Black Box Modelling and Simulating the Dynamic Indoor Air Temperature of a Laboratory Using the Continuous-Time Transfer Function Model

Shamsul Faisal Mohd Hussein¹ Noor Bazila Sharifmuddin² Mohd. Fitri Alif Mohd. Kasai³ Abdulqader Omar Al-Rabeei¹ Amrul Faruq¹,⁴ Siti Munirah Zulkapli¹ Noorazizi Mohd Samsuddin⁵ Sheikh Ahmad Zaki Shaikh Salim⁶ Shahrum Shah Abdullah¹,*

¹Biologically Inspired System and Technology (Bio-iST) iKohza, Department of Electronic Systems Engineering (ESE), Malaysia-Japan International Institute of Technology (MJJIT), Universiti Teknologi Malaysia (UTM) Kuala Lumpur, Jalan Sultan Yahya Petra (Jalan Semarak), 54100 Kuala Lumpur, MALAYSIA.
²Takasago Thermal/Environmental Systems (TTES) iKohza, Department of Mechanical Precision Engineering (MPE), Malaysia-Japan International Institute of Technology (MJJIT), Universiti Teknologi Malaysia (UTM) Kuala Lumpur, Jalan Sultan Yahya Petra (Jalan Semarak), 54100 Kuala Lumpur, MALAYSIA.
³Department of Control and Mechatronics Engineering (CMED), School of Electrical Engineering, Faculty of Engineering, Universiti Teknologi Malaysia (UTM) Kuala Lumpur, Jalan Sultan Yahya Petra (Jalan Semarak), 54100 Kuala Lumpur, MALAYSIA.
⁴Department of Electrical Engineering, Faculty of Engineering, Universitas Muhammadiyah Malang. Jl. Raya Tlogomas 246, 65144 Malang, INDONESIA.
⁵Engineering Department, Razak Faculty of Technology and Informatics, Universiti Teknologi Malaysia (UTM) Kuala Lumpur, Jalan Sultan Yahya Petra (Jalan Semarak), 54100 Kuala Lumpur, MALAYSIA.
⁶Wind Engineering for (Urban, Artificial, Man-made) Environment Laboratory, Department of Mechanical Precision Engineering (MPE), Malaysia-Japan International Institute of Technology (MJJIT), Universiti Teknologi Malaysia (UTM) Kuala Lumpur, Jalan Sultan Yahya Petra (Jalan Semarak), 54100 Kuala Lumpur, MALAYSIA.
*Corresponding author. Email: shahrum@utm.my

ABSTRACT
The air conditioner is one of the devices that uses a high amount of electricity – more electricity consumption means more heat and greenhouse gases emitted to the environment if the electricity is generated by fossil fuel sources such as coal, diesel, natural gas etc. Energy-efficient control algorithms and strategies can be proposed to reduce the power consumption without sacrificing thermal comfort – time and cost can be saved by developing and testing these control algorithms and strategies via computer simulation instead of developing and testing them on the actual site, but this requires the availability of the mathematical model representing the dynamic behaviour of the system that is desired to be simulated. In this research, a black box model representing the dynamic indoor air temperature behaviour of the Industrial Instrumentation Laboratory at Malaysia-Japan International Institute of Technology (MJJIT), Universiti Teknologi Malaysia (UTM) Kuala Lumpur is developed based on the continuous-time transfer function model using the System Identification Toolbox™ in MATLAB® software to get a new model with a simpler model structure for a more practical simulation than the autoregressive–moving-average (ARMA) model developed in the previous research representing the same behaviour without sacrificing a significant level of accuracy. Both the continuous-time transfer function model and the ARMA model are developed based on the actual recorded data from the laboratory and minimal physical characteristics knowledge of the laboratory. The result shows that the optimised continuous-time transfer function model generated by the System Identification Toolbox™ in this research has a significantly simpler model structure than the optimised ARMA model developed in the previous research (standardised eight poles and two zeros for all inputs versus standardised 41 past inputs and outputs for all inputs and output), and only slightly less accurate in terms of percentage of fitting, %Fit value (82.60% and 80.17% versus 87.45% and 80.36% for training and testing data set).

Keywords: Modelling and simulation, black box modelling, building air temperature simulation, building air temperature prediction
1. INTRODUCTION

The air conditioner is one of the power-hungry devices, which contributes to the high usage of electricity – this leads to the high amount of harmful heat and greenhouse gases emitted to the atmosphere if the electricity is generated from the fossil fuel sources like coal, diesel, natural gas etc. Energy-efficient control algorithms and strategies can be designed to minimise the heat and greenhouse gases emission. These proposed control algorithms and strategies can be designed and optimized on the actual site, but doing so is time-consuming and costly. To solve this problem, mathematical models that represent the building’s dynamic hygrothermal (temperature and humidity) behaviour are constructed so that any proposed control algorithm and strategy can be quickly and cost-effectively designed and optimised via computer simulation. Furthermore, the developed mathematical model can also be used to build and optimize predictive controllers such as model predictive controller (MPC).

Just like other plants, the dynamic indoor air temperature behaviour of a building can also be represented using three types of models: (1) the white box model, which is also known as theoretical model; (2) the black box model, which is also known as empirical model; and (3) the grey box model, which is also known as semi-empirical model. The white box model of a system is built based on the physical knowledge of the system [1] – this knowledge is also known as the fundamental knowledge of science and engineering [2]. The black box model is developed by tuning the parameter(s) in linear or nonlinear ‘off-the-shelf’ equation(s) based on the plant’s input(s)-output(s) data to be modelled until the black box model’s output(s) nearly matches (as accurate as possible) the plant’s output(s). The grey box model combines some parts of the white box model and some parts of the black box model – the equation in the grey box model is based on the plant’s theoretical knowledge, but the unknown parameter in the grey box model is tuned using the actual input(s)-output(s) data recorded from the plant. Each model possesses its pro(s) and contra(s). The white box model can be simulated in a wider range of operating conditions, but the development of this model requires a high amount of time and financial resources if the system to be modelled has complex physical characteristic(s) [1]. In certain cases, the white box model is difficult to develop, as the parameter(s) in the model is/are not available, obtainable, or measurable during the model development [1]. The black box model is easier to construct, but it does not extrapolate much beyond the training data set, which is the data used to estimate the model [1]. The available training data set normally does not cover the entire operating condition of the plant to be modelled, so care must be taken when there is a need to simulate the constructed model outside the range of the training data set [1]. However, the black box model is reportedly popular in the industry [1]. Meanwhile, the grey box model provides a similar physical insight to the white box model because the grey box model is also built based on the fundamental knowledge of science and engineering, but the unknown parameter(s) in the grey box model is(are) estimated by using the input(s)-output(s) data of the modelled plant similar to the black box model. The grey box model can therefore be built more quickly and easily than the white box model and can be simulated in a wider range of operating conditions than the black box model [1].

Different researchers have developed different types of mathematical models to represent the dynamic indoor air temperature behaviour of a building. Hussein et al. developed a black box model based on autoregressive-moving-average (ARMA) model to simulate the dynamic indoor air temperature of one of the rooms in iHouse, a smart house testbed belongs to Japan Advanced Institute of Science and Technology (JAIST) in [3] with limited thermal physical knowledge of the smart house testbed. Then, Hussein et al. upgraded the model built in [3] to become a grey box model in [4] based on more thermal-related physical knowledge of the smart house testbed. After that, Hussein et al. developed a black box model based on ARMA model again, but this time to simulate the dynamic indoor air temperature of Industrial Instrumentation Laboratory at Malaysia-Japan International Institute of Technology (MJIIIT), Universiti Teknologi Malaysia (UTM) Kuala Lumpur in [5] also with limited thermal physical knowledge of laboratory – similar works are done by the main author in [3] and [4] are repeated in a different location in [5] so that it will be easier for the main author to expand his work in the future. Ooi et al. built their mathematical models in [6] to develop and optimise a conventional MPC through computer simulation to maintain the air temperature of one of the iHouse bedrooms – this work is then enhanced by Ooi et al. in [7] by using the adaptive MPC which uses two types of online model estimation system, the Kalman filter (KF) state estimator and the linear time-varying Kalman filter (LTVKF) estimator to increase the accuracy of the internal controller’s plant model. Radecki et al. in [8] proposed the implementation of a multimode unscented Kalman filter (UKF) as a generalisable online grey-box model based on the basic physical knowledge of the building combined with measured building data to estimate the dynamics of a multizone building and to identify the unknown time-varying thermal loads of that building – the general and scalable method of developing this control-oriented thermal model can be implemented on a large scale for cost-effective predictive controls.

Previously, the dynamic indoor air temperature behaviour of the Industrial Instrumentation Laboratory at Malaysia-Japan International Institute of Technology (MJIIIT), Universiti Teknologi Malaysia (UTM) Kuala Lumpur was modelled in [5] using an autoregressive–moving-average (ARMA) model, a form of linear black box model – the accuracy of the optimised ARMA model in the previous research is good, but not practical for implementation because its model structure requires many past inputs and output. This research focuses on building a new type of black box model representing the dynamic indoor air temperature behaviour of the same laboratory based on continuous-time transfer function model via MATLAB® System Identification Toolbox™ to obtain a new model with a simpler model structure for a more practical
simulation without significantly sacrificing accuracy. The continuous-time transfer function model developed in this research is compared with the ARMA model representing the same plant developed in the previous research for performance comparison in terms of model structure simplicity and accuracy.

2. RESEARCH METHOD

2.1 Scope of Research

First, the Industrial Instrumentation Laboratory at Malaysia-Japan International Institute of Technology (MJIIT), Universiti Teknologi Malaysia (UTM) Kuala Lumpur is selected as the plant to be modelled in this research and will be described in more detail in Subsection 2.2 (The Research Location). Second, only the weather-related inputs are used for the model developed in this research. The Industrial Instrumentation Laboratory is equipped with air conditioners and ventilation fans, but the data available by the time this research is done is recorded when these thermal comfort devices are not operated — newer data will be recorded while these thermal comfort devices are operated for the model’s future upgrade. Third, the number of pole(s) and zero(s) for the continuous-time transfer function models for all the inputs assigned in this research are standardised to simplify the parameters estimation process and shorten the model’s development time in this research.

2.2 The Research Location

The Industrial Instrumentation Laboratory at Malaysia-Japan International Institute of Technology (MJIIT), Universiti Teknologi Malaysia (UTM) Kuala Lumpur is selected as the plant to be modelled in this research. The laboratory is located at the 7th floor (based on the US-style) or the 6th floor (based on the UK-style) of the MJIIT building. It is also located at the front-side of the building (which is also the western side of the building) on the south wing.

The typical floor plan of the Industrial Instrumentation Laboratory is shown in Figure 1. Based on this figure, it is shown that the Industrial Instrumentation Laboratory is surrounded by the following spaces: (1) the staircase in the north-west; (2) the lift lobby in the north; (3) the corridor in the east; (4) the outdoor space in both the south and the west; (5) a classroom at exactly one level below the laboratory (on the 6th floor based on the US-style or the 5th floor based on the UK-style and is depicted separately in Figure 2); and (6) another classroom at exactly one level above the laboratory (on the 8th floor based on the US-style or the 7th floor based on the UK-style and is depicted separately in Figure 3).

![Figure 1](image_url)

Figure 1 The typical floor plan of the Industrial Instrumentation Laboratory at the 7th floor (based on the US-style) or the 6th floor (based on the UK-style) of the MJIIT building.
The typical floor plan of the classroom exactly below the Industrial Instrumentation Laboratory at the 6th floor (based on the US-style) or the 5th floor (based on the UK-style) of the MJIIT building.

2.3 The Data Recording Devices

Various types of data are recorded at the Industrial Instrumentation Laboratory and its surrounding spaces by using various types of recording devices. The recorded data is divided into two categories, which are the indoor and outdoor data.

The indoor data is recorded using home-made low-cost data loggers built from off-the-shelf sensors connected to single-board microcontrollers (such as Arduino) or single-board computers (such as Raspberry Pi). There are two types of data recorded by these indoor data loggers by the time this research is done, which are the indoor air temperature and the indoor relative humidity. These data loggers are installed in the Industrial Instrumentation Laboratory and the surrounding indoor spaces as mentioned in Subsection 2.2 (The Research Location), which are: (1) the laboratory itself; (2) the staircase in the north-west; (3) the lift lobby in the north; (3) the corridor in the east; (4) the classroom at exactly one level below the laboratory; and (5) the other classroom at exactly one level above the laboratory. The Industrial Instrumentation Laboratory itself is installed with eight sets of temperature and humidity sensors, as shown in Figure 1 [one set of sensors is installed at each of the eight corners of the laboratory (four lower corners at the floor and another four upper corners at the ceiling)] – the air temperature and humidity reading in this laboratory at any given time used in this research is the average temperature and humidity values from all these eight sets of sensors at that given time. By right, the same quantity of sensors should be installed at all the eight corners of each of the spaces surrounding the laboratory. These surrounding spaces however are public places that are accessible all the time unlike the Industrial Instrumentation Laboratory which is locked most of the time and only accessible by authorised persons – it is afraid that the sensors installed at the public places might be vandalised or stolen, so it is decided to reduce the number of sensors installed at each of these locations to only two sets and these sensors are installed at hard-to-notice and hard-to-reach points (which are the ceilings of these surrounding spaces) to reduce the risk of vandalism and theft, as shown in Figure 1, Figure 2, and Figure 3 (either one set on the north and the other set on the south of the spaces or one set on the east and the other set on the west of the spaces) – the air temperature and humidity reading in each of these surrounding spaces at any given time used in this research is the average temperature and humidity values from all these two sets of sensors at each space at that given time.

Meanwhile, the outdoor data is obtained from the readily available weather station belongs to the Wind Engineering for (Urban, Artificial, Man-made) Environment Laboratory, one of the laboratories in MJIIT. The weather station is located at the rooftop of the MJIIT building. There are various types of data recorded by the weather station, including: (1) the outdoor air temperature; (2) the outdoor relative humidity; (3) the wind speed and direction; and (4) the global solar radiation.

For this research, only the indoor and outdoor data related to air temperature and solar radiation is used. The types of solar radiation that are used as the inputs for this research are direct solar radiation and diffuse solar radiation, and their values are obtained by splitting the value of actual recorded global solar radiation into the values of direct solar radiation and diffuse solar radiation through calculation.

2.4 The Data Recording Period

All the data required for this research is recorded for 11 days, from the 5th of February 2019 until the 15th of February 2019 and this data is recorded at every one-minute interval. The recorded data is then divided into two parts, which are: (1) the first group (recorded for the first five days from the 5th of February 2019 until the 9th of February 2019); and (2) the second group (recorded for the remaining six days from the 10th of February 2019 until the 15th of February 2019). It is mentioned earlier in Section 1
(Introduction) that the black box model cannot extrapolate very well when it is simulated beyond the range of the training data set [1]. Based on the recorded data, it is observed that the output data for the second group has a wider maximum-minimum range and overlaps the maximum-minimum range of the output data for the first group. Therefore, it is decided in this research that the second group of the data is assigned as the training data set while the first group of the data is assigned as the testing data set. It is also decided in this research that the duration of the training data set is longer than the testing data set for this research – this is the reason why the number of days for the training data set is one day more than the testing data set (six days versus five days).

2.5 The Inputs-output Data Selection

The inputs-output data selected for the continuous-time transfer function model representing the dynamic indoor air temperature behaviour of the Industrial Instrumentation Laboratory developed in this research is based on the autoregressive-moving-average (ARMA) model representing the dynamic indoor air temperature behaviour of the same laboratory developed in the previous research in [5]. In the previous research, the types of input that are considered able to affect the dynamic indoor air temperature behaviour of the laboratory are listed for the development of the ARMA model, which are: (1) the past and present indoor air temperature of the laboratory itself, $T_{\text{IndLab}}$; (2) the past and present air temperature difference between the staircase in the north-west and the laboratory itself, $\Delta T_{\text{Stair-IndLab}}$; (3) the past and present air temperature difference between the lift lobby in the north and the laboratory itself, $\Delta T_{\text{LiftLobby-IndLab}}$; (4) the past and present air temperature difference between the corridor in the east and the laboratory itself, $\Delta T_{\text{Corridor-IndLab}}$; (5) the past and present air temperature difference between the outdoor space in both the south and the west and the laboratory itself, $\Delta T_{\text{Outdoor-IndLab}}$; (6) the past and present air temperature difference between the classroom at exactly one level below the laboratory and the laboratory itself, $\Delta T_{\text{LowerClass-IndLab}}$; (7) the past and present air temperature difference between the classroom at exactly one level above the laboratory and the laboratory itself, $\Delta T_{\text{UpperClass-IndLab}}$; (8) the past and present direct solar radiation that lands on the southern outer-wall of the laboratory, $I_{\text{D,OuterSouthWall}}$; (9) the past and present direct solar radiation that lands on the western outer-wall of the laboratory, $I_{\text{D,OuterWestWall}}$; and (10) the past and present diffuse solar radiation that lands on both the southern and western outer-walls of the laboratory, $I_{\text{DS,OuterWalls}}$. Meanwhile, only one output is considered for the model in this research, which is the future indoor air temperature of the laboratory itself, $T_{\text{IndLab}}$. The listed inputs and output lead to the construction of the multiple-input and single-output (MISO) model, which is depicted in Figure 4.

However, two problems are identified while estimating the new continuous-time transfer function model in this research based on the similar ARMA model inputs-output structure developed in the previous research, which are: (1) it takes a longer time for the MATLAB® System Identification Toolbox™ to estimate a model with many number of inputs; and (2) it is also difficult for the MATLAB® System Identification Toolbox™ to get a high-accuracy estimated model with many number of inputs (at least while using the data set that is used in this research and the previous research). Therefore, some of the inputs that are considered able to affect the indoor air temperature behaviour of the Industrial Instrumentation Laboratory in the previous research explained in the previous paragraph are revised and combined in this research as follows: (1) the past and present indoor air temperature of the laboratory itself, $T_{\text{IndLab}}$ (no change); (2) the past and present heat gain/loss due to the air temperature difference between the surrounding horizontal spaces and the laboratory itself, $\Delta T_{\text{Walls}}$ is introduced as one of the newly revised inputs and is the combination of the $\Delta T_{\text{Stair-IndLab}}$, $\Delta T_{\text{LiftLobby-IndLab}}$, $\Delta T_{\text{Corridor-IndLab}}$, and $\Delta T_{\text{Outdoor-IndLab}}$ explained in the previous paragraph; (3) the past and present heat gain/loss due to the air temperature difference between the vertical spaces and the laboratory itself, $\Delta T_{\text{Floors}}$ is introduced as one of the newly revised inputs and is the combination of the $\Delta T_{\text{LowerClass-IndLab}}$ and $\Delta T_{\text{UpperClass-IndLab}}$ explained in the previous paragraph; (4) the past and present heat gain due to the direct solar radiation that lands on the outer-wall of the laboratory, $I_{\text{D,OuterWall}}$ is introduced as one of the newly revised inputs and is the combination of the $I_{\text{D,OuterSouthWall}}$ and $I_{\text{D,OuterWestWall}}$ explained in the previous paragraph; and (5) the past and present heat gain due to the diffuse solar radiation that lands on both the outer-walls of the laboratory, $I_{\text{DS,OuterWalls}}$ (no change because the $I_{\text{DS,OuterWalls}}$ is the combination of past and present diffuse solar radiation that lands on the southern outer-wall, $I_{\text{DS,OuterSouthWalls}}$ and western outer-wall, $I_{\text{DS,OuterWestWalls}}$ – this combination has been done and used in the previous research in [5] because both the
southern and western outer-walls in the previous research were already assumed as a single piece of outer-wall due to the amount of diffuse solar radiation per unit area that lands on any vertical surface is the same regardless of which direction the vertical surface is facing). The revised listed inputs and output lead to the construction of the revised MISO model, which is depicted in Figure 5 – the combined inputs of the revised ARMA model are not straight away summed to become the newly introduced inputs, but are multiplied with the surface areas of where the inputs go through prior to the summation and are also depicted in Figure 5, and the surface areas for the corresponding inputs are summarised in Table 1. There are one wall-mounted ventilation fan on the northern wall, another one wall-mounted ventilation fan on the southern wall, two doors on the eastern wall, and windows on the western wall – these items have different heat transfer rate than the wall, but all the walls are assumed as a plain wall without ventilation fans, doors, and windows in this research to maintain model simplicity and estimation time.

![Diagram](image)

**Figure 5** The revised inputs-output model structure of the ARMA model describing the dynamic indoor air temperature of the Industrial Instrumentation Laboratory in this research with some combined inputs to reduce the input numbers for model estimation time reduction and accuracy incrementation while using MATLAB® System Identification Toolbox™.

| No | Item                                    | Length | Width | Height | Area    |
|----|-----------------------------------------|--------|-------|--------|---------|
| 1  | The western part of the northern wall   | -      | 1.20  | 2.90   | 3.4800  |
| 2  | The eastern part of the northern wall   | -      | 3.95  | 2.90   | 11.4550 |
| 3  | The northern part of the eastern wall   | 0.85   | -     | 2.90   | 2.4650  |
| 4  | The southern part of the eastern wall   | 10.15  | -     | 2.90   | 29.4350 |
| 5  | Southern wall                           | -      | 5.15  | 2.90   | 14.9350 |
| 6  | Western wall                            | 11.00  | -     | 2.90   | 31.9000 |
| 7  | Lower floor                             | 11.00  | 5.15  | -      | 56.6500 |
| 8  | Upper floor                             | 11.00  | 5.15  | -      | 56.6500 |

Table 1 The dimension and area of each part of the surfaces for all corresponding inputs of the model.

The ARMA model requires the past output to be used as one of the inputs while the transfer function model does not. Therefore, the past and present indoor air temperature of the laboratory itself, $T_{indInstLab}$ is removed as one of the inputs for the continuous-time transfer function developed in this model and is shown in Figure 6.
equal or greater than the number of zero(s), the number of pole(s) is(are) equal or greater than one, and the number of zero(s) is(are) equal or greater than zero in each continuous-time transfer function model, but the number of pole(s) and zero(s) for all the continuous-time transfer function models in this research are standardised to reduce the searching time for the optimised number of pole(s) and zero(s) during the model estimation process. For a MISO system with \( l \) inputs, one output, and a standardised number of pole(s) and zero(s), the general equation for the continuous-time transfer function model, \( H(s) \) representing the system is shown in Equation (2) below:

\[
H(s) = \sum_{k=0}^{\infty} H_k(s) = H_0(s) + H_1(s) + \ldots + H_{l-1}(s) + H_0(s)
\]

where:

\[
H_k(s) = \frac{\sum_{j=0}^{m} Y_0 j s^{n-j}}{\sum_{j=0}^{n} X_0 j s^{m-j}}
\]

\( x_k(t) \) and \( y_k(t) \) are each the continuous-time input and output signal to and from the continuous-time transfer function model \( H_k(s) \). \( H_0(s) \) is the continuous-time transfer function model producing the output from the first input

\[
H_0(s) = \frac{\sum_{j=0}^{n} Y_0 j s^{n-j}}{\sum_{j=0}^{m} X_0 j s^{m-j}}
\]

and

\[
H_1(s) = \frac{\sum_{j=0}^{n-1} Y_1 j s^{n-j}}{\sum_{j=0}^{m-1} X_1 j s^{m-j}}
\]

\( x_k(t) \) and \( y_k(t) \) are each the continuous-time input and output signal to and from the continuous-time transfer function model \( H_k(s) \) for \( k = 1, 2, \ldots, l-1 \)

\[
H_k(s) = \frac{\sum_{j=0}^{m} Y_k j s^{n-j}}{\sum_{j=0}^{n} X_k j s^{m-j}}
\]

\( x_k(t) \) and \( y_k(t) \) are each the continuous-time input and output signal to and from the continuous-time transfer function model producing the output from the \( l \)th input

\[
H_l(s) = \frac{\sum_{j=0}^{m} Y_l j s^{n-j}}{\sum_{j=0}^{n} X_l j s^{m-j}}
\]

Based on Equation (2) above, the inputs-output data listed in Subsection 2.5 (The Inputs-output Data Selection), and Figure 6, the continuous-time transfer function model for the revised MISO system with standardised \( m \) pole(s) and \( n \) zero(s) representing the indoor dynamic indoor air temperature behaviour of the Industrial Instrumentation Laboratory is written as Equation (3) below:

\[
H_{\text{IndLab}}(s) = H_{\text{Walls}}(s) + H_{\text{Floor}}(s) + H_{\text{Walls}}(s) + H_{\text{Floor}}(s) + H_{\text{Walls}}(s)
\]

where:

\( H_{\text{Walls}}(s) \) is the continuous-time transfer function model producing the output from the heat.
gain/loss due to the air temperature difference between the surrounding horizontal spaces through all the surrounding walls, $\Delta T_{Walls}$

\[
H_{STWalls}(s) = \frac{\sum_{n=0}^{\infty} X_{STWalls}(j^{n-1})}{\sum_{m=0}^{\infty} Y_{STWalls}(j^{m-1})} = x_{STWalls}(s)
\]

$x_{STWalls}(t)$ and $y_{STWalls}(t)$ are each the continuous-time input and output signal to and from the continuous-time transfer function model $H_{STWalls}(s)$. $H_{STFloors}(s)$ is the continuous-time transfer function model producing the output from the heat gain/loss due to the air temperature difference between the vertical spaces through the top and bottom floors, $\Delta T_{Floors}$

\[
H_{STFloors}(s) = \frac{\sum_{n=0}^{\infty} X_{STFloors}(j^{n-1})}{\sum_{m=0}^{\infty} Y_{STFloors}(j^{m-1})} = x_{STFloors}(s)
\]

$x_{STFloors}(t)$ and $y_{STFloors}(t)$ are each the continuous-time input and output signal to and from the continuous-time transfer function model $H_{STFloors}(s)$. $H_{IDWalls}(s)$ is the continuous-time transfer function model producing the output from the heat gain due to the direct solar radiation that lands on the outer-walls of the laboratory, $I_{D\_Walls}$

\[
H_{IDWalls}(s) = \frac{\sum_{n=0}^{\infty} X_{IDWalls}(j^{n-1})}{\sum_{m=0}^{\infty} Y_{IDWalls}(j^{m-1})} = x_{IDWalls}(s)
\]

$x_{IDWalls}(t)$ and $y_{IDWalls}(t)$ are each the continuous-time input and output signal to and from the continuous-time transfer function model $H_{IDWalls}(s)$. $H_{IDSFloors}(s)$ is the optimised transfer function model producing the output from the heat gain due to the diffuse solar radiation that lands on both the outer-walls of the laboratory, $I_{IDS\_Floors}$

\[
H_{IDSFloors}(s) = \frac{\sum_{n=0}^{\infty} X_{IDSFloors}(j^{n-1})}{\sum_{m=0}^{\infty} Y_{IDSFloors}(j^{m-1})} = x_{IDSFloors}(s)
\]

$x_{IDSFloors}(t)$ and $y_{IDSFloors}(t)$ are each the continuous-time input and output signal to and from the continuous-time transfer function model $H_{IDSFloors}(s)$. $m$ and $n$ are each the number of pole(s) and zero(s) for all $H_{STWalls}(s)$, $H_{STFloors}(s)$, $H_{IDWalls}(s)$ and $H_{IDSFloors}(s)$, with $m \geq n$, $m \geq 1$, and $n \geq 0$.

### 2.7 The Black Box Model Estimation and Testing

The parameters that are needed to be estimated in the developed continuous-time transfer function model describing the dynamic indoor air temperature of the Industrial Instrumentation Laboratory in Equation (3) above are $X_{STWalls}$, $Y_{STWalls}$, $X_{STFloors}$, $Y_{STFloors}$, $X_{IDWalls}$, $Y_{IDWalls}$, $X_{IDSFloors}$, $Y_{IDSFloors}$, and $Y_{IDSWalls}$. Since the continuous-time transfer function model in this research is developed with the aid of MATLAB® System Identification Toolbox™, the estimation process of these parameters is done internally using the readily available setting options in the toolbox. Due to time constrain, only the initialisation setting is changed from ‘IV’ (the default setting) to ‘All’ while the rest of the settings are left as default and unexplored. Once the estimation process is done, there is an option in the toolbox to simulate the estimated model and calculate the fitting between the model’s simulated output and the actual output data internally and automatically in the toolbox using the testing data set.

### 2.8 The Black Box Model Optimisation

The MATLAB® System Identification Toolbox™ produces the same numbers and values of the parameters $X_{STWalls}$, $Y_{STWalls}$, $X_{STFloors}$, $Y_{STFloors}$, $X_{IDWalls}$, $Y_{IDWalls}$, and $Y_{IDSWalls}$ in Equation (3) as long as the training data set, the number of pole(s) and zero(s), and the estimation setting options in the toolbox are the same. If the training data set and the estimation setting options in the toolbox are maintained the same during the training process, the only adjustments that can be done for the model’s accuracy adjustment are by changing the numbers of poles(s) and zeros(s), which are the values of $m$ and $n$. Different values of $m$ and $n$ leads to the different numbers and values of the parameters $X_{STWalls}$, $Y_{STWalls}$, $X_{STFloors}$, $Y_{STFloors}$, $X_{IDWalls}$, $Y_{IDWalls}$, $X_{IDSFloors}$, $Y_{IDSFloors}$, and $Y_{IDSWalls}$ in Equation (3). Instead of assigning the values of $m$ and $n$ randomly by using the trial and error method, the possible values of $m$ and $n$ are tried one by one – due to time constraint, the possible values of $m$ and $n$ in this research are tried one by one only from $m = 1$ until $m = 10$ and $n = 0$ until $n = 10$ (with $m \geq n$) – the estimated continuous-time transfer function model with the combination values of $m$ and $n$ that gives the highest best fit value (based on the internal and automatic calculation in the toolbox) is chosen as the optimised model in this research.
2.9 The Black Box Model Verification

It is not yet known in detail regarding how the estimated model’s output simulation and best fit calculation is done internally and automatically in the MATLAB® System Identification Toolbox™ by the time this research is done. Therefore, the optimised continuous-time transfer function model obtained from the toolbox in this research is simulated again manually and externally with both the same training and testing data sets in Simulink® so that the optimised continuous-time transfer function model in this research can be compared fairly using the known manual and external simulation settings with the optimised autoregressive–moving-average (ARMA) model in the previous research in [5] – this is done by calculating the percentage of fitting of the model’s output with the actual data, %Fit by using the following formula:

\[
\% \text{Fit} = \left[1 - \frac{\text{norm}(\hat{T}_{\text{IndInstab}} - \text{meas}(T_{\text{IndInstab}}))}{\text{norm}(\text{meas}(T_{\text{IndInstab}}))}\right]
\]

(4)

where:

\(\hat{T}_{\text{IndInstab}}\) is the output calculated using the optimised black box model, \(T_{\text{IndInstab}}\) is the recorded output data.

3. RESULTS AND ANALYSIS

The optimised continuous-time transfer function model in this research when the value of pole(s) and zero(s) are tested one by one from one until ten for pole(s) and from zero until ten for zero(s) [with the number of zero(s) must be equal or lesser than the number of pole(s) in the continuous-time transfer function model] in the MATLAB® System Identification Toolbox™ is when the numbers of poles and zeros are each six and two – this transfer function model is written as Equation (5) below:

\[
H_{\text{IndInstab}}(s) = H_{\text{STWalls}}(s) + H_{\text{STFloors}}(s) + H_{\text{IDOuterWalls}}(s) + H_{\text{IDSSOuterWalls}}(s)
\]

where:

\(H_{\text{STWalls}}(s)\) is the time transfer function for each type of input.

\(H_{\text{STFloors}}(s)\) is the time transfer function for each type of input.

\(H_{\text{IDOuterWalls}}(s)\) is the time transfer function for each type of input.

\(H_{\text{IDSSOuterWalls}}(s)\) is the time transfer function for each type of input.

The location of poles and zeros for the optimised continuous-time transfer function model for each type of input written in Equation (5) above are summarised in Table 2 and illustrated in pole-zero maps in Figure 7. Based on the location of these poles, it is concluded that the optimised continuous-time transfer function model obtained in this research is stable. It might be possible to approximate this model with a lower order continuous-time transfer function by using the dominant pole approximation technique, but this will be done as one of the publishable improvement items in the future.

Table 2 The location of poles and zeros for the optimised continuous-time transfer function model for each type of input.

| Item | No | \(H_{\text{STWalls}}(s)\) | \(H_{\text{STFloors}}(s)\) | \(H_{\text{IDOuterWalls}}(s)\) | \(H_{\text{IDSSOuterWalls}}(s)\) |
|------|----|-------------------|-------------------|-------------------|-------------------|
| Gain | 1  | 1.98e^{-9}         | -2.57e^{-11}      | -3.28e^{-13}      | -1.26e^{-14}      |
| Zeros 1 | 6.02e^{-4} | -9.26e^{-6} | 1.13e^{-2} | -4.40e^{-5} | -1.16e^{-3} |
| 2   | 3.47e^{-2} | 7.79e^{-3} | -4.00e^{-5} | -4.19e^{-3} |
| Poles 1 | -6.38e^{-8} | -9.31e^{-8} | -2.42e^{-17} | -1.34e^{-7} |
| 2   | -1.02e^{-5} + 1.83e^{-2}/j | -1.27e^{-5} + 9.26e^{-3}/j | -1.26e^{-5} + 9.30e^{-3} | -2.45e^{-5} + 1.82e^{-2}/j |
| 3   | -1.02e^{-5} - 1.83e^{-2}/j | -1.27e^{-5} - 9.26e^{-3}/j | -1.26e^{-5} - 9.30e^{-3} | -2.45e^{-5} - 1.82e^{-2}/j |
| 4   | -5.85e^{-3} + 6.20e^{-2}/j | -2.37e^{-4} - 5.79e^{-3}/j | -2.65e^{-4} - 5.79e^{-3} | -5.09e^{-5} + 5.18e^{-3}/j |
| 5   | -5.85e^{-3} - 6.20e^{-2}/j | -2.37e^{-4} + 5.79e^{-3}/j | -2.57e^{-4} + 5.79e^{-3} | -5.09e^{-5} - 5.18e^{-3}/j |
| 6   | -2.82e^{-1} | -7.27e^{-2} | -4.91e^{-1} | -2.07 |
Figure 7 The pole-zero map for the optimised transfer function in this research for each type of input, which are: (1) $H_{ATWalls}(s)$ (top-left); (2) $H_{ATFloors}(s)$ (top-right); (3) $H_{IPOuterWalls}(s)$ (bottom-left); and (4) $H_{IPSOuterWalls}(s)$ (bottom-right).

Meanwhile, the optimised autoregressive–moving-average (ARMA) model revised in this research and developed in the previous research in [5] when the value of past inputs and outputs, $k$ is tested one by one from $k = 1$ until $k = 200$ are each when $k = 10$ and $k = 41$ – beside the optimised ARMA model developed in the previous research, the result for the optimised ARMA model revised in this research is also presented in this research for a fair comparison with the continuous-time transfer function model in terms of similar types of revised inputs. The calculated percentage of fitness, %Fit values for the simulated outputs for all three models are presented in Table 3 while the simulated outputs for these three models are plotted in Figure 8.

Table 3 The percentage of fitting, %Fit for the optimised models (together with their best model structures) when using the training data set and testing data set.

| No | Model                                           | Optimised Model Structure                                                                 | Percentage of Fitting, %Fit |
|----|-------------------------------------------------|------------------------------------------------------------------------------------------|----------------------------|
| 1  | Previously-developed ARMA Model.               | Standardised 41 past inputs and outputs for all inputs and output.                     | Training Data Set: 87.45   |
|    |                                                 |                                                                                         | Testing Data Set: 80.36    |
| 2  | Newly-revised ARMA Model.                       | Standardised 10 past inputs and outputs for all inputs and output.                     | Training Data Set: 71.01   |
|    |                                                 |                                                                                         | Testing Data Set: 77.18    |
| 3  | System Identification Toolbox – Continuous-time | Standardised six poles and two zeros for all inputs.                                   | Training Data Set: 82.60   |
|    | Transfer Function Model.                        |                                                                                         | Testing Data Set: 80.17    |
Based on the result presented in Table 3 and Figure 8, it can be seen that the simulated output of the optimised continuous-time transfer function model developed in this research is slightly less accurate than the optimised ARMA model developed in the previous research in [5], but slightly more accurate than the simulated output of the optimised ARMA model revised in this research. However, the optimised continuous-time transfer function model developed in this research is more practical to be implemented than both the optimised ARMA model revised in this research and the optimised ARMA model developed in the previous research because the continuous-time transfer function model only requires one past input to calculate the output of one minute ahead while the ARMA model revised in this research requires ten past inputs and outputs and the ARMA model developed in the previous research requires 41 past inputs and outputs just to calculate the output for one minute ahead – the data used in this research (and also the previous research) is recorded at every one minute, this means that the continuous-time transfer function model only requires the data from the past one minute while the ARMA model revised in this research requires the data from the past ten minutes and the ARMA model developed in the previous research requires the data from the past 41 minutes just to calculate the output for one minute ahead.

For future works, it is suggested to do the following things to improve and upgrade the works done in this research and previous research: (1) fine-tuning the existing continuous-time transfer function model developed in this research and ARMA model developed in the previous research for better performance; (2) exploring and investigating the performance of other different types of mathematical models to represent the dynamic indoor air temperature behaviour of the laboratory; and (3) upgrading the existing and future models with controllable inputs from the thermal comfort devices such as the air conditioners, ventilation fans, and motor-operated windows so that the obtained models can be used for simulating the performance of the control algorithms and strategies for the thermal comfort devices in the future. For fine-tuning the existing models, some of the things that can be done are to revise the type(s) of input(s) that may affect the output of the system, to do signal pre-processing and post-processing for the data set that is used to estimate the model, to record more data set with a wider range of operating condition for model estimation and testing, to try different types of parameter estimation algorithm during the model estimation, to introduce signal transport delay element in the model between each type of input and output, to approximate the obtained higher-order continuous-time transfer function model with a lower order continuous-time transfer function model by using the dominant pole approximation technique, etc. – some of these settings can either be programmed for self-programmed algorithm and command line-style toolbox or selected as the available settings in the graphical user interface (GUI)-style toolbox, but are not explored at the moment due to time constrain and will be done in the future. For exploring and investigating the performance of other different types of mathematical models, it is suggested to try modelling the same data with different types of black box model, both linear and non-linear models to investigate their accuracy and practicality (in terms of the simplicity of the model structure). In addition, it is also suggested to try the grey box or even the white box modelling to gain physical insight into the dynamic indoor air temperature behaviour in the laboratory. For upgrading the existing and future models with controllable inputs, the additional
sensors to record the newer data set when the thermal comfort devices are operated is still being set up by the time this research is done and the models will be upgraded once the required data is available in the future.

4. CONCLUSION

The main goal of this study is to develop a more practical but still accurate model for simulating the dynamic indoor air temperature behaviour of Industrial Instrumentation Laboratory, Malaysia-Japan International Institute of Technology (MJIIT), Universiti Teknologi Malaysia (UTM) Kuala Lumpur based on the black box modelling using the continuous-time transfer function model via MATLAB® System Identification Toolbox™. Previously, an accurate black box model representing the dynamic indoor air temperature behaviour of the same laboratory was developed based on the autoregressive-moving-average (ARMA) model in [5], but it is not practical for implementation because it requires too much past inputs and outputs. Even though the continuous-time transfer function model developed in this research is slightly less accurate than the ARMA model developed in the previous research in terms of percentage of fitness, %Fit (82.60% and 80.17% versus 87.45% and 80.36% for training and testing data set), the continuous-time transfer function is much more practical to be implemented than the ARMA model in terms of model structure simplicity (standardised eight poles and two zeros for all inputs versus standardised 41 past inputs and outputs for all inputs and output). The main contribution from this research is in the construction of a new data-driven mathematical model with a simpler model structure for a more practical simulation without sacrificing a significant level of accuracy to simulate a system in a short time with limited system’s physical knowledge. The developed model will be upgraded with controllable input from thermal comfort devices in the future for designing and optimising any proposed energy-efficient control algorithm and strategy quickly and cost-effectively via computer simulation to maintain the thermal comfort in the laboratory while minimising the power consumption to reduce the emission of the environmentally-harmful heat and greenhouse gases generated by the electric power plants that produce electricity from the fossil fuel sources. In addition, this model can also be used for building and optimising predictive controllers such as model predictive controller (MPC)

ACKNOWLEDGMENT

The authors would like to express their appreciation for the support of the: (1) Wind Engineering for (Urban, Artificial, Man-made) Environment Laboratory for supplying the outdoor data recorded using their weather station installed on the rooftop of Malaysia-Japan International Institute of Technology (MJIIT), Universiti Teknologi Malaysia (UTM) Kuala Lumpur; and (2) Biologically Inspired System and Technology (Bio-iST) iKohza for providing the fund to purchase the data recording equipment to record the indoor data and giving access to record the indoor data in the Industrial Instrumentation Laboratory in Malaysia-Japan International Institute of Technology (MJIIT), Universiti Teknologi Malaysia (UTM) Kuala Lumpur.

REFERENCES

[1] D. E. Seborg, D. A. Mellichamp, T. F. Edgar, and F. J. Doyle, Process Dynamics and Control. John Wiley & Sons Incorporated, 2010.

[2] N. S. Nise, Control Systems Engineering, 6th ed. John Wiley & Sons Incorporated, 1995.

[3] S. F. Mohd Hussein, H. Nguyen, S. S. Abdullah, Y. Lim, and Y. Tan, “Black Box Modelling the Thermal Behaviour of iHouse Using Auto Regressive and Moving Average (ARMA) Model,” J. Teknol., vol. 78, no. 6–13, pp. 51–58, 2016.

[4] S. F. M. Hussein et al., “Simplifying the Auto Regressive and Moving Average (ARMA) Model Representing the Dynamic Thermal Behaviour of iHouse Based on Theoretical Knowledge,” vol. 752, 2017, pp. 697–711.

[5] S. F. M. Hussein et al., “Black Box Modelling and Simulating the Dynamic Indoor Air Temperature of a Laboratory using Autoregressive–moving-average (ARMA) Model,” Indones. J. Electr. Eng. Comput. Sci., 2021, in press.

[6] E. Sian, M. Yoshiki, F. Yuan, L. Yuto, and T. Yasuo, “Study of predictive thermal comfort control for cyber-physical home systems,” in 2018 13th System of Systems Engineering Conference, SoSE 2018, 2018, pp. 444–451.

[7] S. E. Ooi, Y. Fang, Y. Lim, and Y. Tan, “Study of adaptive model predictive control for cyber-physical home systems,” in Lecture Notes in Electrical Engineering, 2019, vol. 481, pp. 165–174.

[8] P. Radecki and B. Hencey, “Online Model Estimation for Predictive Thermal Control of Buildings,” IEEE Trans. Control Syst. Technol., vol. 25, no. 4, pp. 1414–1422, 2017.