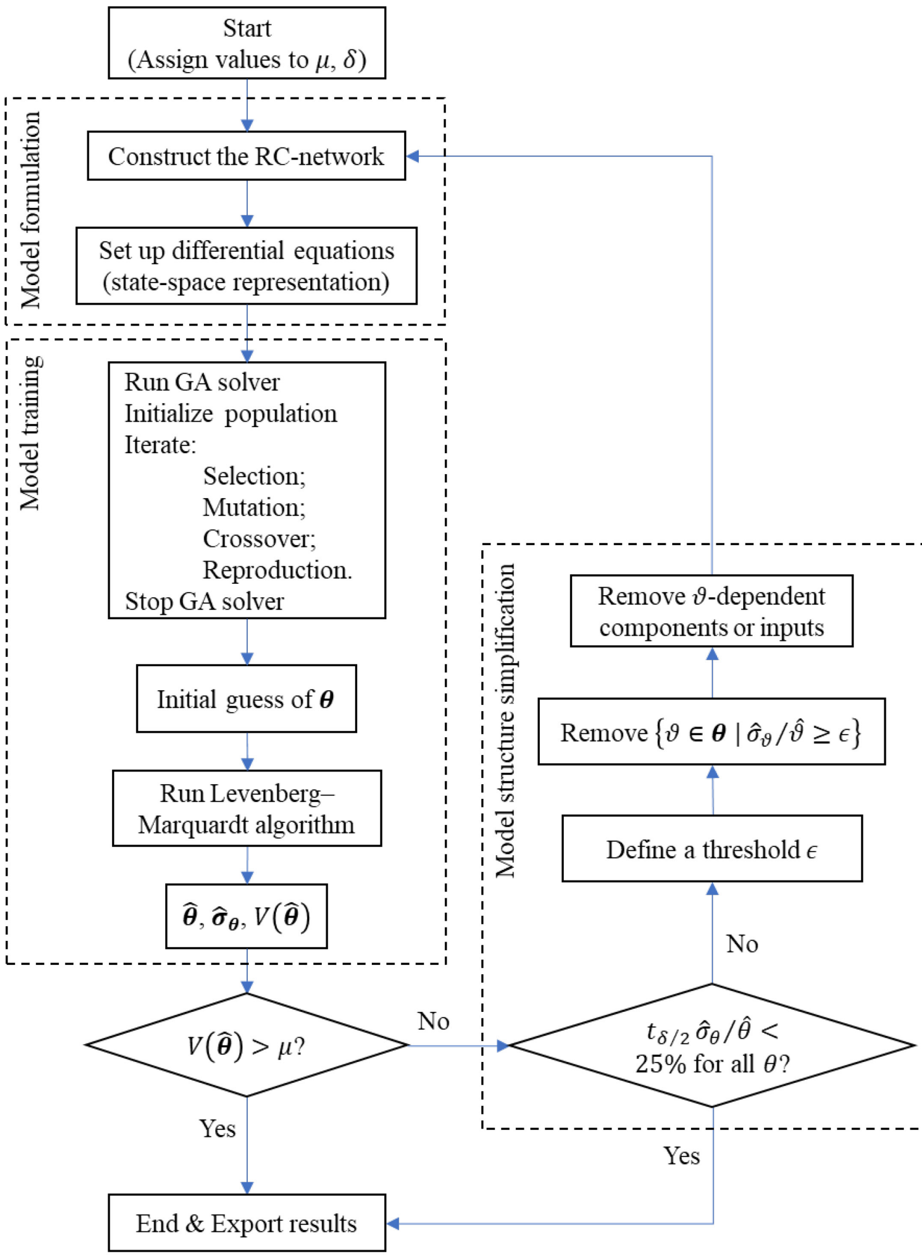


Figure 3.1 Flow chart of the model development methodology (model formulation, model training, and model structure simplification).

3.3. Case Study

The proposed methodology is applied to a low-energy house [150, 151] to create a simple and sufficient RC model. The house is a wood-framed two-story single-detached home (including a basement) located in Eastman, Quebec, Canada. It has large glazing areas facing south. Energy system and locations of thermal mass are shown in Figure 3.2. There is a significant amount of thermal mass from the concrete floor/slab/walls in both the living room and the basement. Space heating is mainly provided by a geothermal heat pump through forced hot air. In the basement, there is a ventilation concrete slab which can also provide some space heating when there is under-floor warm air circulation.

This house was monitored for a period of several years. Raw data obtained during monitoring was preprocessed (formatting, synchronization, deleting outliers, etc.) into measurements [152] that can be used to train RC models. To account for temperature stratification of the house, a three-zone model is adopted in this case study: basement, first floor, and second floor. Following the model development methodology presented above, a complex preliminary model structure is formulated and trained first, then simplified to a simple and sufficient structure. This section will present the measured data, demonstrate the methodology, and evaluate the modelling results.

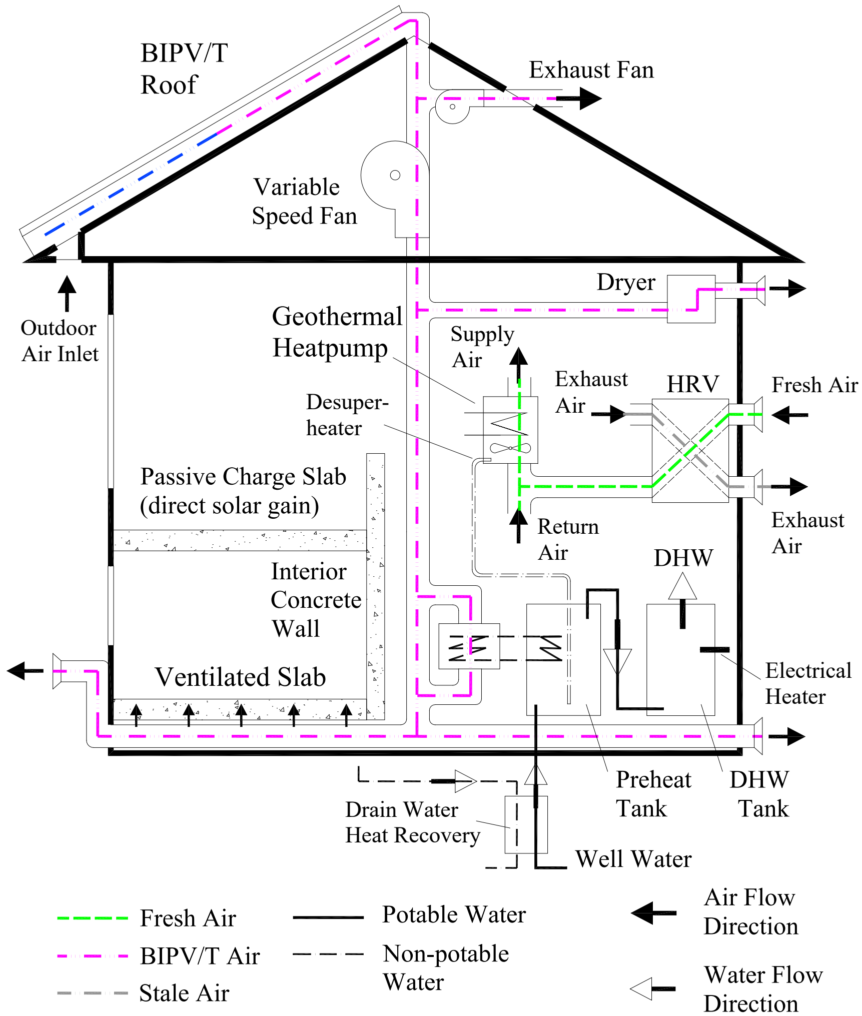


Figure 3.2 Energy system and location of thermal mass in ÉcoTerra™ house [150]

3.3.1. Measured data

Relevant input variables are summarized in Table 3.1. Global irradiation on the south façade (ϕ_s) and gross electricity demand (Q_{elec}) are used to approximate effective solar heat gains and internal heat gains [109], respectively. Heating power provided by the geothermal heat pump (Q_{hp}) is calculated based on the measured air flow rate and temperature difference between supply air and return room air. The thermal energy charged to the slab (Q_{vcs}) is calculated based on the measured temperature difference between the inlet and outlet air and the flow rate of the air, which was heated by the roof-mounted solar thermal system. Ground temperature is assumed to be constant (i.e., $T_g = 13$°C). The desired model outputs are the zone air temperature T_{2f} for the second floor, T_{1f} for the first floor, and T_{0b} for the basement.

Table 3.1 Summary of model input and output variables

Variables	**Unit**	**Description**
T_o	℃	Outdoor air temperature
T_g	℃	Ground temperature
ϕ_s	kW/m^2	Global irradiation on the south façade
Q_{hp}	kW	Heating power provided by the geothermal heat pump
Q_{elec}	kW	Gross electricity demand
Q_{vcs}	kW	Heat charged to the ventilation concrete slab
T_{0b}, T_{1f}, T_{2f}	℃	Air temperatures of the basement, first floor, and second floor

All measurements are sampled every 10 minutes. This sampling interval is chosen to be sufficiently small to avoid overlooking of important thermal dynamics. The measurements consist of the entire dataset and are then divided into one training dataset and two testing datasets (see Figure A.2) for testing 1 and 2, respectively. The training dataset contains sufficient difference among various inputs and outputs so that non-identifiability is hardly caused by the data quality [111]. Two testing datasets represent two different situations. In the first testing dataset, zonal temperatures have significant variations within 24 hours, but the daily pattern of the temperatures barely changes. In the second testing dataset, zonal temperatures are less excited and in relatively low frequencies (i.e., more variations from day to day). A suitable RC model should perform well on both scenarios.

3.3.2. Model development

The measured data show that in the training dataset, *RMSE*s of the temperature differences between the basement and the main floor and between the main floor and the second floor are 1.03℃ and 1.68℃, respectively. Hence, the zonal temperature differences are not negligible, and a three-zone model is necessary. The house is divided into three zones: basement (notation: 0b), main floor (notation: 1f), and second floor (notation: 2f).

The complex preliminary model structure is created based on physical principles and shown in Figure 3.3 (a). In this model structure, there are five types of model parameters: thermal resistances (R), thermal capacitances (C), solar gain factors (F), internal gain factors (p), and heating power distribution factors (α). Thermal resistances (R) and capacitances (C) describe the thermal dynamics between temperature nodes. Solar gain factors (F) are applied on global irradiation to approximate effective solar heat gains. They are closely related to window areas (often interpreted as solar apertures). Gross electricity demand (Q_{elec}) is weighted by internal gain factors (p) to approximate effective internal heat gains. Sum of p should be no greater than one ($\sum_{j=1}^{3} p_j \leq 1$). Heating power ($Q_{hp}$) is distributed to each zone through a distribution factor (α). The distribution factors, instead of being estimated from the measurements, are pre-calculated proportionally to each zone's heated floor area. For the basement, first-floor, and second-floor, the distribution factors are respectively, $\alpha_0 = 0.35$, $\alpha_1 = 0.38$, and $\alpha_2 = 0.27$ ($\sum_{j=1}^{3} \alpha_j = 1$).

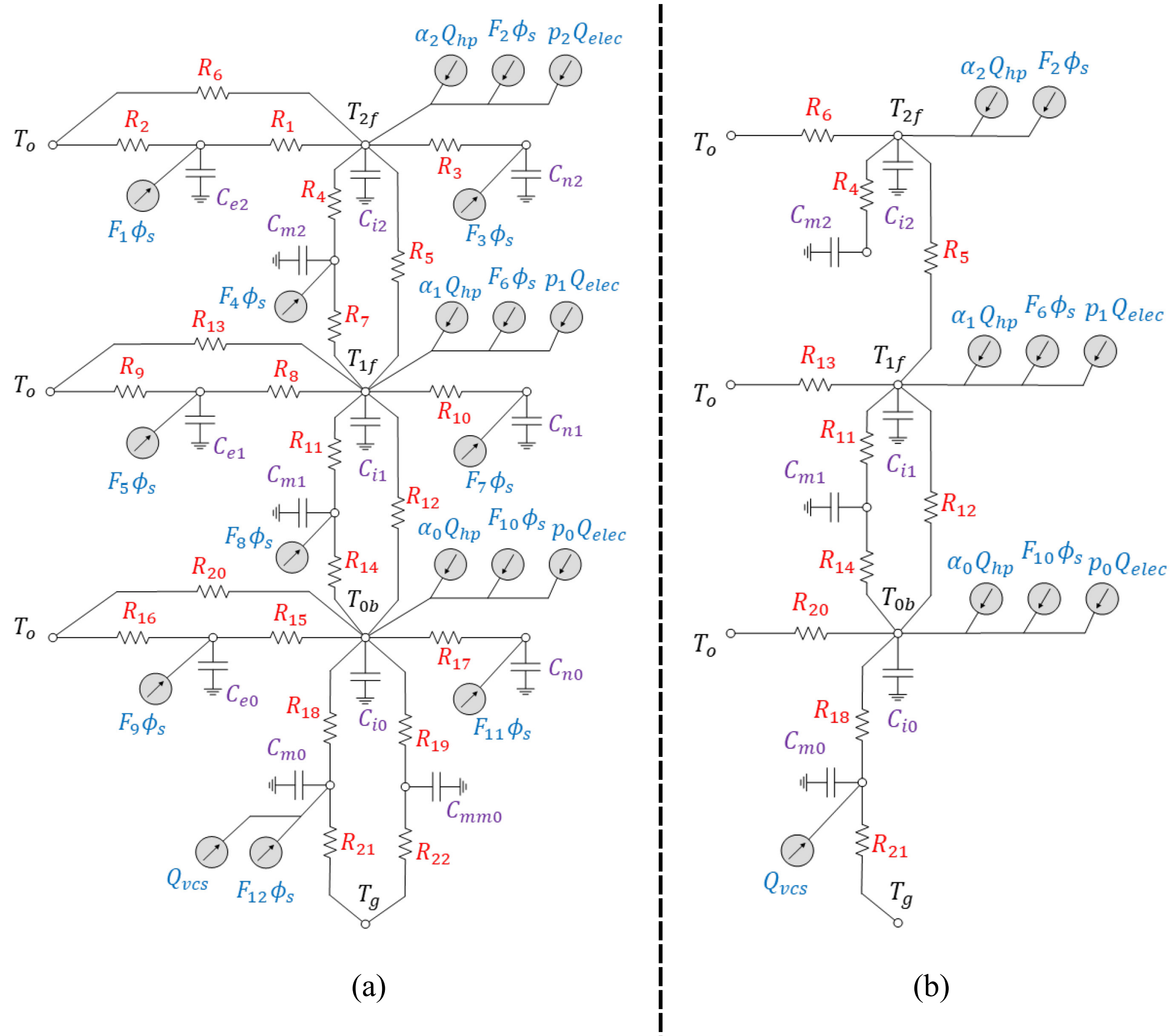


Figure 3.3 Model structures: (a) the complex three-zone model structure, and (b) the simplified three-zone model structure

In the complex model structure, each level's building envelop is modelled by three resistances and one capacitance ($3R1C$), including a fast response path (e.g., R_6) and a slow response path (e.g., R_2-C_{e2}-R_1). The fast response path captures the thermal impact through windows and ventilation while the slow response path is mainly to characterize the transient conduction through thick walls, roof, and ceilings. Inside each zone, it is modeled by a $2C1R$ where one thermal capacitance C_i is for indoor air (i denotes indoor air) and the other capacitance

C_n belongs to building fabric (e.g., walls and floors). Between zones, it is modeled by a $3R1C$ with a fast response path (e.g., R_5) and a slow response path (e.g., R_4-C_{m2}-R_7). The two paths respectively represent convective and transient conductive heat transfer between adjacent floors. For the basement, specifically, the slab thermal mass is modeled by two thermal capacitances: C_{m0} is for the ventilated slab and C_{mm0} for the non-ventilated slab.

Using the presented methodology (Section 3.2), the complex preliminary model structure is simplified shown in Figure 3.3 (b). Details of the simplification process and the parameter estimates are provided in Table B.1 in the appendix where non-identifiable parameters to be removed in each turn are highlighted. The simplification has significantly reduced the complexity level of the model structure. For example, slow response paths of the building envelop are removed for all three zones. Thermal capacitances associated with the building fabric are discarded. For the influence of solar radiation, only the solar gains to indoor air remain in the model structure. Moreover, the internal heat gains of the second floor are estimated as zero. Each zone ends up being modelled by two capacitances and three to four resistances. In summary, the simplified model structure is of $6C10R$ (i.e., 6 thermal capacitances and 10 thermal resistances) compared to the complex preliminary structure of $13C22R$ (i.e., 13 thermal capacitances and 22 thermal resistances). The model complexity measured in terms of the number of thermal capacitances and resistances has been reduced by more than 50%.

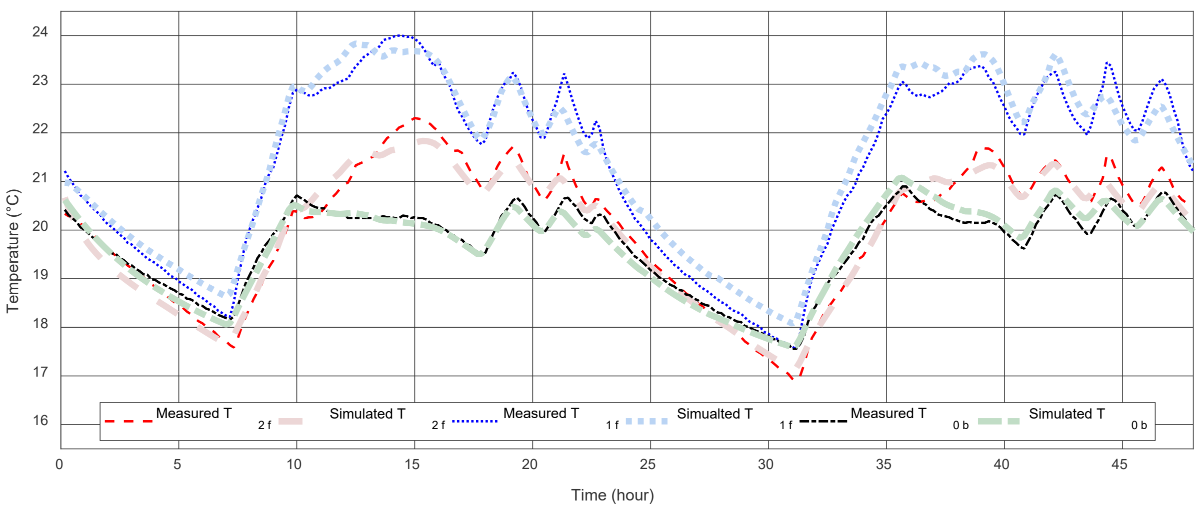
3.3.3. Results and analysis

The results of model training and testing are summarized in Table 3.2, which shows that simplification of the complex preliminary model structure causes no significant loss of training/testing accuracy measured concerning both RMSE and *Fit*. The simplified model exhibits favourable testing performance (with all *Fit* larger than 60%). This can also be seen in Figure 3.4, where the simulated zonal temperatures are displayed against the measurements for two typical days in both testings.

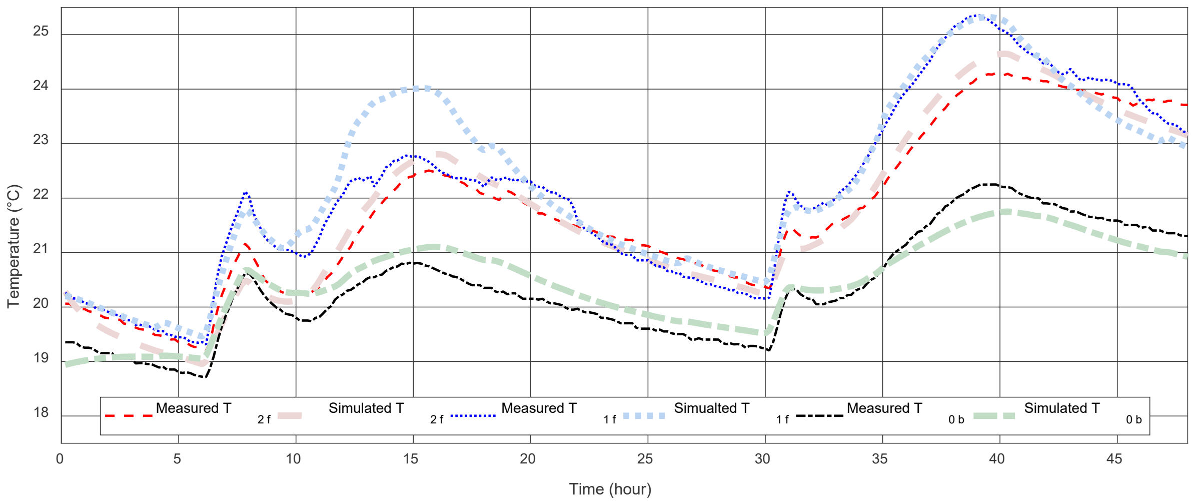
Table 3.2 Training and testing performances of the three-zone model

Model	**Zone air temperature**	**Training**		**Testing 1**		**Testing 2**	
		RMSE (℃)	*Fit* (%)	RMSE (℃)	*Fit* (%)	RMSE (℃)	*Fit* (%)
Complex	T_{2f}	0.275	77.0	0.295	71.5	0.623	60.7
	T_{1f}	0.296	79.7	0.376	74.6	0.439	70.8
	T_{ob}	0.139	80.9	0.231	74.5	0.221	79.5
Simplified	T_{2f}	0.277	76.8	0.319	69.1	0.632	60.1
	T_{1f}	0.300	79.5	0.416	71.9	0.427	71.6
	T_{0b}	0.145	80.2	0.235	74.0	0.272	74.7

There is a relatively large decrease in fitting accuracy for second-floor air temperature (T_{2f}) from training to testing 2, e.g., *Fit* drops from 76.8% to 60.1% for the simplified model (see Table 3.2). Yet, a similar decrease of fitting accuracy happens to the complex model structure as well (*Fit* drops from 77.0% to 60.7%). In addition, an obvious discrepancy occurs between the measured and the simulated first-floor air temperatures (T_{1f}) around 3:00 pm on the first testing day (hour 15), as shown in Figure 3.4 (b). However, such discrepancy is not observed around 3:00 pm on the next day (hour 39), which has similar weather conditions. Therefore, the decrease of fitting accuracy and the temperature discrepancy is not due to the model simplification, but unmodelled influences that are only notable in the testing 2, such as occupant activities, e.g., closing blinds, that result in little or no solar radiation getting transmitted through the windows. Certainly, the model's simulation accuracy can be increased by integrating influences due to occupant activities. Nevertheless, even with the decrease of fitting accuracy and occasionally discrepancies, the largest RMSE of the simplified model structure is merely 0.632℃ (for T_{2f} on testing 2), which is acceptable for the model to acquire the primary thermal dynamics of the house when compared to the variation of indoor air temperatures (around 5℃).



(a)

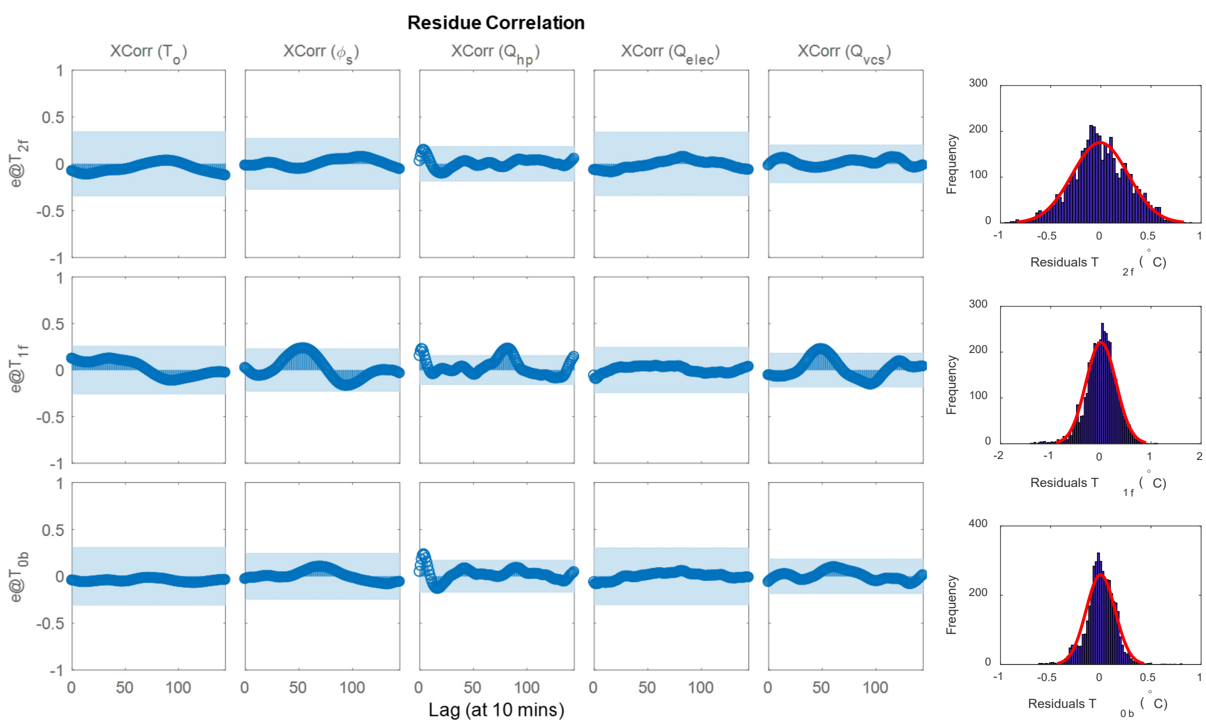


(b)

Figure 3.4 Comparison between measured and simulated zonal temperatures for the first two days in (a) testing 1, and (b) testing 2.

From the complex model structure to the simplified one, a total of 35 non-identifiable parameters are removed (Table B.1), and all the remaining parameters are identifiable, i.e., for each parameter, $t_{\delta/2}\,\hat{\sigma}_{\theta}/\hat{\theta} < 25\%$. This suggests that parameter estimates of the simplified model structure are of enough confidence so that they can be used for evaluating the actual performance of the thermal system. Moreover, the simplification process is finished in 10 rounds. In other words, only 10 model structures need to be constructed to simplify the complex model structure which has 50 unknown parameters. This indicates the presented methodology is very efficient towards finding a suitable model structure.

For the simplified model structure, estimated cross-correlations of residuals ($e = y - \hat{y}$) with input signals on training dataset are shown in Figure 3.5 (a). The 99% confidence interval under the null hypothesis that the residuals and inputs are statistically independent is also shown (by shaded regions). The maximum time lag here is 144 corresponding to 24 hours (1440 minutes) at a 10-minutes sampling interval. Since only a small number of cross-correlations exceed the confidence region, it is reasonable to accept the hypothesis that the residuals are uncorrelated to input signals. This further indicates that the simplified model structure has captured the essential part of the thermal dynamics from inputs to outputs. Histograms of the simulation residuals are also shown in Figure 3.5 (b). For each zone, the residuals can be well fitted to a normal distribution. This agrees with the assumption during model training that the simulation errors are jointly Gaussian.



(a) (b)

*Figure 3.5 Residual analysis: (a) cross-correlation (XCorr) of residuals "**e**" with input signals and the 99% confidence region marking statistically insignificant correlations displayed as a shaded region around the X-axis. (b) histograms of residuals fitted well with the normal distribution*

To justify the capability of the simplified model structure to address temperature stratification, the RMSEs from the two testing's (see Table 3.2) are compared with the inter-zone RMSEs (i.e., RMSEs of temperature differences between adjacent zones). Most testing RMSEs in Table 3.2 are less than 0.5℃ (except for the RMSE in testing 2 for the second floor). However, the inter-zone RMSEs between the basement and the main floor and between the main floor and the second floor are 1.42℃ and 1.51℃ in testing dataset 1, and 0.88℃ and 1.44℃ in testing dataset 2, respectively. Thus, most inter-zone RMSEs are greater than 1℃. This implies that the simplified three-zone model structure can effectively distinguish adjoining zonal temperatures. This is also visualized in Figure 3.4. It can be seen that for most of the time, temperature differences between adjacent zone air temperatures are precisely captured by the RC model with a simplified model structure.

Some may suggest that each floor level should be further divided into a north and a south zone to take into consideration the nonuniform solar gains. However, due to continuous air circulation and open space design (except for the bedrooms on the second floor) in this house, temperatures of adjacent south and north sections are nearly the same though slightly different during sunny daytime. Slight temperature differences can hardly be recognizable. Furthermore, a model of more zones will contain more model parameters and require more input variables. Therefore, separating a floor level into more zones, in this case, is not necessary.

3.4. Discussion

The results presented in Section 3.3.3 "Results and analysis" show that by removing all the non-identifiable parameters, the model uncertainty is significantly reduced. In other words, the parameter estimates are obtained with enough confidence and tend to be more physically interpretable. For example, the wood-framed building envelop serves as an excellent insulator, but it has insignificant thermal capacitance. Thus, the slow response path of the building envelop is removed during simplification. However, the removal of slow response path does not apply between the zones that contain massive concrete. Inside the house, there are a concrete slab and a wall exposed to both the main floor and the basement air, so they contribute thermal capacitance to both the zones. Hence, a common thermal capacitator has remained between the main floor and the basement. Moreover, the second floor contains few electric appliances and therefore internal heat gains indicated by the gross electricity demand is negligible (i.e., $p_2 = 0$).

In addition, the parameters in the RC model are not exact but equivalent. Their estimates only enable the RC network to imitate building thermal dynamics and do not match precisely with the apparent quantities. For example, the apparent value of indoor air thermal capacitance (specific heat times its volume) for the main floor zone is about $0.1\,kWh/$°C whereas C_{i1}(as reported in Table B.1, Model 10) is $3.3\,kWh/$°C. Therefore, C_{i2}, C_{i1}, and C_{i0} do not represent indoor air thermal capacities. Instead, each of them can be interpreted as an effective thermal capacitance including the indoor air, furniture, and surface layers of the building fabric. On the other hand, C_{m2}, C_{m1}, and C_{m0} can be explained as thermal capacities of inner layers of the building fabric. Thus, thermal resistances R_4, R_{11}, R_{14}, and R_{18} reflect how effectively the inner layers of building fabric influence indoor air temperatures (smaller resistances, more effective). Usually, these thermal resistances can be reduced by increasing exposure of internal thermal mass (e.g., uncovered concrete slab) to indoor air.

As illustrated by the case study, the proposed model development methodology is easy to implement and fulfills our goal to obtain a reliable RC model structure for characterization of building thermal dynamics. In practice, the methodology can be further enhanced by addressing the following points. First, the training data must be sufficient and in good quality to guarantee the estimated asymptotic covariance matrix is reliable. Data that contain insufficient or unfavourable information can cause the wrong removal of parameters. Second, heuristic optimization algorithms (e.g., GA) in global search may not at once give satisfactory initial estimates (within limited generations) to be used by the local search. If necessary, the GA solver can be run for multiple times until a valid result is obtained. Third, nonlinear local search can be slow for large model structures in reaching the optima. To proceed, one can limit the local search time and conservatively remove non-identifiable parameters (e.g., remove one at a time).

The RC models obtained using the proposed methodology have broad potential applications. Since all or most of the parameters are identifiable, their estimates are of adequate confidence. Valuable knowledge of, such as passive heat storage and effective solar gains, can be identified. The obtained models, therefore, can be used for evaluating the actual performance of houses and informing future designs. Since the models have satisfactory simulation accuracy, they can also be used for model predictive control, fault detection and diagnosis, ongoing commissioning, and so on.

3.5. Summary

A data-based model development methodology has been proposed to acquire reliable RC model structures for thermal dynamic analysis of houses. The methodology is conceptually simple and yet practically effective. In principle, it starts with creating a complex preliminary model structure based on physical principles, and then non-identifiable parameters are progressively removed to obtain a simplified model structure. In this methodology, a global search algorithm (i.e., GA) is adopted for better initial estimates, and non-identifiable parameters are quantified using asymptotic standard errors. The proposed methodology is illustrated through the development of a three-zone RC model for a low-energy house.

The simplified model structure for the house is validated. The results reveal that the simulation accuracy is favourable and has not been jeopardized after simplification. Besides, training residuals of the simplified model structure are insignificantly correlated with the inputs, suggesting thermal dynamics of the house are well modelled. Furthermore, this three-zone model has demonstrated a satisfactory capability of characterizing temperature stratification across the floors. Finally, the simplified model structure and the simplification process are justified from a physical perspective. Therefore, the obtained RC model proves to be reliable for simulation of building thermal dynamics as well as physical interpretation of the thermal characteristics.

Chapter 4. Application of RC Models to Building Performance Evaluation

4.1. Introduction

RC models are especially suitable for evaluating building thermal behaviours. Compared to autoregressive models (e.g., ARX models), parameters of the RC models can be directly physically interpretable [23, 25]. For example, the estimated R's value can represent the effective thermal resistance of a building envelop. Furthermore, RC models allow for direct interpretation of internal thermal mass (e.g., concrete wall and slabs) through the estimated thermal capacitances that cannot be identified on steady-state conditions or using autoregressive models.

The majority of the previous studies of RC models focused on predicting building load or indoor air temperatures for model-based control of space heating and cooling [6-8, 40]. Though some studies employed data-driven models to estimate thermal properties of building components, the estimation is limited to thermal transmittance of building envelop and solar heat gain coefficient [36-38]. A comprehensive interpretation of thermal capacitances has not yet seen in the literature. Besides, it was shown that data-driven models could be applied to analyzing energy use or qualifying energy management in buildings [17, 68]. However, few studies have used data-driven RC models to address passive thermal energy storage in buildings.

The objective of this chapter is to apply RC models to evaluating the actual thermal performance of buildings. It is organized as follows. Section 4.2. "Methodology" briefly describes the methodology of this chapter. Section 4.3. "Parameter Interpretation of RC Models" associates the estimated parameters of RC models to specific building components through design and modelling of a single-zone room. Then, Section 4.4. "Evaluation of Building Thermal

Performance” applies a three-zone RC model to a real house to evaluate its thermal performance by inferring important thermal properties, analyzing energy flow paths, and investigate the function of internal thermal mass. In the end, the results are discussed in Section 4.5. "Discussion” and the key findings are summarized in Section 4.6. “Summary”.

4.2. Methodology

The methodology consists of two phases of studies in order to employ RC models to evaluate the actual performance of existing designs in buildings and identify potential design improvement.

The first phase of study investigates how parameter estimates in RC models are related to design configurations. First, a simple single-zone room is designed with different configurations. Under each configuration and given inputs, the room’s air temperature (i.e., output) is simulated by the finite difference method. Then, the given inputs and the obtained output are used as “measurements” to train a low-order RC network. Finally, parameters of the RC network are interpreted by comparing their estimates under different configurations and relate them to passive thermal energy storage (TES) design.

The second phase of the study is to develop a three-zone RC model for a real house, then apply it to evaluate the house’s design and thermal performance. First, some essential thermal properties are derived from the estimated parameters. Second, different energy flow paths are identified along with their relative significance. Third, parameter values of the three-zone RC model are altered to create different scenarios, and their influences on indoor air temperature fluctuation and energy consumption are compared. Finally, the comparison results are used to evaluate the passive TES system of the house and identify potential design improvement.

4.3. Parameter Interpretation of RC Models

In order to investigate the relationships between the apparent building properties and the estimated parameters of RC models, a simple single-zone room is designed with different configurations. A 31-day weather dataset is used to simulate the thermal performance of this room. The simulated indoor air temperature is recorded as system output and used to train a low-order RC network. By varying the design configurations (i.e., the thickness and area of the concrete slab), their influences on the estimates of the RC parameters can be revealed.

4.3.1. Design of a single-zone room

A single-zone room is designed with three sets of configurations (or cases), as shown in Figure 4.1. For each case, the building envelop is composed of wooden material and built upon a slab foundation. Within the building envelop, there is a large and clear single-lane window facing south. The slab is partially or fully concrete, and there is a layer of insulation between the concrete and the soil. The design configurations being varied are the thickness and area of the concrete slab. In case 1, the concrete slab occupies half of the foundation area, and the other half is filled with dense insulation ($R = \infty$). The concrete slab's thickness is doubled in case 2 but its area remains the same. Case 3 has a concrete slab of the same thickness as in case 1, but the area is doubled. These three cases are designed mainly to interpret parameters related to internal thermal mass.

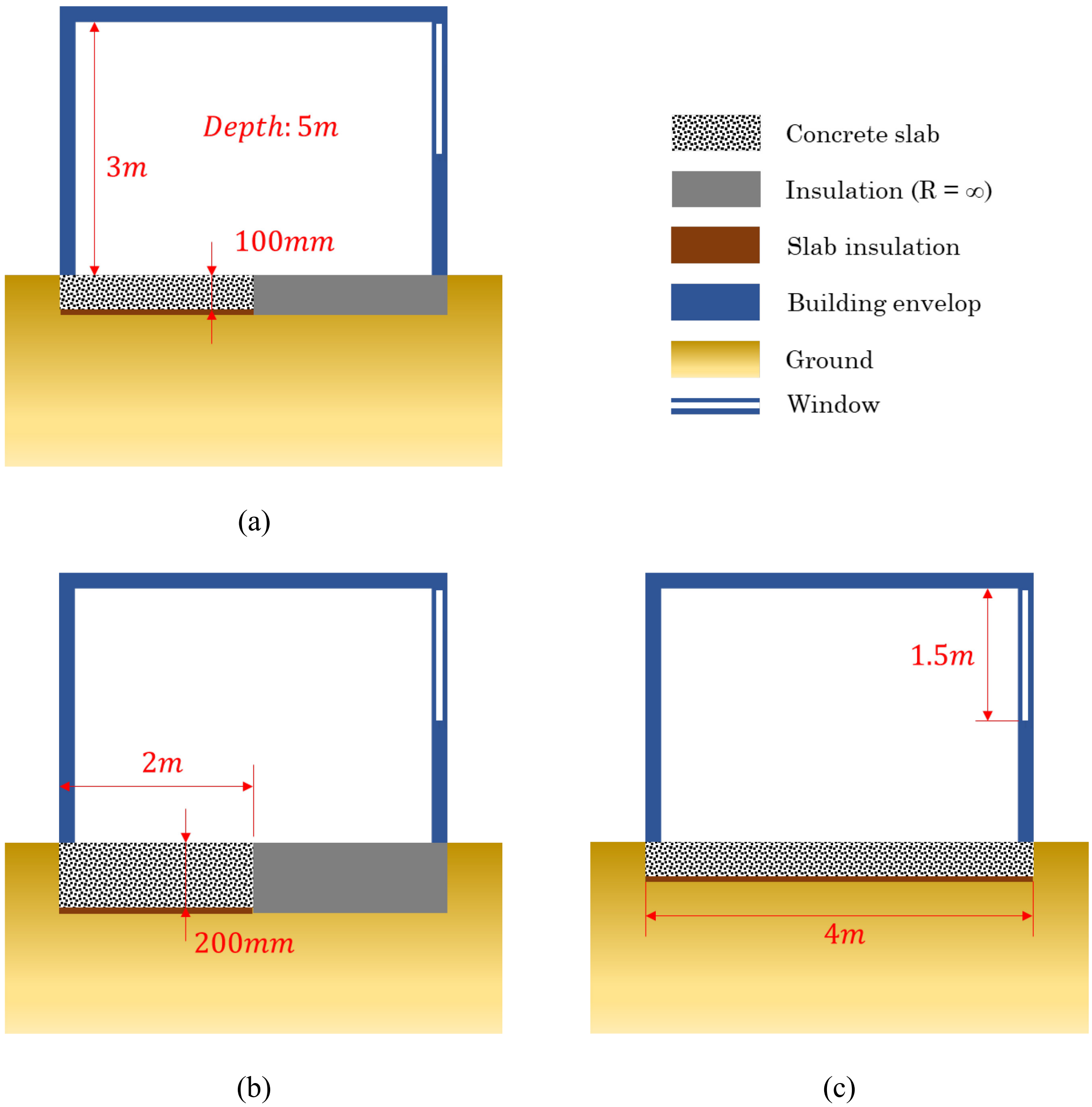


Figure 4.1 Three different configurations in the design of a single-zone space: (a) case 1, concrete slab with half the area; (b) case 2, concrete slab with half the area but doubled thickness; (c) case 3, concrete slab with the full area.

The design specifications are summarized in Table 4.1. Thermal dynamics of the single-zone room are simulated by the finite difference method. One-dimensional heat transfer normal to the component surfaces is assumed. Inputs to the room (i.e., outdoor air temperature, global solar radiation, ground temperature, and heating power) are retrieved from the training dataset in Figure A.2. The simulated indoor air temperature is recorded as an output every 10 minutes for 31 days.

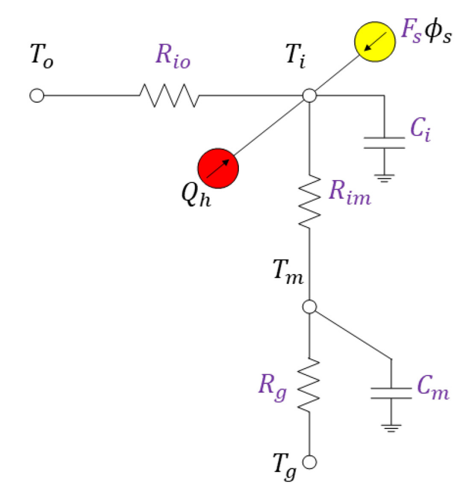
Table 4.1 Design specifications for three cases of the single-zone house

		Case 1	Case 2	Case 3
Thermal resistance ($°C/kW$)	Indoor air – Outdoor air	31.7	31.7	31.7
	Indoor air – Slab top	11.1	11.1	5.6
	Concrete slab	12.5	25.0	6.3
	Slab bottom – Soil	113.3	113.3	56.7
Thermal capacity ($kWh/°C$)	Indoor air	0.2167	0.2167	0.2167
	Building envelop	0.9806	0.9806	0.9806
	Concrete slab	0.5	1.0	1.0
Window (facing south)	Area (m^2)	7.5	7.5	7.5
	Average SHGC	0.519	0.519	0.519

4.3.2. Training of a 3R2C network

A 3R2C network, as shown in Table 4.2, is proposed to represent a thermal system of the single-zone room. It is trained using the modelled output and inputs. The model training method can be found in Section 3.2.2. "Model training". In the network, parameters to be estimated include the thermal resistance between indoor and outdoor air (R_{io}), the thermal resistance between indoor air and concrete slab (R_{im}), the thermal resistance between the concrete slab and ground (R_g), the thermal capacitance of indoor air (C_i), the thermal capacitance of the concrete slab (C_m), as well as the solar aperture (F_s). The training results are shown in Table 4.2.

Table 4.2 Structure and parameter estimates of the 3R2C network for the single-zone room



Parameter	Case 1	Case 2	Case 3
R_{io} (°C/kW)	31.79	32.35	31.82
R_{im} (°C/kW)	6.032	9.024	4.127
R_g (°C/kW)	133.2	111.5	63.86
C_i (kWh/°C)	0.3823	0.4535	0.3816
C_m (kWh/°C)	0.6857	0.7487	1.111
F_s (m^2)	3.971	3.936	3.983

4.3.3. Interpretation of model parameters

A comparison of Table 4.1 and Table 4.2 indicates that the initial descriptions of C_i and C_m are not precise. For example, every estimated C_i (0.3823 kWh/°C for case 1, 0.4535 kWh/°C for case 2, or 0.3816 kWh/°C for case 3) is over 70% larger than the designed indoor air thermal capacity (0.2167 kWh/°C). This implies that C_i represents not only the thermal capacitance of indoor air but also some thermal capacitance from the building envelop and the concrete slab. It should be interpreted as the thermal capacitance of indoor air, furniture, and some surface layers of the building fabric. This is further confirmed when the thickness of the concrete slab is doubled from case 1 to case 2, the estimate of C_i also increases notably (from 0.3823 kWh/°C to 0.4535 kWh/°C). Thicker concrete has allowed more surface layers to have temperatures close to that of the indoor air. However, only increasing the surface area of the concrete slab (as from case 1 to case 3) has negligible influence on C_i. This is because that larger area of the concrete slab makes it harder to bring up the surface layers' temperatures.

Correspondingly, C_m represents inner layers of the building fabric that can not be easily penetrated by the indoor air temperature. Its estimate is affected by both the building envelop and the concrete slab. For example, the estimated C_m for case 1 ($0.6857\ kWh/$°C) is larger than the designed concrete slab thermal capacity ($0.5\ kWh/$°C). Thus, part of C_m must have been contributed by the building envelop. However, since the building envelop is more of an insulator than a capacitor, C_m depends mostly on the amount of concrete. Furthermore, the estimate of C_m only reflects the effective thermal mass. From case 1 to case 2, the amount of concrete is doubled, but the estimate of C_m has not been increased for more than 10% (from $0.6857\ kWh/$°C to $0.7487\ kWh/$°C). From case 2 to case 3, the amount of concrete remains the same but the estimate of C_m mainly increases due to the doubling of concrete surface area. Therefore, the concrete slab in case 2 may have been over-designed and ineffectively used.

Thermal resistances are comparatively more straightforward for interpretation: R_{io} is the thermal resistance associated with the building envelop; R_g is the thermal resistance associated with the slab and ground; R_{im} denotes the thermal resistance between the indoor air and inner layers of the building fabric. The estimated R_{io} (31.79 °C$/kW$ for case 1, 32.35 °C$/kW$ for case 2, or 31.82 °C$/kW$ for case 3) aligns well with the designed indoor air – outdoor air thermal resistance (31.7 °C$/kW$). In addition, the estimated R_{im} is highly related to the concrete area exposed to the surrounding air. For example, from case 2 to case 3, doubling the surface area of the concrete slab results in a much smaller value of R_{im} (from 9.024 °C$/kW$ to 4.127 °C$/kW$).

This reduction of R_{im}'s value induces greater heat transfer between the indoor air and inner layers of the building fabric.

Finally, F_s represents the solar aperture or the solar gain factor that determines the effective solar gains to the indoor. For a well-insulated building, F_s can be interpreted as the multiplication of the total glazing area and the solar heat gain coefficient (SHGC) of the window. In all the three cases, the estimated F_s ($3.971\ m^2$, $3.936\ m^2$, and $3.983\ m^2$) agree precisely with its apparent value ($\text{Area} \cdot \text{SHGC} = 7.5m^2 \times 0.519 = 3.893m^2$). Comparatively, R_{io} and F_s are two parameters in the 3R2C network that have the most precise physical meanings.

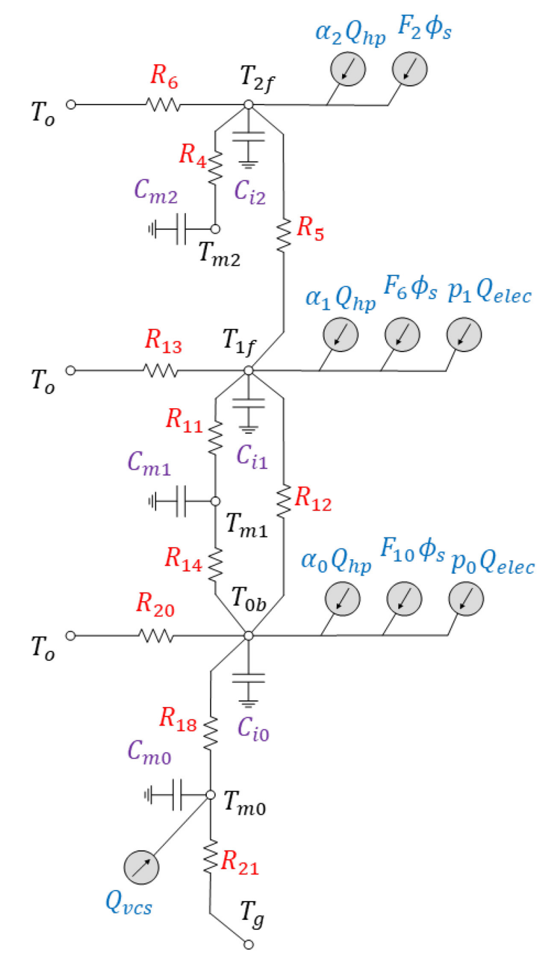
4.4. Evaluation of Building Thermal Performance

In this section, the three-zone RC model developed in Section 3.3 "Case Study" is applied to evaluating the thermal performance of ÉcoTerra™ house through energy balance analysis and investigation of internal thermal mass. Figure 3.2 gives the energy system and location of thermal mass in the low energy house. A description of measurements and a summary of the input and output variables can be found in Section 3.3.1.

4.4.1. Inferring building thermal properties

The three-zone model has a similar zonal structure as the RC network in Table 4.2. For the parameters, C_{i2}, C_{i1}, and C_{i0} fall in the category of C_i; C_{m2}, C_{m1}, and C_{m0} fall in the category of C_m; R_4, R_{11}, R_{18}, and R_{20} correspond to R_{im}; R_6, R_{13}, and R_{20} correspond to R_{io}; R_{21} corresponds to R_g; F_2, F_6, and F_{10} fall in the category of F_s. As such, parameters of the three-zone model in Table 4.3 have similar meanings as the parameters in the 3R2C network.

Table 4.3 The simplified model structure and the parameter values



Parameter	Value	Parameter	Value
R_4	1.0	C_{i1}	2.0
R_5	1.3	C_{m1}	19.5
R_6	14.4	C_{i0}	2.9
R_{11}	3.4	C_{m0}	10.0
R_{12}	7.1	F_2	2.6
R_{13}	33.7	F_6	4.2
R_{14}	10.5	F_{10}	1.3
R_{18}	2.1	p_2	0.4
R_{20}	24.7	p_3	0.1
R_{21}	13.9	α_0	0.35
C_{i2}	1.8	α_1	0.38
C_{m2}	3.3	α_2	0.27
Units: R (°C/kW); C (kWh/°C); F (m^2)			

Some important thermal properties can be derived from the three-zone model. For example, the overall thermal transmittance of the above-grade building envelop of the studied house can be calculated by

$$U_{tot} = \frac{1}{R_6} + \frac{1}{R_{13}} + \frac{1}{R_{20}} = 139.6\, W/\text{°C} \qquad (4.1)$$

The studied house has a reported building envelop area of $354m^2$ (walls: $227.5m^2$; windows and doors: $39.5m^2$; ceilings: $87m^2$). Using the trade-off compliance path recommended by [153], the overall thermal transmittance for a reference building is calculated to be $148.8\,W/°\mathrm{C}$. If integrated with mechanical ventilation and infiltration/exfiltration effects, this reference value will be even larger. Since U_{tot} is already smaller than the reference value, it can be concluded that the building envelop is performing in good quality.

In addition, the solar gain factor can be calculated by

$$F_{tot} = F_2 + F_6 + F_{10} = 8.1\,m^2 \qquad (4.2)$$

The studied house has a reported south glazing area of $20.9m^2$. Dividing F_{tot} by the south glazing area, an average SHGC of 0.39 is retrieved. The house is installed with clear low-e triple glazing windows. According to Chapter 15 Table 10 of [12], the SHGC of 0.39 is rather small compared to similar windows. Therefore, the windows may not be operating at their best state.

4.4.2. Model-based energy analysis

Based on physical meanings of the model parameters, a more definite understanding of the house's thermal behaviours can be achieved by conducting energy balance analysis on the three-zone model to identify the relative importance of different energy flow paths. Energy flow paths can be found by integrating difference equations based on the three-zone RC network for a long enough period (e.g., a month). Let Δt be the sampling interval; t_1 and t_2 are respectively, the starting and ending time of a month. Take the first floor for example, $\alpha_1 \sum_{t=t_1}^{t_2} Q_{hp,t}\Delta t$ and $F_6 \sum_{t=t_1}^{t_2} \phi_{s,t}\Delta t$ are heat gains from space heating and solar radiation, respectively; ${R_{13}}^{-1} \sum_{t=t_1}^{t_2} \left(T_{1f,t} - T_{o,t}\right)\Delta t$ is heat loss to the outdoor; ${R_5}^{-1} \sum_{t=t_1}^{t_2} \left(T_{1f,t} - T_{2f,t}\right)\Delta t$ and ${R_{12}}^{-1} \sum_{t=t_1}^{t_2} \left(T_{1f,t} - T_{0b,t}\right)\Delta t$ are heat losses to adjacent zones; ${R_{11}}^{-1} \sum_{t=t_1}^{t_2} \left(T_{1f,t} - T_{m1,t}\right)\Delta t$

denotes net heat storage in the inner layers of building fabric. For foundation, $\sum_{t=t_1}^{t_2} Q_{vcs,t}\Delta t$ is heat gain from the VCS and ${R_{21}}^{-1}\sum_{t=t_1}^{t_2}\left(T_{m0,t} - T_{g,t}\right)\Delta t$ is heat loss to the ground.

To calculate heat gain/loss terms in the energy balance equations, node temperatures or system states (i.e., T_{2f}, T_{1f}, T_{0b}, T_{2m}, T_{1m}, T_{0m}) must be estimated over the data duration. Given the parameter values in Table 4.3, matrices in the state-space representation can be calculated (see Section 2.2.1 for Eqn. (2.3) and Table 2.2). The estimation can be performed as

$$\hat{\mathbb{x}}_{t+1} = \boldsymbol{A}\hat{\mathbb{x}}_t + \boldsymbol{B}\boldsymbol{u}_t \qquad (4.3a)$$

$$\hat{\boldsymbol{x}}_t = \hat{\mathbb{x}}_t + \boldsymbol{\Gamma}_2\boldsymbol{u}_t \qquad (4.3b)$$

where

$\hat{\boldsymbol{x}} = \left[\hat{T}_{2f};\ \hat{T}_{1f};\ \hat{T}_{0b};\ \hat{T}_{2m};\ \hat{T}_{1m};\ \hat{T}_{0m}\right]$;

u is the vector of inputs to the simplified model structure; and

$\boldsymbol{A}$, $\boldsymbol{B}$, and $\boldsymbol{\Gamma}_2$ are matrices of the simplified model structure.

The system states are estimated by Eqn. (4.3) using the dataset in Figure A.2 (a), Then, the heat losses and gains are calculated and listed in Table 4.4. The total heat loss for the studied house (for March 2010) is $2224\ kWh$. Around 85% of this energy is lost through the building envelop and ventilation to the outdoor while the remaining 15% is lost to the ground. The heat loss to the outdoor referred to the building envelop area ($354\ m^2$) is $5.3\ kWh/m^2$. The heat loss to the ground referred to the foundation slab area ($90\ m^2$) is $3.7\ kWh/m^2$. The phenomenon that the above-grade heat loss is faster than that below the grade implies a potential of reducing heat loss by enhancing building envelop or recovering more heat from the exhausting air.

Total heat input to the house is $759\ kWh$ in which solar gains ($759\ kWh$) represent 34.5%, the heat pump ($853\ kWh$) represents 38.8%, and internal heat sources represent 21.6%. Although the VCS contributes only $113\ kWh$ to space heating, it is able to balance out over 30% of heat loss to the ground. Additionally, the second floor draws a lot of energy from the first floor (481 kWh) and eventually loses it to the outdoor through large areas of walls and ceilings. This energy draw is even greater than the sum of heating supply and solar gains to the second floor. It is mostly caused by the mechanical air circulation that boosts the heat exchange between the second and the first floor. A possible solution is to provide more heating to the second floor by proper control.

Table 4.4 Energy flow paths: heat losses and gains (unit: kWh)

Energy flow \ Location		**Zone 2f**	**Zone 1f**	**Zone 0b**	**Foundation**	**Total**
Heat loss to	Outdoor	939	420	532	----	1891
	Ground	----	----	----	333	333
Heat gain from	Heating	209	339	305	----	853
	Solar	246	395	118	----	759
	Internal	0	373	102	----	475
	VCS	----	----	----	113	113
Building fabric	Charging	209	146	240	----	595
(inner layers)	Discharging	210	88	97	----	395
	Net storage	1 (gain)	58 (loss)	143 (loss)	----	200
From adjacent zone		481 (gain)	629 (loss)	148 (gain)	----	0

The net energy storage in inner layers of the building fabric can be regarded as a combined effect of thermal charging and discharging. Inner layers of the building fabric behave like a heat battery. This "battery" is charged when the indoor air is warmer (e.g., $T_{1f} > T_{m1}$) and discharged when the indoor air is cooler (e.g., $T_{1f} < T_{m1}$). This effect is visualized in Figure 4.2, where the estimated T_{m1} and T_{1f} are displayed for two days together with solar gains and heating supply to the first floor. It can be seen that some energy has been shifted from daytime to nighttime through charging and discharging. This shifted energy helps achieve a stable indoor air temperature and avoid temperature peaks during sunny noon.

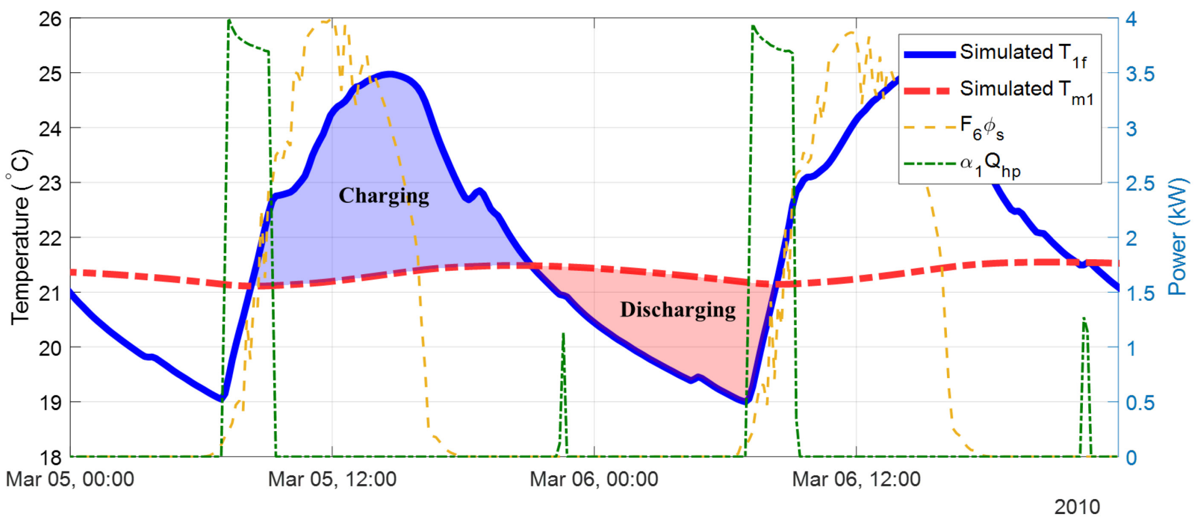


Figure 4.2 The estimated thermal mass temperature (T_{m1}) displayed for two days in March together with the estimated zone temperature (T_{1f}), solar gains ($F_6\phi_s$), and heating supply ($\alpha_1 Q_{hp}$).

As shown in Table 4.4, the energy shifted by charging ($595\ kWh$) and discharging ($395\ kWh$) is quite considerable. It suggests that inner layers of the building fabric have been effectively used for energy balance and stabilizing the indoor air temperature. However, due to the coupling of the building fabric (concrete wall and slabs in particular) and the ground, some energy charged to the thermal mass does not discharge to the indoor air but lost into the surrounding ground. This results in some negative net energy storage in the first and second floor. Therefore, reducing heat loss to the ground can also migrate the difference between daytime charging and nighttime discharging, which leads to more effective energy shifting.

4.4.3. Investigation of internal thermal mass

Due to varied heat losses and gains, air temperatures in a building always fluctuate. For the studied house in a typical day, the heat pump is turned on at dawn to heat the indoor air (e.g., see Figure 4.2). As a result of solar gains, the air temperature continues to increase during the day and ultimately drops after sunset. This diurnal change pattern results in a significant fluctuation of the indoor air temperature. Reducing this fluctuation can enhance the indoor thermal comfort.

To investigate the capability of internal thermal mass of stabilizing indoor air temperatures, the parameter estimates of internal thermal mass are altered to create different scenarios. In the three-zone model, there are three categories of parameters that are associated with the internal thermal mass: C_i's (C_{i2}, C_{i1}, and C_{i0}), R_{im}'s (R_4, R_{11}, R_{18}, and R_{20}), and C_m's (C_{m2}, C_{m1}, and C_{m0}). There are eleven alteration scenarios:

i. reference with no alterations ($C_i/R_{im}/C_m$);

ii. increase C_i by 50% ($\uparrow C_i = 1.5C_i$);

iii. decrease C_i by 50% ($\downarrow C_i = 0.5C_i$);

iv. increase ${R_{im}}^{-1}$ by 50% ($\downarrow R_{im} = R_{im}/1.5$);

v. decrease ${R_{im}}^{-1}$ by 50% ($\uparrow R_{im} = R_{im}/0.5$);

vi. increase C_m by 50% ($\uparrow C_m = 1.5C_m$);

vii. decrease C_m by 50% ($\downarrow C_m = 0.5C_m$);

viii. increase the thickness of building fabric: $\uparrow C_i$, $\uparrow R_{im}$, $\uparrow C_m$;

ix. decrease thickness of building fabric: $\downarrow C_i$, $\downarrow R_{im}$, $\downarrow C_m$;

x. increase the surface area of building fabric: $\downarrow R_{im}$, $\uparrow C_m$;

xi. decrease surface area of building fabric: $\uparrow R_{im}$, $\downarrow C_m$.

The operation of increase (↑) or decrease (↓) is performed on one category of parameters, e.g., $\uparrow C_i$ is to increase the values of all C_{i2}, C_{i1}, and C_{i0} by 50%.

All other inputs to the simplified model structure (i.e., T_o, ϕ_s, Q_{vcs}, Q_{elec}, and T_g) are unchanged and from the training dataset. For all the scenarios, heating power (Q_{hp}) and indoor air temperatures (T_{2f}, T_{1f}, and T_{0b}) are simulated with a simple control algorithm. The control algorithm is defined as follows:

Heating setpoints (reference: zone 1f):

Daytime (7:00 am to 11:00 pm): $22\ °C$

Nighttime (11:00 pm to 7:00 am): $18\ °C$

Dead band:

The control calls for heating when the reference temperature (T_{1f}) drops $1.0\ °C$ *below the setpoint and remains the heat on until it is* $0.5\ °C$ *above the setpoint.*

Heating power (by heat pump):

The maximum heating power $Q_{hp,max} = 9.6\ kW$

Heating power pulse: $0 \to \frac{1}{2} Q_{hp,max} \to Q_{hp,max} \to \cdots \to Q_{hp,max} \to 0$

The sampling interval is 10 minutes. Control commands (e.g., turn the heat on) at time t are made depending on the observed T_{1f} at time $t-1$. As the system states are initialized, the simulation can proceed iteratively.

To evaluate temperature fluctuation, variances of the simulated T_{2f}, T_{1f}, and T_{0b} are computed for each scenario. Take zone 1f for example, the variance is defined as

$$var(T_{1f}) = \frac{1}{N-1}\sum_{t=1}^{N}\left(T_{1f,t} - \frac{\sum_{t=1}^{N} T_{1f,t}}{N}\right)^2 \qquad (4.4)$$

where,

N is the number of samples (here, $N = t_2 - t_1 + 1$); and

$T_{1f,t}$ is the simulated T_{1f} at the t^{th} time step.

The variance index measures how far the simulated indoor air temperatures spread out from their averages. Large values of $var(\cdot)$ signify significant temperature fluctuations.

The calculated variances of simulated T_{2f}, T_{1f}, and T_{0b} are listed in Table 4.5 for each scenario. It is noticed (scenarios 1~7) that for every zone the rise of temperature variance by $\downarrow C_i$, $\uparrow R_{im}$, or $\uparrow C_m$ is larger than the reduction of variance by $\uparrow C_i$, $\downarrow R_{im}$, or $\downarrow C_m$. For example, $var(T_{1f})$ is raised by 38% (from 2.13 to 2.95) after decreasing C_i but reduced by only 27% (from 2.13 to 0.58) after increasing C_i. This phenomenon can be also observed in Figure 4.3 (a, b, & c), where the simulated T_{1f}'s under different scenarios are plotted for two days. It indicates that the house is designed with sufficient internal thermal mass, so that continuing to improve the thermal mass will become less beneficial.

In practice, alteration of C_i, R_{im}, or C_m is not independent. Changing internal thermal mass is often accomplished by changing the thickness and/or surface area of building fabric (concrete wall and slabs in particular), which often results in alterations of more than one parameter. Scenarios 1 & 8~11 are used to examine the combined effect of altering two or more parameters. They are proposed based on the results of Table 4.2 (the building fabric corresponds to the concrete slab in Figure 4.1): when the surface area of the building fabric increases, $\downarrow R_{im}$ and $\uparrow C_m$; when the thickness of the building fabric increases, $\uparrow C_i$, $\uparrow R_{im}$, and $\uparrow C_m$. As indicated in Table 4.5 and Figure 4.3 (d & e), increasing the thickness or surface area causes less change in the variance of indoor air temperatures than decreasing the thickness or surface area. This suggests that both the thickness and surface area of building fabric are designed to be sufficient that further increasing them is not necessary, and non-redundant that further decreasing them leads to considerate raise in the variance of indoor air temperatures.

Table 4.5 Temperature variance and heating energy usage calculated based on simulation of the simplified model structure under different scenarios

Scenarios	$var(T_{2f})$	$var(T_{1f})$	$var(T_{0b})$	$\sum_{t=t_1}^{t_2} Q_{hp}\,\Delta t\ (kWh)$
Reference ($C_i/R_{im}/C_m$)	1.48	2.13	0.58	876
$\uparrow C_i$	1.05	1.55	0.39	886
$\downarrow C_i$	2.01	2.95	0.91	850
$\uparrow R_{im}$	2.33	3.20	0.85	846
$\downarrow R_{im}$	1.15	1.59	0.43	897
$\uparrow C_m$	1.26	2.04	0.54	862
$\downarrow C_m$	2.08	2.37	0.66	882
$\uparrow C_i, \uparrow R_{im}, \uparrow C_m$	1.53	2.16	0.51	853
$\downarrow C_i, \downarrow R_{im}, \downarrow C_m$	2.63	2.63	0.74	876
$\downarrow R_{im}, \uparrow C_m$	0.88	1.47	0.39	890
$\uparrow R_{im}, \downarrow C_m$	2.69	3.31	0.92	857

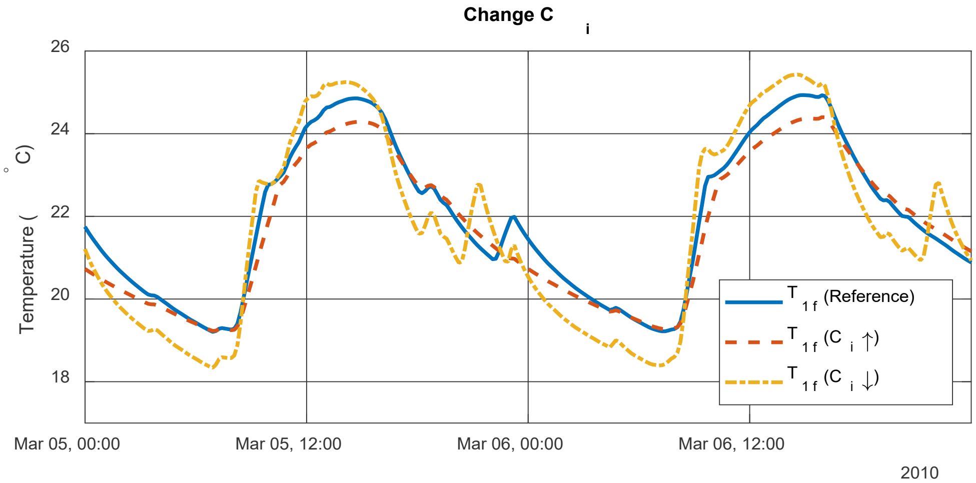
Although the studied house is well designed regarding its internal thermal mass, the mechanisms of changing C_i, R_{im}, and C_m's values to stabilize indoor air temperatures is still worth exploring in order to inform future designs.

C_i (C_{i2}, C_{i1}, or C_{i0}) represents the thermal capacity of indoor air, furniture, and some surface layers of building fabric. A larger value of C_i makes it harder to vary the indoor air temperature, hence reduces fluctuations of the indoor air temperature. However, strategies for increasing the value of C_i can be hard to implement: increasing the amount of indoor air is impractical; adding furniture or surface layer components (e.g., partition walls) may not lead to a significant change of C_i but will sacrifice the living space.

R_{im} (R_4, R_{11}, R_{18}, or R_{20}) is the thermal resistance between indoor air and inner layers of building fabric. A smaller value of R_{im} makes it easier to charge and discharge the inner layers of building fabric. Decreasing its value allows more energy to be stored at daytime and released at nighttime, and thus stabilizes the indoor air temperature. To reduce R_{im}'s value, the best strategy is to increase the exposure of building fabric to its surrounding air such as using raw concrete walls and circulating indoor air within concrete slabs.

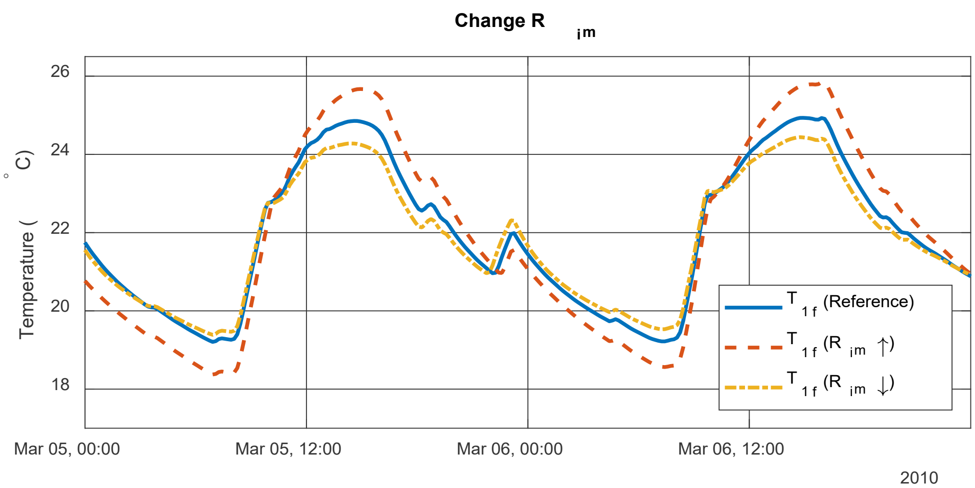
C_m (C_{m2}, C_{m1}, or C_{m0}) denotes the thermal capacity of inner layers of building fabric. Similar to C_i, large values of C_m makes it difficult to vary the temperature of building fabric's inner layers. Thus, an increased value of C_m will enlarge temperature difference between the indoor air and the building fabric, which results in more charging or discharging and more energy to be shifted from daytime to nighttime. Increasing its value typically involves integrating more thermal mass into the building fabric (e.g., more concrete).

(a)



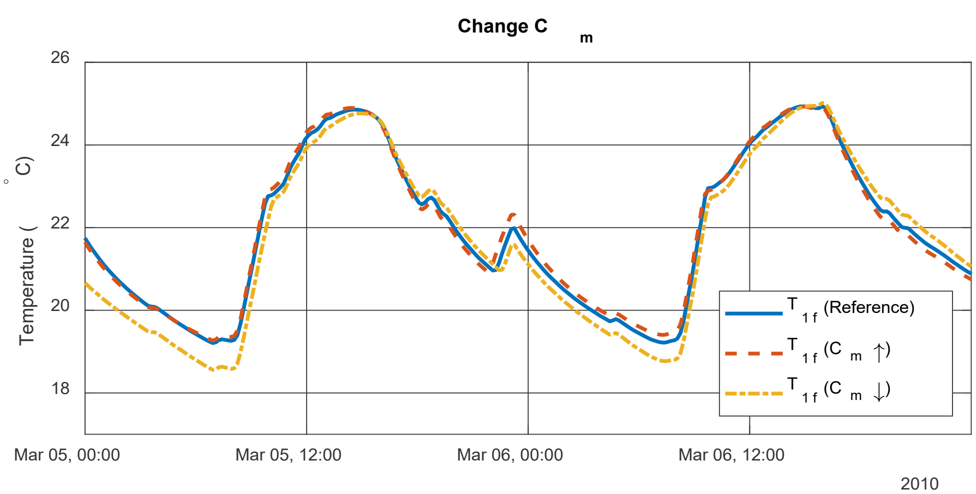
Change C_i
Temperature (°C)
26
24
22
20
18
Mar 05, 00:00
Mar 05, 12:00
Mar 06, 00:00
Mar 06, 12:00
2010
T_{1f} (Reference)
T_{1f} (C_i ↑)
T_{1f} (C_i ↓)

(b)



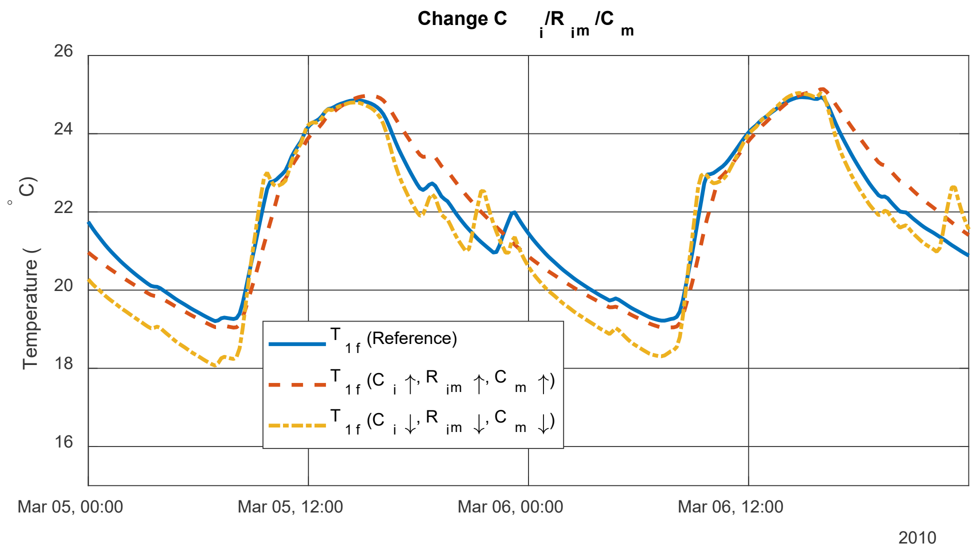
Change R_{im}
Temperature (°C)
26
24
22
20
18
Mar 05, 00:00
Mar 05, 12:00
Mar 06, 00:00
Mar 06, 12:00
2010
T_{1f} (Reference)
T_{1f} (R_{im} ↑)
T_{1f} (R_{im} ↓)

(c)



Change C_m
Temperature (°C)
26
24
22
20
18
Mar 05, 00:00
Mar 05, 12:00
Mar 06, 00:00
Mar 06, 12:00
2010
T_{1f} (Reference)
T_{1f} (C_m ↑)
T_{1f} (C_m ↓)

(d)



(e)

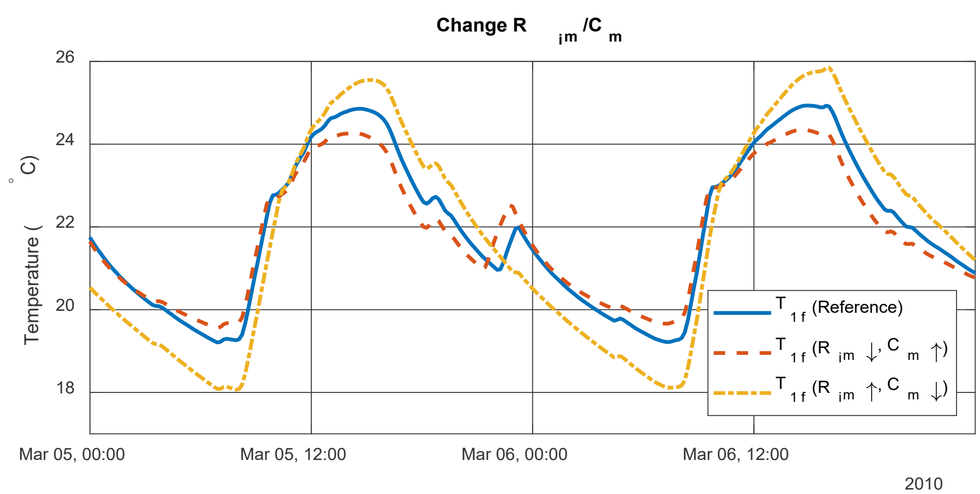


Figure 4.3 Simulated zone temperature (T_{1f}) for two days under different scenarios: (a) change C_i; (b) change R_{im}; (c) change C_m; (d) change C_i, R_{im}, and C_m; (e) change R_{im} and C_m. Increase is ↑ and decrease is ↓.

At last, changing values of C_i, R_{im}, or C_m does not induce a great change of heating energy usage. As shown in Table 4.5, the estimated heating energy for each scenario lies within $870 \pm 30\ kWh$. This is because C_i, R_{im}, and C_m, parameters of internal thermal mass, mainly affect the fluctuation of indoor air temperatures rather than heat losses to the outdoor or the ground. Since other sources of heat gains are never changed, the energy for heating should not, in a significant way, depend on the values of such parameters.

4.5. Discussion

The data-driven RC model has demonstrated a strong ability of physical interpretation and thermal performance evaluation of buildings.

Typically, the RC models have two types of parameters for interpretation: steady-state parameters (e.g., R_{io}'s and F_s's) and dynamic parameters (e.g., C_i's, R_{Im}'s, and C_m's). Steady-state parameters can be identified even when there are no dynamics (i.e., all system states are time-invariant constants). Dynamic parameters can only be inferred when the training data contains sufficient difference among various inputs and outputs. Physical meanings of dynamic parameters are relatively obscure. For example, the boundary between the surface layers and inner layers of building fabric is difficult to define. Even though, the basic interpretations of dynamic parameters are already enough to perform analyses on the internal thermal mass.

In thermal performance evaluation, an RC model can approximate the net energy storage in building fabric as well as the shifted energy from day to night by charging and discharging. This advantage of RC models allows them to provide useful information for building operation to save energy use and avoid overheating during sunny days. Also, energy shifting can be regarded as a way of inner layers of the building fabric to constrain the fluctuations of indoor air temperatures. RC models can be used to investigate such constraining effects by altering their parameter values. In particular, the thickness of surface area of building fabric (concrete wall/slabs) can be evaluated regarding their adequacy and effectiveness. With a suitable RC model, such evaluation approaches apply to buildings in general.

4.6. Summary

First, parameters in RC models are thoroughly interpreted through the design of a simple single-zone room and training of a low-order RC network. Parameter estimates in the RC network are related to different design configurations. Then, a three-zone RC model of a low energy house is adopted to evaluate the actual thermal performance of the house. By performing energy balance analysis, the relative importance of different energy flow paths is identified; energy storage by inner layers of the building fabric is assessed. By altering the estimated parameter values and simulating the three-zone RC model under different scenarios, parameters of internal thermal mass are investigated concerning their capability of reducing the fluctuations of indoor air temperatures. In conclusion, the data-driven RC models are powerful in deriving important thermal properties, evaluating the actual thermal performance, and inform future designs of buildings.

Chapter 5. Conclusion

This chapter summarizes the research content, marks the major contributions, and provides recommendations for future work.

5.1. Research Summary

The goal of this research is to identify suitable data-driven models for thermal dynamic analysis of buildings. In pursuing this goal, Chapter 2 critically reviews the data-driven approach including data-driven models (i.e., RC, TF, and AI models) that are the most commonly employed within building applications along with the corresponding model training and selection techniques. Based on the results of the literature review, Chapter 3 proposes a simple yet effective methodology to acquire RC model structures that are reliable for physical interpretation. It is shown that the obtained model is readily physically interpretable and has a favourable learning ability from on-site measurements. Chapter 4 applies the obtained RC model to evaluating the thermal performance of a low energy house. The estimated model parameters are associated with design configurations. By adjusting the parameter values, the influences of changing the current design of internal thermal mass on reducing temperature fluctuations are investigated. Important thermal properties (e.g., the equivalent thermal transmittance of the house) are derived. The relative importance of different energy flow paths is identified and compared. In conclusion, the obtained RC model proves to be powerful at both characterizing and assessing the thermal behaviours of buildings.

5.2. Major Contributions

This research presents a comprehensive study of data-driven modelling of and its application in building thermal systems. In this context, the major contributions of this thesis include:

i. Conducting a critical review of the data-driven approach for modelling building thermal dynamics. This includes summarizing typical input/output variables used in data-driven modelling, comparing different types of data-driven models by their formulation, prediction accuracy, and physical interpretation, and examining some model training and selection techniques.

ii. Proposing a methodology for developing RC models that are reliable for physical interpretation and suitable for thermal dynamic analysis of buildings. This involves integrating the genetic algorithm into model training and detecting non-identifiable parameters for removal. The proposed methodology is effective, efficient, and easy to implement.

iii. Applying the developed RC model to evaluating the design and actual thermal performance of occupied buildings. The model parameters are related to design configurations through an in-depth physical interpretation. The model's ability to evaluate building designs is justified through an investigation of the internal thermal mass, an energy balance analysis, and inferring important thermal properties.

The proposed model development methodology is verified with the development of a multi-zone model with on-site measurements from a low energy house. The thermal performance evaluation is performed on the same house using the same dataset.

5.3. Future Work

Current research focuses on developing data-driven models for building thermal systems. This work can be extended into the following directions:

i. Design software or toolboxes that allow for easy collection of data and automation of the model development process.
ii. Develop time-variant models to consider time-varying effects and investigate the potential improvement of model performance.
iii. Incorporate data-driven models into conventional building operating systems for model predictive control.
iv. Apply data-driven models with online training to fault detection and diagnosis of complex mechanical systems.

References

[1] International Energy Agency, Transition to sustainable buildings: strategies and opportunities to 2050, International Energy Agency, Paris, 2013.

[2] L. Pérez-Lombard, J. Ortiz, C. Pout, A review on buildings energy consumption information, Energy & Buildings. 40 (2008) 394-398.

[3] T.A. Reddy, Applied data analysis and modeling for energy engineers and scientists, Springer, New York, 2011.

[4] S. Prívara, J. Cigler, Z. Váňa, F. Oldewurtel, C. Sagerschnig, E. Žáčeková, Building modeling as a crucial part for building predictive control, Energy & Buildings. 56 (2013) 8-22.

[5] Y. Zhu, Multivariable system identification for process control, 1. ed. ed., Pergamon, Amsterdam [u.a.], 2001.

[6] Y. Ma, Model Predictive Control for Energy Efficient Buildings (2013).

[7] K. Lee, J.E. Braun, Model-based demand-limiting control of building thermal mass, Building and Environment. 43 (2008) 1633-1646.

[8] M. Sourbron, C. Verhelst, L. Helsen, Building models for model predictive control of office buildings with concrete core activation, Journal of Building Performance Simulation. 6 (2013) 175-198.

[9] W.Y. Lee, et al, Fault detection in an air-handling unit using residual and recursive parameter identification methods, Winter meeting of American Society of Heating, Refrigeration and Air Conditioning Engineers, Atlanta, GA (United States), 17-21 Feb 1996 (1996).

[10] J. Wang, H. Wang, Y. Chen, C.W.H. Chan, J. Qin, Online model-based fault detection and diagnosis strategy for VAV air handling units, Energy & Buildings. 55 (2012) 252-263.

[11] M. Yalcintas, Energy-savings predictions for building-equipment retrofits, Energy & Buildings. 40 (2008) 2111-2120.

[12] ASHRAE, 2017 ASHRAE Handbook Fundamentals (2017).

[13] R.E. Hedegaard, S. Petersen, Evaluation of Grey-Box Model Parameter Estimates Intended for Thermal Characterization of Buildings, Energy Procedia. 132 (2017) 982-987.

[14] K. Gaspar, M. Casals, M. Gangolells, A comparison of standardized calculation methods for in situ measurements of façades U-value, Energy & Buildings. 130 (2016) 592-599.

[15] I. Hazyuk, C. Ghiaus, D. Penhouet, Optimal temperature control of intermittently heated buildings using Model Predictive Control: Part I – Building modeling, Building and Environment. 51 (2012) 379-387.

[16] I. Hazyuk, C. Ghiaus, D. Penhouet, Optimal temperature control of intermittently heated buildings using Model Predictive Control: Part II – Control algorithm, Building and Environment. 51 (2012) 388-394.

[17] O. Mejri, E. Palomo Del Barrio, N. Ghrab-Morcos, Energy performance assessment of occupied buildings using model identification techniques, Energy & Buildings. 43 (2011) 285-299.

[18] W.J. Cole, K.M. Powell, E.T. Hale, T.F. Edgar, Reduced-order residential home modeling for model predictive control, Energy & Buildings. 74 (2014) 69-77.

[19] C. Ghiaus, A. Chicinas, C. Inard, Grey-box identification of air-handling unit elements, Control Engineering Practice. 15 (2007) 421-433.

[20] S. Prívara, Z. Váňa, E. Žáčeková, J. Cigler, Building modeling: Selection of the most appropriate model for predictive control, Energy and Buildings. 55 (2012) 341-350.

[21] J. Hu, P. Karava, A State-Space Modeling Approach and Subspace Identification Method for Predictive Control of Multi-Zone Buildings with Mixed-Mode Cooling, Proceedings of International High Performance Buildings Conference (2014).

[22] R. Sonderegger, Diagnostic tests determining the thermal response of a house, ASHRAE meeting, Atlanta, GA, USA, 29 Jan 1978 (1977).

[23] H. Madsen, J. Holst, Estimation of continuous-time models for the heat dynamics of a building, Energy & Buildings. 22 (1995) 67-79.

[24] A. Saberi Derakhtenjani, J.A. Candanedo, Y. Chen, V.R. Dehkordi, A.K. Athienitis, Modeling approaches for the characterization of building thermal dynamics and model-based control: A case study, Science and Technology for the Built Environment. 21 (2015) 824-836.

[25] J.M. Penman, Second order system identification in the thermal response of a working school, Building and Environment. 25 (1990) 105-110.

[26] D.A. Coley, J.M. Penman, Second order system identification in the thermal response of real buildings. Paper II: Recursive formulation for on-line building energy management and control, Building and Environment. 27 (1992) 269-277.

[27] J. Date, A.K. Athienitis, M. Fournier, A Study of Temperature Set Point Strategies for Peak Power Reduction in Residential Buildings, Energy Procedia. 78 (2015) 2130-2135.

[28] S.F. Fux, A. Ashouri, M.J. Benz, L. Guzzella, EKF based self-adaptive thermal model for a passive house, Energy & Buildings. 68 (2014) 811-817.

[29] P. Andersen, M. Jiménez, H. Madsen, C. Rode, Characterization of heat dynamics of an arctic low-energy house with floor heating, Build. Simul. 7 (2014) 595-614.

[30] Z. Liao, A.L. Dexter, A simplified physical model for estimating the average air temperature in multi-zone heating systems, Building and Environment. 39 (2004) 1013-1022.

[31] R. Bălan, J. Cooper, K. Chao, S. Stan, R. Donca, Parameter identification and model based predictive control of temperature inside a house, Energy & Buildings. 43 (2011) 748-758.

[32] J.E. Braun, N. Chaturvedi, An Inverse Gray-Box Model for Transient Building Load Prediction, HVAC&R Research. 8 (2002) 73-99.

[33] Q. Zhou, S. Wang, X. Xu, F. Xiao, A grey-box model of next-day building thermal load prediction for energy-efficient control, International Journal of Energy Research. 32 (2008) 1418-1431.

[34] S. Wang, X. Xu, Simplified building model for transient thermal performance estimation using GA-based parameter identification, International Journal of Thermal Sciences. 45 (2006) 419-432.

[35] J. Date, et al, Control-oriented Modelling of Thermal Zones in a House: a Multi-level Approach, International High Performance Buildings Conference (2016).

[36] P. Biddulph, V. Gori, C.A. Elwell, C. Scott, C. Rye, R. Lowe, T. Oreszczyn, Inferring the thermal resistance and effective thermal mass of a wall using frequent temperature and heat flux measurements, Energy & Buildings. 78 (2014) 10-16.

[37] P.H. Baker, H.A.L. van Dijk, PASLINK and dynamic outdoor testing of building components, Building and Environment. 43 (2008) 143-151.

[38] M.J. Jiménez, B. Porcar, M.R. Heras, Estimation of building component UA and gA from outdoor tests in warm and moderate weather conditions, Solar Energy. 82 (2008) 573-587.

[39] J.E. Seem, S.A. Klein, W.A. Beckman, J.W. Mitchell, Transfer Functions for Efficient Calculation of Multidimensional Transient Heat Transfer, Journal of Heat Transfer. 111 (1989) 5.

[40] J. Široký, F. Oldewurtel, J. Cigler, S. Prívara, Experimental analysis of model predictive control for an energy efficient building heating system, Applied Energy. 88 (2011) 3079-3087.

[41] H. Harb, N. Boyanov, L. Hernandez, R. Streblow, D. Müller, Development and validation of grey-box models for forecasting the thermal response of occupied buildings, Energy & Buildings. 117 (2016) 199-207.

[42] P. Bacher, H. Madsen, Identifying suitable models for the heat dynamics of buildings, Energy & Buildings. 43 (2011) 1511-1522.

[43] G.F. Franklin, J.D. Powell, M.L. Workman, Digital control of dynamic systems, Addison Wesley, Menlo Park, Calif., 1998.

[44] H. Kwakernaak, R. Sivan, Linear optimal control systems, Wiley, New York <>, 1972.

[45] S. Prívara, J. Široký, L. Ferkl, J. Cigler, Model predictive control of a building heating system: The first experience, Energy & Buildings. 43 (2011) 564-572.

[46] K.J. Keesman, System identification: an introduction, Springer, London ;New York, 2011.

[47] G. Mustafaraj, J. Chen, G. Lowry, Development of room temperature and relative humidity linear parametric models for an open office using BMS data, Energy Build. 42 (2010) 348-356.

[48] M.J. Jiménez, H. Madsen, K.K. Andersen, Identification of the main thermal characteristics of building components using MATLAB, Building and Environment. 43 (2008) 170-180.

[49] W.J.N. Turner, A. Staino, B. Basu, Residential HVAC fault detection using a system identification approach, Energy & Buildings. 151 (2017) 1-17.

[50] P. Moroşan, R. Bourdais, D. Dumur, J. Buisson, Building temperature regulation using a distributed model predictive control, Energy & Buildings. 42 (2010) 1445-1452.

[51] P.R. Armstrong, S.B. Leeb, L.K. Norford, Control with Building Mass -- Part I: Thermal Response Model, ASHRAE Trans. 112 (2006) 449-461.

[52] P.R. Armstrong, S.B. Leeb, L.K. Norford, Control with Building Mass-Part II: Simulation, ASHRAE Transactions. 112 (2006) 462.

[53] P.R. Armstrong, Model identification with application to building control and fault detection (2004).

[54] T.Y. Chen, Application of adaptive predictive control to a floor heating system with a large thermal lag, Energy & Buildings. 34 (2002) 45-51.

[55] T.Y. Chen, Real-time predictive supervisory operation of building thermal systems with thermal mass, Energy & Buildings. 33 (2001) 141-150.

[56] R.Z. Freire, G.H.C. Oliveira, N. Mendes, Predictive controllers for thermal comfort optimization and energy savings, Energy & Buildings. 40 (2008) 1353-1365.

[57] J.E. Seem, Modeling of Heat Transfer in Buildings (1987).

[58] T. McKelvey, Identification of state-space models from time and frequency data, Univ Linköping, Linköping, 1995.

[59] J.A. Candanedo, V.R. Dehkordi, M. Stylianou, Model-based predictive control of an ice storage device in a building cooling system, Applied Energy. 111 (2013) 1032.

[60] S. Royer, et al, Black-box modeling of buildings thermal behavior using system identification, Proceedings of the 19th IFAC World Congress (2014).

[61] G.J. Ríos-Moreno, M. Trejo-Perea, R. Castañeda-Miranda, V.M. Hernández-Guzmán, G. Herrera-Ruiz, Modelling temperature in intelligent buildings by means of autoregressive models, Automation in Construction. 16 (2007) 713-722.

[62] G. Mustafaraj, G. Lowry, J. Chen, Prediction of room temperature and relative humidity by autoregressive linear and nonlinear neural network models for an open office, Energy & Buildings. 43 (2011) 1452-1460.

[63] K. Yun, R. Luck, P.J. Mago, H. Cho, Building hourly thermal load prediction using an indexed ARX model, Energy & Buildings. 54 (2012) 225.

[64] H.A. Nielsen, H. Madsen, Modelling the heat consumption in district heating systems using a grey-box approach, Energy Build. 38 (2006) 63-71.

[65] H.A. Nielsen, H. Madsen, Predicting the Heat Consumption in District Heating Systems using Meteorological Forecasts (2000).

[66] P. Bacher, H. Madsen, H.A. Nielsen, B. Perers, Short-term heat load forecasting for single family houses, Energy & Buildings. 65 (2013) 101.

[67] K. Subbarao, Building parameters and their estimation from performance monitoring (1985).

[68] A. Rabl, Parameter Estimation in Buildings: Methods for Dynamic Analysis of Measured Energy Use, Journal of Solar Energy Engineering. 110 (1988) 52.

[69] T. Chen, A methodology for thermal analysis and predictive control of building envelope heating systems (1997).

[70] T.Y. Chen, A.K. Athienitis, Investigation of practical issues in building thermal parameter estimation, Building and Environment. 38 (2003) 1027-1038.

[71] J.F. Kreider, Predicting hourly building energy use: The great energy predictor shootout -- Overview and discussion of results, 1994 American Society of Heating, Refrigerating, and Air

Conditioning Engineers (ASHRAE) annual meeting, Orlando, FL (United States), 25-29 Jun 1994 (1994).

[72] J.S. Haberl, S. Thamilseran, Great energy predictor shootout II: Measuring retrofit savings -- overview and discussion of results, 1996 annual meeting of the American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE), Inc., San Antonio, TX (United States), 22-26 Jun 1996 (1996).

[73] M.Q. Raza, A. Khosravi, A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings, Renewable and Sustainable Energy Reviews. 50 (2015) 1352-1372.

[74] C. Deb, et al, A review on time series forecasting techniques for building energy consumption, Renewable and Sustainable Energy Reviews. 74 (2017) 902-924.

[75] M.A. Mat Daut, M.Y. Hassan, H. Abdullah, H.A. Rahman, M.P. Abdullah, F. Hussin, Building electrical energy consumption forecasting analysis using conventional and artificial intelligence methods: A review, Renewable and Sustainable Energy Reviews (2016).

[76] G. Dudek, Artificial Immune System for Short-Term Electric Load Forecasting, in: Artificial Intelligence and Soft Computing – ICAISC 2008, Springer Berlin Heidelberg, Berlin, Heidelberg, 2008, pp. 1007-1017.

[77] G. Dudek, Short-Term Load Forecasting Using Random Forests (2015) 821-828.

[78] J.F. Kreider, D.E. Claridge, P. Curtiss, R. Dodier, J.S. Haberl, M. Krarti, Building Energy Use Prediction and System Identification Using Recurrent Neural Networks, Journal of Solar Energy Engineering. 117 (1995) 161.

[79] M. Kawashima, et al, Hourly thermal load prediction for the next 24 hours by ARIMA, EWMA, LR and an artificial neural network, American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE) winter meeting and exhibition, Chicago, IL (United States), 28 Jan - 1 Feb 1995 (1995).

[80] S. Karatasou, M. Santamouris, V. Geros, Modeling and predicting building's energy use with artificial neural networks: Methods and results, Energy & Buildings. 38 (2006) 949-958.

[81] R. Yokoyama, T. Wakui, R. Satake, Prediction of energy demands using neural network with model identification by global optimization, Energy Conversion and Management. 50 (2009) 319-327.

[82] P.A. González, J.M. Zamarreño, Prediction of hourly energy consumption in buildings based on a feedback artificial neural network, Energy & Buildings. 37 (2005) 595-601.

[83] P.A. González Lanza, J.M. Zamarreño Cosme, A short-term temperature forecaster based on a state space neural network, Engineering Applications of Artificial Intelligence. 15 (2002) 459-464.

[84] S.A. Kalogirou, M. Bojic, Artificial neural networks for the prediction of the energy consumption of a passive solar building, Energy. 25 (2000) 479-491.

[85] G. Zhang, B. Eddy Patuwo, M. Y. Hu, Forecasting with artificial neural networks:: The state of the art, International Journal of Forecasting. 14 (1998) 35-62.

[86] J. Yang, H. Rivard, R. Zmeureanu, On-line building energy prediction using adaptive artificial neural networks, Energy & Buildings. 37 (2005) 1250-1259.

[87] R.H. Dodier, G.P. Henze, Statistical Analysis of Neural Networks as Applied to Building Energy Prediction, Journal of Solar Energy Engineering. 126 (2004) 592.

[88] A.E. Ruano, E.M. Crispim, E.Z.E. Conceição, M.M.J.R. Lúcio, Prediction of building's temperature using neural networks models, Energy & Buildings. 38 (2006) 682-694.

[89] P.M. Ferreira, A.E. Ruano, S. Silva, E.Z.E. Conceição, Neural networks based predictive control for thermal comfort and energy savings in public buildings, Energy & Buildings. 55 (2012) 238-251.

[90] Weijie Mai, et al, Electric load forecasting for large office building based on radial basis function neural network, PESGM (2014) 1-5.

[91] Q. Li, Q. Meng, J. Cai, H. Yoshino, A. Mochida, Predicting hourly cooling load in the building: A comparison of support vector machine and different artificial neural networks, Energy Conversion and Management. 50 (2009) 90-96.

[92] D.F. Specht, A general regression neural network, IEEE Transactions on Neural Networks. 2 (1991) 568-576.

[93] W. Lee, J.M. House, N. Kyong, Subsystem level fault diagnosis of a building's air-handling unit using general regression neural networks, Applied Energy. 77 (2004) 153-170.

[94] J.W. Moon, S.K. Jung, Y. Kim, S. Han, Comparative study of artificial intelligence-based building thermal control methods – Application of fuzzy, adaptive neuro-fuzzy inference system, and artificial neural network, Applied Thermal Engineering. 31 (2011) 2422-2429.

[95] Z. Du, B. Fan, J. Chi, X. Jin, Sensor fault detection and its efficiency analysis in air handling unit using the combined neural networks, Energy & Buildings. 72 (2014) 157-166.

[96] M. Mohandes, Support vector machines for short-term electrical load forecasting, International Journal of Energy Research. 26 (2002) 335-345.

[97] Q. Li, Q. Meng, J. Cai, H. Yoshino, A. Mochida, Applying support vector machine to predict hourly cooling load in the building, Applied Energy. 86 (2009) 2249-2256.

[98] F. Zhang, C. Deb, S.E. Lee, J. Yang, K.W. Shah, Time series forecasting for building energy consumption using weighted Support Vector Regression with differential evolution optimization technique, Energy & Buildings. 126 (2016) 94-103.

[99] V.N. Vapnik, The nature of statistical learning theory, 2. ed. ed., Springer, New York [u.a.], 2000.

[100] J. Che, J. Wang, Short-term load forecasting using a kernel-based support vector regression combination model, Applied Energy. 132 (2014) 602-609.

[101] B. Dong, C. Cao, S.E. Lee, Applying support vector machines to predict building energy consumption in tropical region, Energy & Buildings. 37 (2005) 545-553.

[102] B. Schölkopf, A.J. Smola, Learning with kernels, MIT Press, Cambridge, Mass. [u.a.], 2002.

[103] Y. Yao, Z. Lian, S. Liu, Z. Hou, Hourly cooling load prediction by a combined forecasting model based on Analytic Hierarchy Process, International Journal of Thermal Sciences. 43 (2004) 1107-1118.

[104] C. Fan, F. Xiao, S. Wang, Development of prediction models for next-day building energy consumption and peak power demand using data mining techniques, Applied Energy. 127 (2014) 1-10.

[105] T. Senjyu, P. Mandal, K. Uezato, T. Funabashi, Next day load curve forecasting using hybrid correction method, TPWRS. 20 (2005) 102-109.

[106] D.K. Chaturvedi, A.P. Sinha, O.P. Malik, Short term load forecast using fuzzy logic and wavelet transform integrated generalized neural network, International Journal of Electrical Power and Energy Systems. 67 (2015) 230-237.

[107] M. Rana, I. Koprinska, Forecasting electricity load with advanced wavelet neural networks, Neurocomputing. 182 (2016) 118-132.

[108] P. Fazenda, P. Lima, P. Carreira, Context-based thermodynamic modeling of buildings spaces, Energy & Buildings. 124 (2016) 164-177.

[109] G. Reynders, J. Diriken, D. Saelens, Quality of grey-box models and identified parameters as function of the accuracy of input and observation signals, Energy & Buildings. 82 (2014) 263-274.

[110] F. Oldewurtel, D. Sturzenegger, M. Morari, Importance of occupancy information for building climate control, Applied Energy. 101 (2013) 521-532.

[111] Yashen Lin, et al, Issues in identification of control-oriented thermal models of zones in multi-zone buildings, CDC (2012) 6932-6937.

[112] A. Deconinck, S. Roels, Is stochastic grey-box modelling suited for physical properties estimation of building components from on-site measurements?, Journal of Building Physics. 40 (2017) 444-471.

[113] L. Ljung, System identification, Prentice Hall, Upper Saddle River, N.J., 1999.

[114] R.B. Gopaluni, R.S. Patwardhan, S.L. Shah, MPC relevant identification—tuning the noise model, Journal of Process Control. 14 (2004) 699-714.

[115] R.B. Gopaluni, R.S. Patwardhan, S.L. Shah, Bias distribution in MPC relevant identification, IFAC Proceedings Volumes. 35 (2002) 435-440.

[116] J. Rehor, V. Havlena, Grey-box model identification – control relevant approach, IFAC Proceedings Volumes. 43 (2010) 117-122.

[117] J. Zhao, Y. Zhu, R. Patwardhan, Identification of k-step-ahead prediction error model and MPC control, Journal of Process Control. 24 (2014) 48-56.

[118] S. Prívara, J. Cigler, Z. Váňa, F. Oldewurtel, E. Žáčeková, Use of partial least squares within the control relevant identification for buildings, Control Engineering Practice. 21 (2013) 113-121.

[119] T. Söderström, P. Stoica, Some properties of the output error method, Automatica. 18 (1982) 93-99.

[120] Y. Tomita, A.A.H. Damen, Hof, Paul M. J. Van Den, Equation error versus output error methods, Ergonomics. 35 (1992) 551-564.

[121] D. Eckhard, A.S. Bazanella, C.R. Rojas, H. Hjalmarsson, Cost function shaping of the output error criterion, Automatica. 76 (2017) 53-60.

[122] N.R. Kristensen, H. Madsen, S.B. Jørgensen, Parameter estimation in stochastic grey-box models, Automatica. 40 (2004) 225-237.

[123] K.J. Åström, Maximum likelihood and prediction error methods, Automatica. 16 (1980) 551-574.

[124] T. Soderstrom, H. Fan, B. Carlsson, S. Bigi, Least squares parameter estimation of continuous-time ARX models from discrete-time data, TAC. 42 (1997) 659-673.

[125] L. Ljung, Some Classical and Some New Ideas for Identification of Linear Systems, J Control Autom Electr Syst. 24 (2013) 3-10.

[126] S.J. Qin, An overview of subspace identification, Computers and Chemical Engineering. 30 (2006) 1502-1513.

[127] P. Van Overschee, B. De Moor, Subspace Identification for Linear Systems : Theory - Implementation - Applications, Springer, Boston, 1996.

[128] I.D. Landau, R. Lozano, M. M'Saad, A. Karimi, Adaptive Control: Algorithms, Analysis and Applications, Springer, London, 2011.

[129] T. Söderström, L. Ljung, I. Gustavsson, A theoretical analysis of recursive identification methods, Automatica. 14 (1978) 231-244.

[130] L. Ljung, Analysis of a general recursive prediction error identification algorithm, Automatica. 17 (1981) 89-99.

[131] D. Antonucci, et al, Building performance evaluation through a novel feature selection algorithm for automated arx model identification procedures, Energy and Buildings. 150 (2017) 432-446.

[132] L.M. Candanedo, V. Feldheim, D. Deramaix, Data driven prediction models of energy use of appliances in a low-energy house, Energy and Buildings (2017).

[133] I. Naveros, C. Ghiaus, Order selection of thermal models by frequency analysis of measurements for building energy efficiency estimation, Applied Energy. 139 (2015) 230-244.

[134] L. Ljung, System Identification Toolbox™ User's Guide, MathWorks, Natick, MA, 2017.

[135] M. Hudson Beale, M. T. Hagan, H. B. Demuth, Neural Network Toolbox™ User's Guide, MathWorks, 2017.

[136] R. Juhl, et al, ctsmr - Continuous Time Stochastic Modeling in R (2016).

[137] P.C. Young, et al, The CAPTAIN Toolbox for Matlab (2009).

[138] H. Garnier, et al, The CONTSID toolbox for Matlab: extensions and latest developments, 15th IFAC Symposium on System Identification, SYSID 2009 (2009).

[139] L. Ljung, P. E. Caines, Asymptotic normality of prediction error estimators for approximate system models, 1978 IEEE Conference on Decision and Control including the 17th Symposium on Adaptive Processes (1978) 927-932.

[140] K.K. Andersen, H. Madsen, L.H. Hansen, Modelling the heat dynamics of a building using stochastic differential equations, Energy & Buildings. 31 (2000) 13-24.

[141] A. Fonti, G. Comodi, S. Pizzuti, A. Arteconi, L. Helsen, Low Order Grey-box Models for Short-term Thermal Behavior Prediction in Buildings, Energy Procedia. 105 (2017) 2107-2112.

[142] A. Raue, C. Kreutz, T. Maiwald, J. Bachmann, M. Schilling, U. Klingmüller, J. Timmer, Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood, Bioinformatics. 25 (2009) 1923-1929.

[143] A. Raue, C. Kreutz, T. Maiwald, U. Klingmuller, J. Timmer, Addressing parameter identifiability by model-based experimentation, IET systems biology. 5 (2011) 120-130.

[144] H. Madsen, et al, Thermal Performance Characterization using Time Series Data - IEA EBC Annex 58 Guidelines (2015).

[145] S. Vajda, H. Rabitz, E. Walter, Y. Lecourtier, Qualitative and Quantitative Identifiability Analysis of Nonlinear Chemical Kinetic Models, Chemical Engineering Communications. 83 (1989) 191-219.

[146] T. Quaiser, M. Mönnigmann, Systematic identifiability testing for unambiguous mechanistic modeling – application to JAK-STAT, MAP kinase, and NF-κB signaling pathway models, BMC Systems Biology. 3 (2009) 50.

[147] J. Guedj, R. Thiébaut, D. Commenges, Practical Identifiability of HIV Dynamics Models, Bull. Math. Biol. 69 (2007) 2493-2513.

[148] L. Ljung, P.E. Caines, Asymptotic Normality of Prediction Error Estimators for Approximate System Models, Stochastics. 3 (1980) 29.

[149] Jay M Ver Hoef, Who Invented the Delta Method?, The American Statistician. 66 (2012) 124.

[150] Y. Chen, A.K. Athienitis, K. Galal, Modeling, design and thermal performance of a BIPV/T system thermally coupled with a ventilated concrete slab in a low energy solar house: Part 1, BIPV/T system and house energy concept, Solar Energy. 84 (2010) 1892-1907.

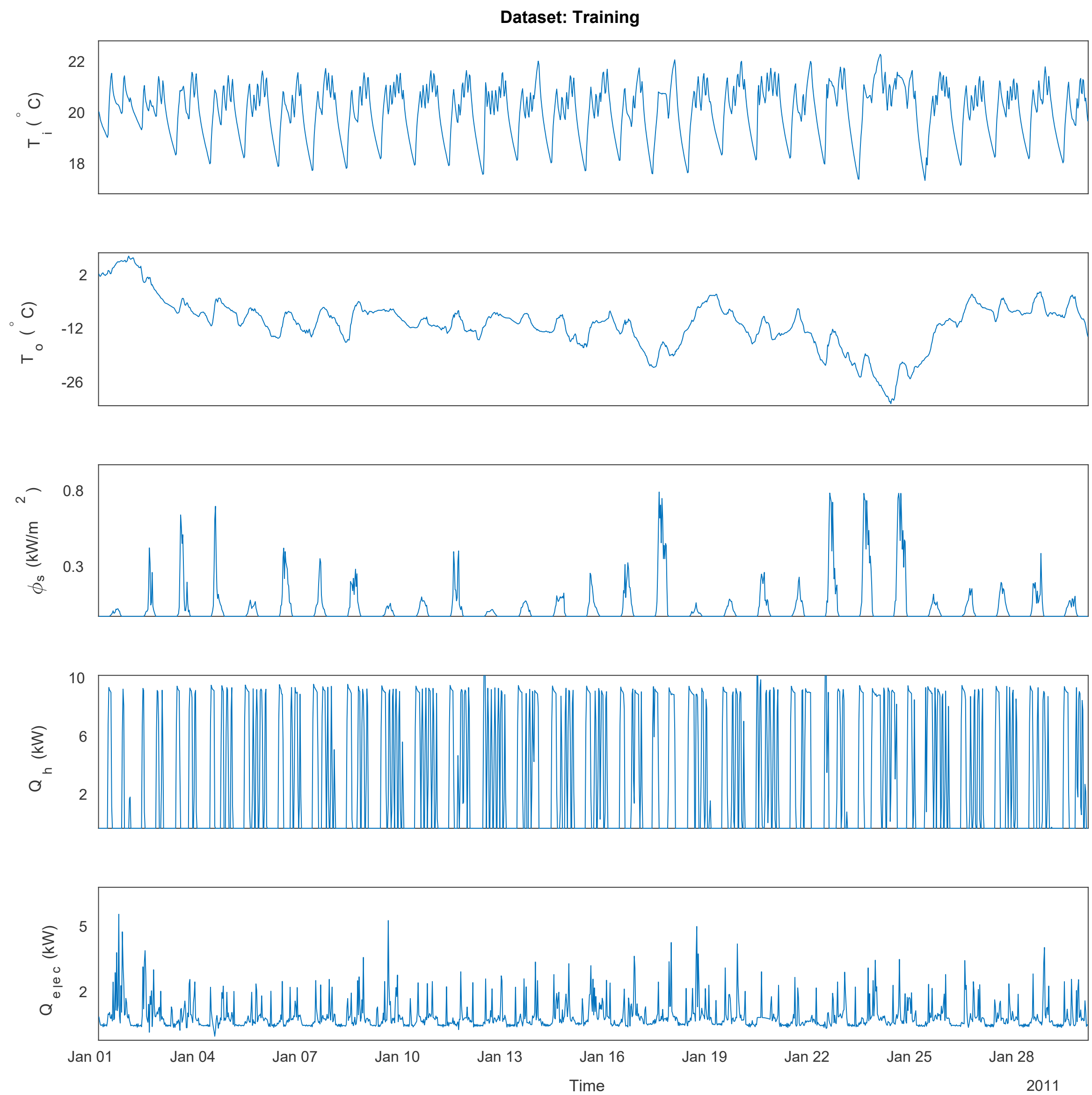
[151] Y. Chen, K. Galal, A.K. Athienitis, Modeling, design and thermal performance of a BIPV/T system thermally coupled with a ventilated concrete slab in a low energy solar house: Part 2, ventilated concrete slab, Solar Energy. 84 (2010) 1908-1919.

[152] M.A. Doiron, Whole-Building Energy Analysis and Lessons Learned for a Near Net-Zero Energy Solar House (2011).

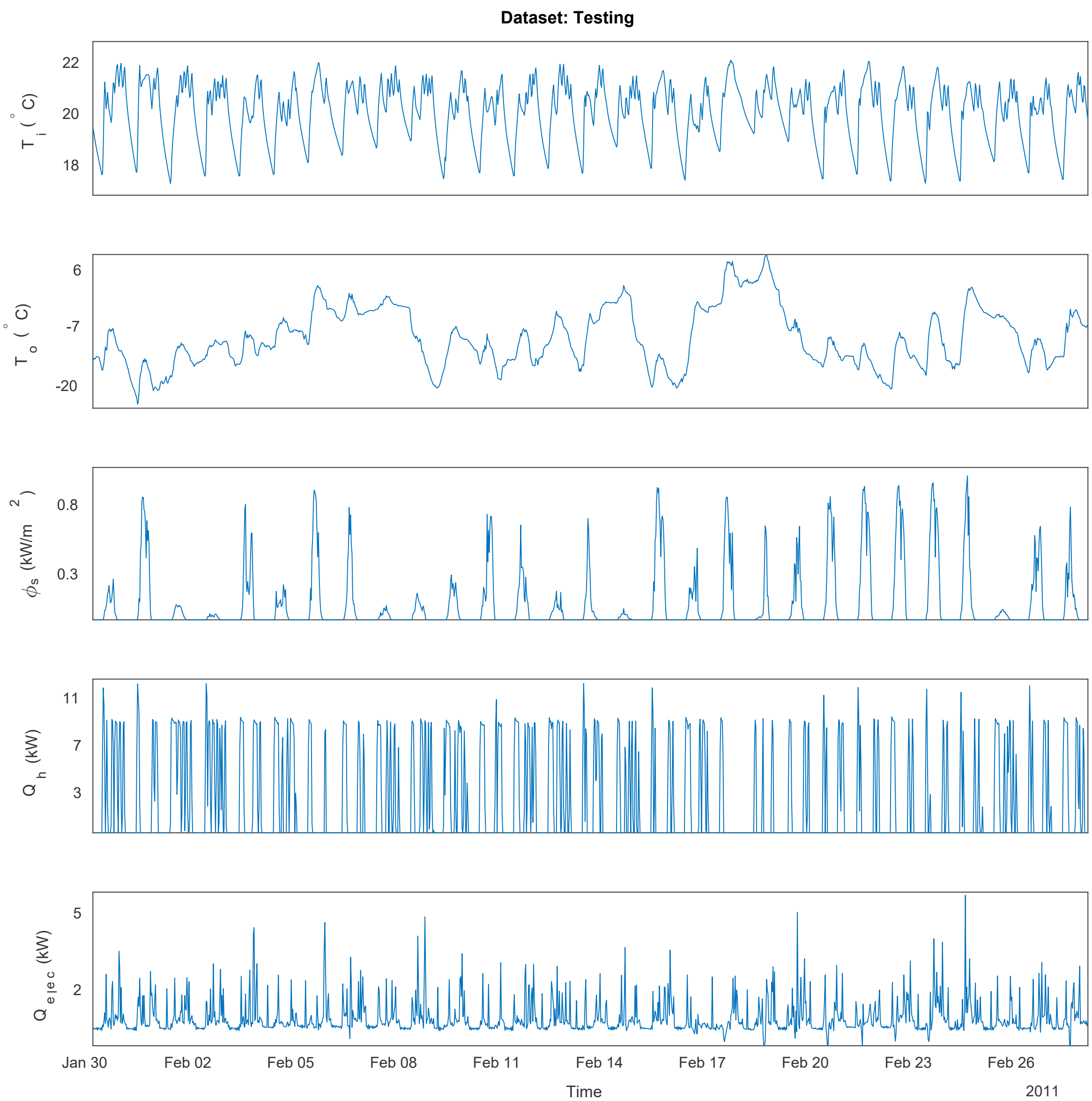
[153] National Energy Code of Canada for Buildings 2011 (2011).

Appendix A. Training/Testing Datasets

A.1. Datasets for the case study in Section 2.4



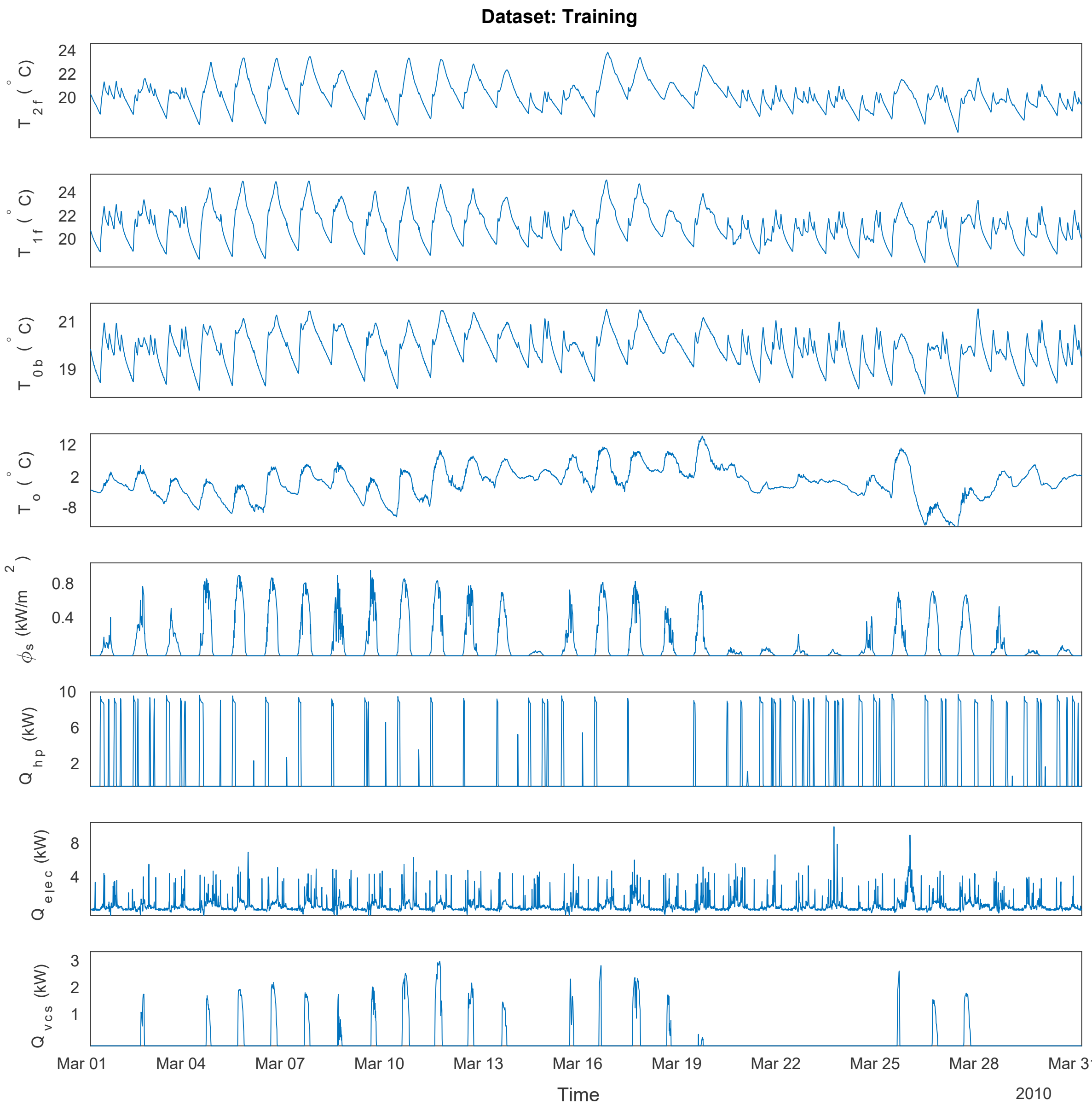
(a)



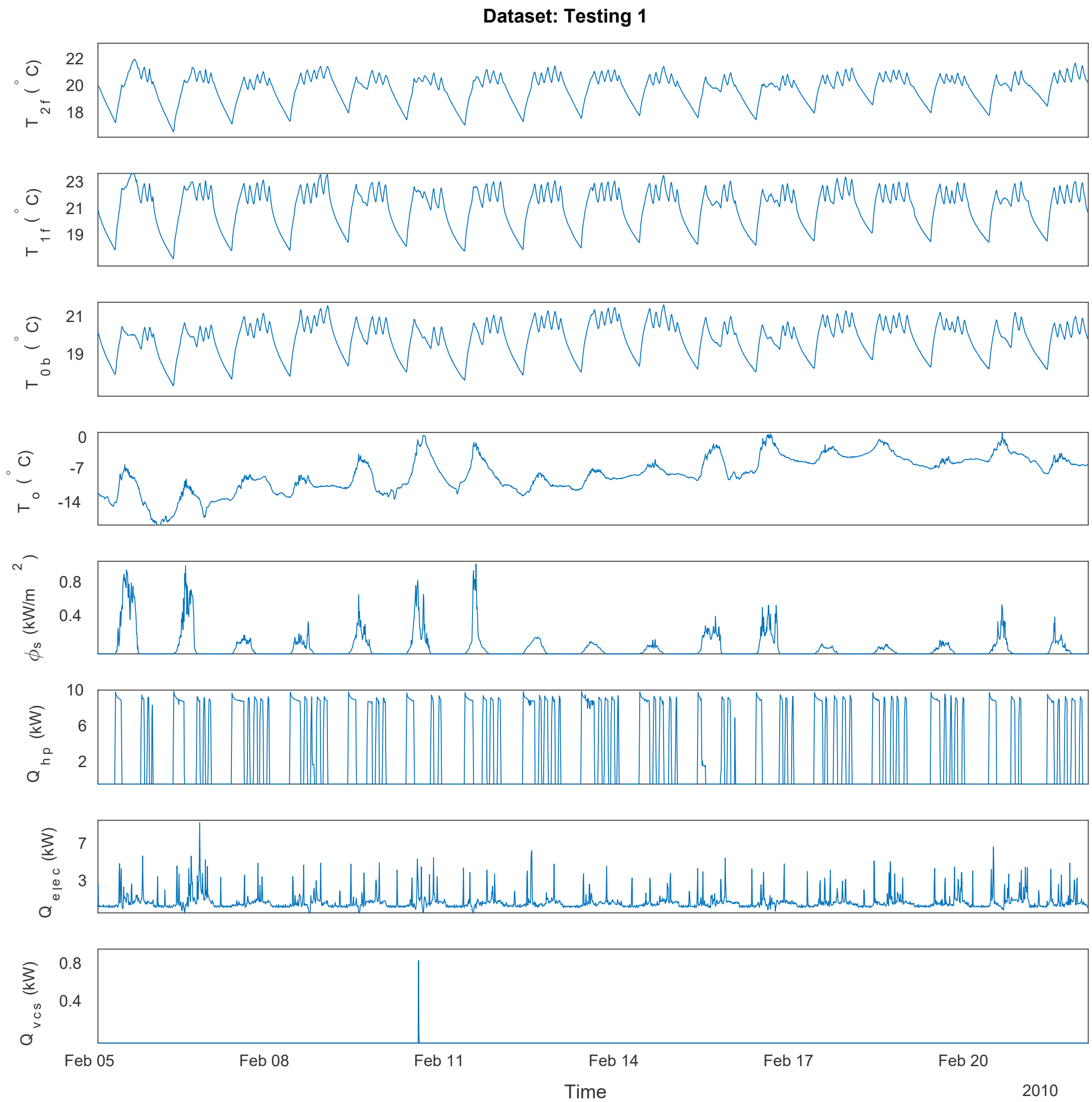
(b)

Figure A.1 Datasets for the case study in Section 2.4 "Case study": (a) training (30 days), and (b) testing (30 days). Sampling interval: 0.5 hour

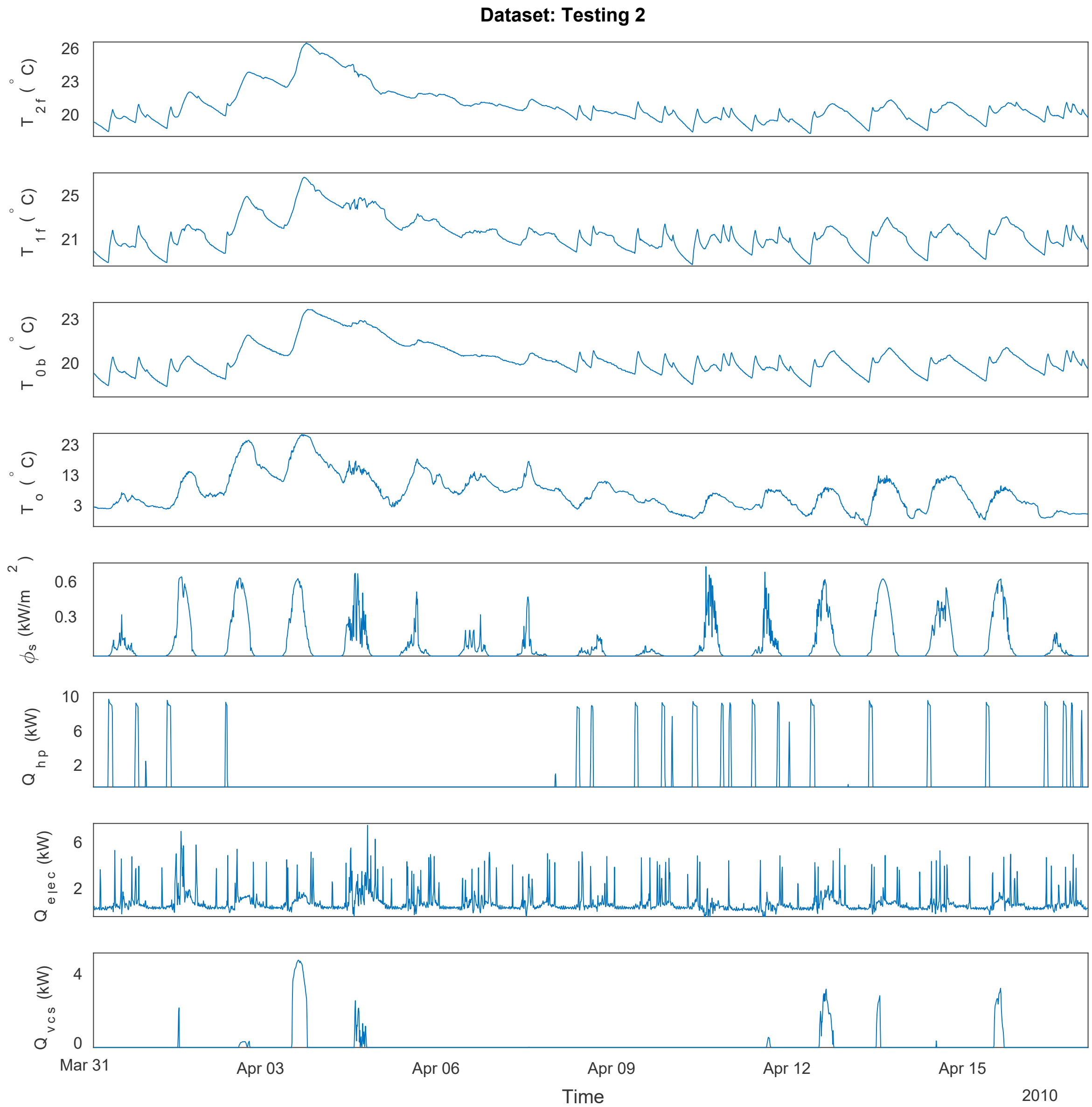
A.2. Datasets for the case study in Section 3.3



(a)


Dataset: Testing 1
T 2f (°C)
22
20
18
T 1f (°C)
23
21
19
T 0b (°C)
21
19
T o (°C)
0
-7
-14
ϕs (kW/m 2)
0.8
0.4
Q hp (kW)
10
6
2
Q ejec (kW)
7
3
Q vcs (kW)
0.8
0.4
Feb 05
Feb 08
Feb 11
Feb 14
Feb 17
Feb 20
Time
2010

(b)



(c)

Figure A.2 Dataset for the Case Study in Section 3.3 "Case Study": (a) training (31 days), (b) testing 1 (17 days), and (c) testing 2 (17 days). Sampling interval: 10 minutes.

Appendix B. Simplification Details

Table B.1 Simplification process of a three-zone RC model structure: the parameters to be removed in each round are highlighted. Model 1 is the complex preliminary model structure, and Model 10 is the simplified model structure.

Parameters	Model 1 Estimate	Standard Error	Ratio
R1	17.3	344.4	39.03
R2	14.6	332.1	44.67
R3	6.0	40.9	13.37
R4	1.9	0.4	0.44
R5	1.6	0.5	0.57
R6	21.2	1.6	0.15
R7	10.4	6.7	1.26
R8	17.0	18020.0	2077.00
R9	28.0	19152.0	1342.20
R10	26.5	4149.5	306.53
R11	13.3	10393.0	1527.50
R12	4.8	0.6	0.25
R13	21.4	1.8	0.16
R14	19.5	6415.7	645.41
R15	26.0	14135.0	1065.30
R16	14.1	14252.0	1982.50
R17	20.5	2422.7	231.65
R18	2.7	1.5	1.10
R19	14.7	691.8	92.43
R20	32.8	1.3	0.08
R21	14.9	18.6	2.45
R22	17.4	2923.0	330.18
C1	8.7	24.9	5.62
C2	2.0	0.1	0.07
C3	19.7	48.3	4.80
C4	3.4	0.4	0.23
C5	26.7	10014.0	735.49
C6	1.9	0.0	0.04
C7	34.1	6609.2	380.04
C8	27.6	12811.0	910.81
C9	15.3	7169.8	918.20
C10	2.8	0.0	0.02
C11	20.1	1494.5	145.55
C12	9.8	5.5	1.11
C13	7.8	643.0	162.16
F1	0.4	12.4	62.01
F2	2.1	0.1	0.12
F3	0.0	3.3	Inf
F4	2.4	0.5	0.37
F5	0.0	1348.3	Inf
F6	3.9	0.1	0.06
F7	12.9	120.2	18.19
F8	0.0	683.1	Inf
F9	1.3	1026.4	1504.30
F10	1.2	0.0	0.07
F11	4.1	264.5	126.16
F12	0.7	0.9	2.51
p1	0.0	0.0	Inf
p2	0.2	0.0	0.21
p3	0.0	0.0	0.68

Threshold Inf

Objective Function 4.05E-05

Parameters	Model 2 Estimate	Standard Error	Ratio
R1	21.4	62.4	5.72
R2	27.2	91.4	6.59
R3	11.7	18.5	3.09
R4	2.2	0.2	0.17
R5	15.9	11.9	1.46
R6	22.6	0.8	0.07
R7	2.9	0.3	0.20
R8	18.3	139.0	14.89
R9	29.5	254.9	16.91
R10	35.8	464.2	25.44
R11	12.3	11.9	1.90
R12	6.2	0.7	0.21
R13	18.1	0.6	0.06
R14	23.8	46.6	3.84
R15	18.0	251.8	27.48
R16	19.6	282.7	28.24
R17	35.9	761.5	41.58
R18	2.5	0.3	0.22
R19	21.9	204.9	18.33
R20	29.6	0.6	0.04
R21	14.6	5.3	0.71
R22	9.6	297.8	60.92
C1	26.4	138.0	10.24
C2	1.7	0.0	0.05
C3	10.8	34.3	6.24
C4	3.3	0.2	0.10
C5	16.8	46.5	5.42
C6	1.8	0.0	0.04
C7	34.2	457.3	26.24
C8	4.8	6.3	2.59
C9	15.2	14.2	1.83
C10	2.8	0.0	0.01
C11	31.6	559.4	34.68
C12	6.9	0.8	0.24
C13	8.4	179.8	42.11
F1	4.6	17.8	7.60
F2	2.5	0.1	0.06
F4	1.0	0.2	0.45
F6	3.7	0.1	0.04
F7	12.5	32.8	5.12
F9	2.7	37.0	26.43
F10	1.3	0.0	0.03
F11	10.2	46.2	8.84
F12	0.0	0.2	Inf
p2	0.2	0.0	0.18
p3	0.0	0.0	0.41

Threshold Inf

Objective Function 2.41E-05

Parameters	Model 3 Estimate	Standard Error	Ratio
R1	25.4	160.7	12.39
R2	35.5	165.4	9.12
R3	10.3	26.8	5.07
R4	2.8	0.2	0.16
R5	30.5	32.9	2.12
R6	24.9	0.9	0.07
R7	2.2	0.1	0.13
R8	30.5	2123.4	136.65
R9	39.0	1727.0	86.90
R10	36.9	2788.8	148.25
R11	43.3	711.4	32.22
R12	5.1	0.5	0.19
R13	18.0	0.6	0.07
R14	9.9	200.0	39.77
R15	20.2	171.3	16.61
R16	12.2	158.6	25.49
R17	33.1	37069.0	2195.70
R18	2.7	0.2	0.18
R19	35.9	45593.0	2487.60
R20	32.5	0.6	0.03
R21	40.8	39.9	1.92
R22	33.8	16237.0	941.95
C1	9.0	19.6	4.25
C2	1.8	0.0	0.05
C3	18.8	42.6	4.45
C4	3.6	0.2	0.11
C5	36.6	1711.0	91.60
C6	1.7	0.0	0.04
C7	41.0	2459.6	117.59
C8	30.1	814.7	53.11
C9	11.5	53.2	9.04
C10	2.8	0.0	0.01
C11	48.4	54343.0	2200.10
C12	6.4	0.7	0.21
C13	36.8	52474.0	2795.40
F1	2.8	12.5	8.90
F2	2.5	0.1	0.06
F4	0.9	0.3	0.61
F6	3.8	0.1	0.04
F7	12.6	314.6	48.87
F9	0.5	9.3	34.39
F10	1.2	0.0	0.05
F11	25.7	347.0	26.41
p2	0.2	0.0	0.18
p3	0.0	0.0	0.67

Threshold 2.00E+03

R22 is removed since C13 is removed

F11 is removed since C11 is removed

Objective Function 2.57E-05

Parameters	Model 4 Estimate	Standard Error	Ratio
R1	9.1	26.5	5.73
R2	49.5	37.9	1.50
R3	29.7	286.7	18.89
R4	2.8	0.2	0.15
R5	77.9	187.5	4.72
R6	24.4	0.8	0.07
R7	2.2	0.1	0.12
R8	25.8	494.3	37.59
R9	41.9	850.5	39.76
R10	82.3	4628.9	110.27
R11	31.6	139.5	8.66
R12	4.2	0.1	0.07
R13	16.9	0.6	0.06
R14	10.8	74.7	13.52
R15	10.1	64.6	12.48
R16	6.6	62.4	18.42
R18	2.5	0.1	0.08
R20	32.8	0.5	0.03
R21	108.4	106.2	1.92
C1	21.0	57.9	5.42
C2	1.8	0.0	0.05
C3	7.1	66.3	18.43
C4	3.6	0.2	0.08
C5	35.6	784.1	43.16
C6	1.7	0.0	0.03
C7	77.8	5888.1	148.29
C8	48.0	241.7	9.87
C9	123.4	409.6	6.51
C10	2.8	0.0	0.01
C12	8.4	0.4	0.10
F1	2.5	2.1	1.67
F2	2.5	0.1	0.06
F4	1.1	0.3	0.44
F6	3.8	0.1	0.04
F7	38.9	997.4	50.20
F9	21.0	199.0	18.55
F10	1.2	0.0	0.04
p2	0.2	0.0	0.16
p3	0.0	0.0	0.49

Threshold 100

F7 is removed since C7 is removed

Objective Function 2.83E-05

Parameters	Model 5 Estimate	Standard Error	Ratio
R1	1.2	0.1	0.14
R2	37.0	12.4	0.65
R3	3.8	1.3	0.65
R4	56.6	291.2	10.08
R5	0.9	0.0	0.07
R6	19.2	2.8	0.28
R7	156.9	913.6	11.41
R8	59.5	289.8	9.54
R9	84.0	1961.7	45.77
R11	96.7	801.3	16.24
R12	4.4	0.2	0.08
R13	29.0	1.8	0.12
R14	6.8	7.1	2.03
R15	27.7	132.4	9.38
R16	5.4	136.0	49.71
R18	2.4	0.1	0.11
R20	31.0	0.5	0.03
R21	162.0	397.0	4.80
C1	4.1	0.2	0.11
C2	2.1	0.1	0.07
C3	23.3	16.2	1.36
C4	87.3	866.2	19.45
C5	59.3	681.6	22.51
C6	1.7	0.0	0.03
C8	90.8	90.6	1.96
C9	48.3	986.9	40.08
C10	2.8	0.0	0.01
C12	8.5	0.6	0.13
F1	2.3	0.3	0.29
F2	2.2	0.1	0.13
F4	16.5	145.7	17.34
F6	3.8	0.1	0.03
F9	19.3	488.5	49.60
F10	1.2	0.0	0.04
p2	0.2	0.0	0.22
p3	0.0	0.0	0.19

Threshold 40

R15 is removed since C9 is removed

Objective Function 3.22E-05

Parameters	Model 6 Estimate	Standard Error	Ratio
R1	3575.7	748000.0	410.14
R2	464.1	31000000.0	131000.00
R3	7.2	2.9	0.81
R4	2.2	0.2	0.14
R5	2573.1	189000.0	144.02
R6	21.4	0.6	0.06
R7	2.2	0.1	0.12
R8	1770.8	56755.0	62.82
R11	7.0	0.9	0.24
R12	6.3	0.3	0.10
R13	18.3	0.5	0.05
R14	4.2	0.4	0.17
R18	3.0	0.2	0.14
R20	25.8	0.2	0.02
R21	13.0	0.4	0.07
C1	147.7	8790000.0	117000.00
C2	1.8	0.0	0.05
C3	9.4	6.2	1.29
C4	3.2	0.1	0.09
C5	1.5	47.9	62.67
C6	1.7	0.0	0.03
C8	17.2	1.1	0.13
C10	2.8	0.0	0.01
C12	6.5	0.5	0.15
F1	547.0	32400000.0	116000.00
F2	2.7	0.1	0.06
F4	0.4	0.2	0.96
F6	3.9	0.1	0.04
F10	1.3	0.0	0.03
p2	0.2	0.0	0.13
p3	0.1	0.0	0.12

Threshold 1.00E+05

R1 is removed since C1 is removed

Objective Function 3.63E-05

Parameters	Model 7 Estimate	Standard Error	Ratio
R3	9.0	0.7	0.15
R4	1.8	0.2	0.18
R5	4.1	0.9	0.44
R6	21.2	0.6	0.06
R7	91.5	443.3	9.50
R8	1.5	0.1	0.16
R11	4.1	0.2	0.11
R12	5.6	0.2	0.07
R13	20.0	0.4	0.04
R14	60.7	29.0	0.94
R18	1.7	0.0	0.04
R20	26.2	0.2	0.01
R21	14.6	0.3	0.04
C2	1.8	0.0	0.05
C3	8.5	1.4	0.31
C4	1.8	0.2	0.23
C5	1.2	0.1	0.18
C6	1.5	0.0	0.05
C8	11.0	1.2	0.21
C10	2.9	0.0	0.01
C12	14.0	0.3	0.04
F2	2.6	0.1	0.06
F4	0.2	0.1	1.36
F6	4.3	0.0	0.02
F10	1.3	0.0	0.03
p2	0.4	0.0	0.06
p3	0.1	0.0	0.08

Threshold 1

Objective Function 4.08E-05

Parameters	Model 8 Estimate	Standard Error	Ratio
R3	6.5	0.7	0.21
R4	0.8	0.1	0.13
R5	1.0	0.0	0.08
R6	12.0	0.3	0.05
R8	31969.0	25100000.0	1538.90
R11	7.3	1.2	0.31
R12	6.6	0.3	0.10
R13	48.9	4.9	0.20
R14	4.3	0.5	0.22
R18	3.1	0.3	0.16
R20	25.0	0.2	0.01
R21	13.8	0.5	0.07
C2	1.8	0.1	0.08
C3	8.3	1.9	0.44
C4	3.3	0.2	0.09
C5	5703.7	121000000.0	41659.00
C6	1.8	0.0	0.03
C8	14.6	1.4	0.18
C10	2.9	0.0	0.01
C12	6.9	0.6	0.18
F2	3.0	0.1	0.06
F6	3.9	0.1	0.03
F10	1.3	0.0	0.04
p2	0.3	0.0	0.16
p3	0.1	0.0	0.11

Threshold 1.00E+03

Objective Function 4.25E-05

Parameters	Model 9 Estimate	Standard Error	Ratio
R3	8.7	1.3	0.29
R4	0.9	0.1	0.13
R5	1.2	0.1	0.10
R6	13.2	0.4	0.05
R11	4.2	0.3	0.13
R12	7.4	0.3	0.09
R13	41.2	3.3	0.16
R14	8.4	1.1	0.25
R18	2.2	0.1	0.07
R20	24.6	0.2	0.01
R21	13.5	0.3	0.05
C2	1.8	0.1	0.08
C3	5.6	1.3	0.44
C4	3.1	0.2	0.11
C6	1.9	0.0	0.03
C8	16.8	1.4	0.16
C10	2.9	0.0	0.01
C12	10.1	0.4	0.07
F2	2.9	0.1	0.05
F6	4.1	0.1	0.03
F10	1.3	0.0	0.03
p2	0.4	0.0	0.06
p3	0.1	0.0	0.06

Threshold 0.4

R3 is removed since C3 is removed

Objective Function 4.75E-05

Parameters	Model 10 Estimate	Standard Error	Ratio
R4	1.0	0.1	0.12
R5	1.3	0.1	0.11
R6	14.4	0.4	0.06
R11	3.4	0.2	0.10
R12	7.1	0.3	0.07
R13	33.7	2.3	0.13
R14	10.5	1.1	0.21
R18	2.1	0.1	0.05
R20	24.7	0.2	0.01
R21	13.9	0.3	0.04
C2	1.8	0.1	0.07
C4	3.3	0.1	0.08
C6	2.0	0.0	0.03
C8	19.5	2.0	0.20
C10	2.9	0.0	0.01
C12	10.0	0.3	0.06
F2	2.6	0.0	0.03
F6	4.2	0.1	0.03
F10	1.3	0.0	0.02
p2	0.4	0.0	0.06
p3	0.1	0.0	0.06

All parameters are identifiable

Objective Function 5.03E-05

Note:

$\mu = 10E - 5$, $\delta = 0.05$

*Ratio = 1.96*Standard Error/Estimate*

$[C1; C2; C3; C4; C5; C6; C7; C8; C9; C10; C11; C12; C13] =$

$[C_{e2}; C_{i2}; C_{n2}; C_{m2}; C_{e1}; C_{i1}; C_{n1}; C_{m1}; C_{e0}; C_{i0}; C_{n0}; C_{m0}; C_{mm0}]$

Parameter units: $R(°\text{C}/kW)$, $C(kWh/°\text{C})$, $F(m^2)$