EVALUATION OF SELECTED DYNAMIC MODELS IN COMPARISON WITH HOT BOX TEST RESULTS FOR MEASUREMENT OF BUILDING ENVELOPE THERMAL PROPERTIES

Xinrui Lu[1] and Ali M. Memari[2]

ABSTRACT

This research is mainly focused on the experimental measurement of R-value by several different models. Building energy consumption accounts for about 40% of the total energy use in the U.S, and therefore accurate energy simulation is desired. The R-value is one of the key parameters that can influence the energy simulation results and therefore is of great importance. The Average Model has long been the most widely accepted method to measure the thermal properties of building components. However, its steady-state assumption and dependence on temperature difference limit its use especially for in-situ measurement. In this study, several dynamic models, including the Pentaur Model and R-C Network Models, are studied with test data obtained from a series of hot box tests performed in the Building Enclosure Testing Laboratory. The results show that the 3R2C model has the best performance and a desirable stability of accuracy with respect to different levels of temperature difference, and therefore is recommended for practical measurement. The results also indicate that unlike the Average Model, the accuracy of dynamic models does not necessarily depend on the level of temperature difference.

KEYWORDS

hot box test, thermal performance, building envelope, dynamic data analysis

1. INTRODUCTION

In recent years a significant amount of effort has been made to reduce building energy consumption (Ma *et al.*, 2012). While there are many factors influencing the energy performance of buildings, such as the climate zone, HVAC system and renewable energy production, the building envelope system plays a critical role as space heating/cooling energy represents a large amount (35% to 70%) of total building energy consumption (Atsonios *et al.*, 2017). The thermal resistance (R-value) is a key parameter of the building envelope system, and accurate

1. Department of Civil and Environmental Engineering, Ph.D. Candidate, Penn State University, University Park, PA, 16802, luxinruipsu@gmail.com (corresponding author)

2. Department of Architectural Engineering and Department of Civil and Environmental Engineering, Professor and Bernard and Henrietta Hankin Chair of Residential Construction, Penn State University, University Park, PA, 16802. (814) 863-9788, memari@engr.psu.edu

quantification of R-value is highly desirable. The theoretical calculation (design R-value) is useful but may vary due to construction irregularities, material deterioration and quality of workmanship; therefore, in-situ measurement of R-value is needed, in particular for specification of products. For example, Peng and Wu (2008) proposed a study showing that the R-value calculated from the in-situ measurement deviates up to 24% compared with the design R-value. For accurate whole-building energy simulation and energy retrofit of existing buildings, the real R-value is also a necessity. Cesaratto and De Carli (2013) found that the difference between the measured and design R-value is about 20% for new or refurbished buildings, which leads to a difference of 11% to 14% in the net heating energy demand by building energy simulation.

The traditional well-established model to measure the thermal properties for building envelope systems is the "Average Model," which assumes steady-state heat flow passing through the specimen and therefore the thermal storage effect can be neglected. Typical applications of the Average Model includes the Hot Box Test (HBT) Method as suggested by ASTM C1363-11 Standard for measurement in laboratory, and the Heat Flow Meter (HFM) Method. The HFM method is internationally accepted and the most widely used one for in-situ measurement. Desogus *et al.* (2011) used the HFM Method to compare the non-destructive and destructive in-situ test method for a brick wall. Ahmad *et al.* (2014) measured the in-situ thermal resistance of hollow reinforced precast concrete walls by the HFM Method. According to the literature, it is well understood that as the HFM Method is based on the steady-state assumption, it requires a large temperature difference between interior and exterior environment and a relatively long testing period. Therefore, for buildings without capacity to effectively heat up the interior space and for cases when a shorter testing period is desired, the HFM Method is not suitable.

The HBT Method, on the other hand, does not have such limitations. The HBT Method generally includes the use of two controlled chambers: the climate chamber and the metering chamber. The metering chamber is used to simulate the indoor environment and the climate chamber is used to simulate the outdoor environment. With the capability of controlling both the interior and exterior environments, a desired level of temperature difference can be created for application of the Average Model. Although it is relatively easy for the HBT Method to reach high accuracy and quantify uncertainties, it requires large facilities and building envelope mock-up assemblies as also indicated by Meng *et al.* (2017), which is not suitable for the in-situ measurement of existing buildings.

Based on the above discussion, the Average Model is not appropriate for non-steady state conditions, especially when the dynamic behavior of building envelope systems is of concern, such as the in-situ environment condition. Moreover, for cases when the documented thermal properties are not accurate or the documentation is missing, which is a common situation for existing buildings, the in-situ measurement for both R-values and thermal capacitances is needed. Therefore, models that can consider the dynamic performance of building envelope systems are of great interest. In the past decade, considerable efforts have been made to explore the application of different dynamic models as introduced in detail in Section 2. However, due to the complexity of dynamic models, a general consensus on the choice of the model that performs better or on the accuracy of the dynamic models compared to the commonly accepted Average Model is somewhat missing. A comparison of the dynamic models with the Average Model by a hot box test is then needed before the dynamic models can be widely used for in-situ measurement. As a hot box test is performed under a controlled environment and thus can reach desirable accuracy for the Average Model, such a comparison provides direct validation for the feasibility of dynamic models. This paper aims to provide a comparative study for the application of several dynamic models by using a series of hot box tests performed in the Building Enclosure Testing Laboratory (BETL).

2. DYNAMIC MODELS

There are several types of dynamic models that can be used for the in-situ measurement of building thermal properties. Anderlind (1992) published a regression model taking the dynamic behavior of the measured wall into consideration as a linear relationship with past temperature changes. This model is later named by the author as the Pentaur Model. Jiménez *et al.* (2008) proposed a study about the identification of thermal properties by R-C Network Models using the MATLAB IDENT toolbox. The authors also described how the R-C Network models can be related to the general ARMAX models. Naveros *et al.* (2014) used a grey-box model to measure the thermal resistance and capacitance of a simple homogeneous wall under real weather conditions. It should be noted that the R-C Network models are also grey-box models described by state-space equations. Deconinck and Roels (2016) did a comparative study regarding the Average Model, the Pentaur (Anderlind's) Model, the AR(MA)X-models and the R-C Network Models. The authors found that the R-C Network Models generally lead to slightly lower estimates of R-values than other dynamic models, while all dynamic models show improved performances compared to the Average Model. The authors then concluded that the Average Model is easy-to-use and reliable for winter measurements, whereas the dynamic models are more complex to use but offer a more versatile applicability. It is worth mentioning that as also stated by the authors, the ARX-models are black-box, meaning that the model parameters have no direct physical interpretations. In other words, the ARX-models are purely data-driven by means of multiple linear regression. In contrast, the R-C Network Models are grey-box, meaning that the model is constructed based on physical knowledge and therefore physical interpretations can be attached to the model parameters. It should also be noted that while the Pentaur Model is also an ARX-model, its stationary part is based on heat transfer principles and therefore provides information for the R-value.

Biddulph *et al.* (2014) also performed a study to compare the Average Model and the R-C Network Models by in-situ measurement. The dataset was collected during winter. The authors concluded that compared with the steady-state Average Model, although with higher uncertainty in results, the R-C Network Models can be used for in-situ measurement to significantly reduce the monitoring period and do not require a consistently high temperature difference as the Average Model, which provides the potential for summer-time measurement.

A new study conducted by Gori and Elwell (2018) also showed that while the error of Average Model increases with the decrease of temperature difference, the use of dynamic models can compensate for this limitation and therefore is applicable for in-situ measurement outside the winter period when the indoor-outdoor temperature difference is generally low. Fonti *et al.* (2017) specifically studied the performance of R-C Network Models of different orders by the MATLAB IDENT toolbox. By comparing the identification parameters of the models such as the FPE (Final Prediction Error) and RMSE (Root-Mean-Square Error), the authors concluded that the second-order R-C Network Model (3R2C) has the best performance.

In this paper, aside from the Average Model, the Pentuar Model and the R-C Network Models are selected for analysis. R-C Network Models of different orders are explored, including first-order (2R1C), second-order (3R2C) and third-order (4R3C). The selected models together with the Average Model are briefly introduced as follows.

TABLE 1. Model Attributes.

Model Types	Attributes
Average Model	Steady-State Calculation
Pentuar Model	Multiple Regression
R-C Network Models	Grey-Box Estimation

2.1 The Average Model

The Average Model is the most widely accepted model for measuring the thermal properties of building components, both on site and in laboratory. This model is based on a steady-state assumption, meaning that heat storage of the tested wall can be ignored, and therefore the amount of heat flux passing through each layer inside the wall remains the same. Hence, the heat flux can be described as:

$$q^i = \frac{1}{R}\left(T^i_{int} - T^i_{ext}\right) \tag{1}$$

where q is the measured heat flux, T_{int} and T_{ext} are the interior and exterior surface temperatures, and R is the thermal resistance to be calculated. The superscript i represents the i^{th} measurement. The Average Model has become the standard measurement method for thermal properties of building components, as suggested by ASTM C1363 and ISO 9869 standards. However, as also indicated by Atsonios *et al.* (2017) and Deconinck and Roels (2016), the in-situ application of the Average Model is often seasonally bounded as it requires a large temperature difference between the interior and exterior environment. It also requires a sufficiently long testing period for the Average Model to provide valid results on-site, which can also limit the use of this model as one will generally want a shorter testing period as possible especially for existing buildings with occupants. It is also worth mentioning that, even though the Average Model is capable of calculating the R-value, it provides no information about the thermal capacitance, nor the dynamic behavior of the building envelope system. For cases when whole-building simulation is needed with building properties measured on site, having only the R-value is not enough, as most building simulation software tools require both the R-value and the thermal capacitance.

2.2 Pentaur Model

The Pentaur Model was developed by Anderlind (1992) for calculating the R-value considering the dynamic performance of the measured wall. The heat flux passing through the wall is described as three separate parts: the first part is a steady-state stationary behavior including the R-value to be calculated, the second and the third parts represents the dynamic performance of the wall relating the current heat flux to the past temperature changes of interior and exterior surfaces. The model is descried as in Equation (2).

$$q^i = \frac{1}{R}\left(T^i_{surf,int} - T^i_{surf,ext}\right) + \sum_{l=1}^{p} A_l\left(T^{i-p+l}_{surf,int} - T^{i-p+l-1}_{surf,int}\right) + \sum_{l=1}^{p} B_l\left(T^{i-p+l}_{surf,ext} - T^{i-p+l-1}_{surf,int}\right) \tag{2}$$

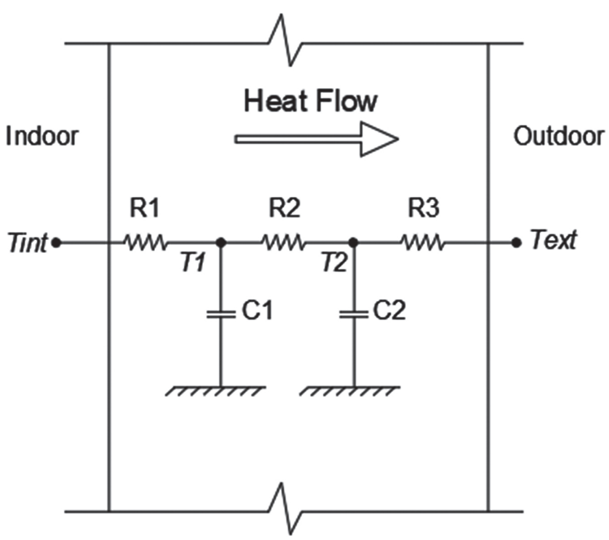
As Equation (2) shows, q is the heat flux, R is the thermal resistance to be calculated, $T_{surf,int}$ and $T_{surf,ext}$ stands for the interior and exterior surface temperature and superscript i means the i^{th} measurement point. A and B are the regression coefficients and p is the number of past data points one chooses to use. As mentioned above, the first part of Equation (2) describes the steady-state behavior of the wall, while the second and third parts consider the dynamic influence in the measured heat flux q^i as a summation of the past surface temperature changes. Users need to determine the number of historical points to use in the Pentaur Model, i.e., the p value. For example, if $p = 5$, the Pentaur Model considers the heat flux measured at current step to be influenced by the past five surface temperature changes. A detailed introduction of the Pentaur Model can also be found in another study performed by Lu and Memari (2018). As can be observed from the above equation, there is no thermal capacitance term in the Pentaur Model, meaning that the dynamic part of this model is not based on physics principles, but rather is data-driven. Therefore, the Pentaur Model can be regarded as a type of ARX-model partially based on physics.

2.3 R-C Network Models

The R-C Network Models are lumped resistance-capacitance models based on real energy balance relationships. This type of model considers the total thermal resistance of the wall as a series of separate thermal resistances between nodes and the total thermal capacitance as separate capacitances lumped at each node. It is an analogy to the electric circuit and with the lumped capacitance assumption, the thermal storage effect can be treated only at each node. R-C Network Models can be utilized with different orders such as 2R1C or 3R2C, however, higher orders do not necessarily lead to better accuracy. Figure 1 shows a typical second-order R-C Network Model (3R2C).

The energy balance at node 1 and 2 is given by Equations (3) and (4), with T_1^j and T_2^j the temperatures at node 1 and 2 for the j^{th} step.

FIGURE 1. 3R2C Model.



$$C_1 \frac{dT_1^j}{dt} = \frac{T_{int}^j - T_1^j}{R_1} - \frac{T_1^j - T_2^j}{R_2} \tag{3}$$

$$C_2 \frac{dT_2^j}{dt} = \frac{T_1^j - T_2^j}{R_2} - \frac{T_2^j - T_{ext}^j}{R_3} \tag{4}$$

The heat flux input can be described as Equation (5):

$$q^j = \frac{1}{R_1}\left(T_{int}^j - T_1^j\right) \tag{5}$$

Rearranging the above equations to get the matrix form of the grey-box model with R_1, R_2, R_3, C_1 and C_2 as identifiable parameters results in the following:

$$\begin{bmatrix} \frac{dT_1^j}{dt} \\ \frac{dT_2^j}{dt} \end{bmatrix} = \begin{bmatrix} -\frac{1}{C_1R_1} - \frac{1}{C_1R_2} & \frac{1}{C_1R_2} \\ \frac{1}{C_2R_2} & -\frac{1}{C_2R_2} - \frac{1}{C_2R_3} \end{bmatrix} \begin{bmatrix} T_1^j \\ T_2^j \end{bmatrix} + \begin{bmatrix} \frac{1}{C_1R_1} & 0 \\ 0 & \frac{1}{C_2R_3} \end{bmatrix} \begin{bmatrix} T_{int}^j \\ T_{ext}^j \end{bmatrix} \tag{6}$$

$$q = \begin{bmatrix} -\frac{1}{R_1} & 0 \end{bmatrix} \begin{bmatrix} T_1^j \\ T_2^j \end{bmatrix} + \begin{bmatrix} \frac{1}{R_1} & 0 \end{bmatrix} \begin{bmatrix} T_{int}^j \\ T_{ext}^j \end{bmatrix} \tag{7}$$

This study uses the MATLAB IDENT toolbox to solve for the above state-space equations. The IDENT toolbox is capable of estimating the parameters in the state-space equations (*greyest* function) as well as providing uncertainty information for the solution, such as the Root-Mean-Square Error (RMSE) and the Final Prediction Error (FPE). It should be noted that for the grey-box estimation, some measured data need to be used as the input for estimation and other measured data need to be used as the output for validation, i.e., to evaluate how good the estimated parameters are. The goal of this kind of estimation is to minimize the difference between the predicted output and the measured output. For this study, the measured interior and exterior surface temperatures are selected as the input and the measured heat flux is selected as the output. The choices of input and output are totally arbitrary.

The main difference between the R-C Network Models and the Pentaur Model is that R-C Network models are based on energy balance equations; therefore, the real value of the wall thermal capacitance can be calculated. In contrast, the Pentaur Model does not provide any information regarding the thermal capacitance. It is also worth mentioning that the nodes in the R-C Network Models are conceptual and do not represent any specific locations inside the wall. Likewise, the R parameters do not represent thermal resistance of any specific layers. Only the summation of the R parameters has a practical meaning as the overall thermal resistance of the wall.

3. EXPERIMENTAL SETUP

This section includes an introduction about the test facility, the specimen, and measurement instrumentation. All the tests were performed in the hot box facility of the Building Enclosure Testing Laboratory (BETL) at the Pennsylvania State University. A series of 14 hot box tests were carried out, each lasting for around 36 hours with temperature differences ranging from 10°C to 40°C. The hot box test facility provides approximately 9.3 m^2 of controlled environment and includes two distinct chambers: the metering chamber used to simulate the indoor environment and the climate chamber used to simulate the outdoor chamber. Heaters and air-conditioners were used to control the temperatures in each chamber. The limitation of the facility in this study was that the metering chamber could only be heated to around 30°C and therefore could not provide a higher temperature difference between two chambers.

The specimen is 110 cm * 70.1 cm in size and the material properties of the specimen are listed in Table 2 (Wolfgang, 2010).

Fenwall 192–103LEW-A01 thermistors were used for temperature measurement with an accuracy of ±0.1°C. A HFP01 HukseFlux heat flux sensor (HFP 01) was used for heat flux measurement with an accuracy of ±5% on walls according to the manufacture's product manual. The 16-bit FP AI-110 was used as the Analog-Digital Converter controlled by Labview. For both the interior and exterior surfaces, four thermistors were used and the readings were averaged to get the surface temperatures. Sensors were mounted on a representative part of the specimen avoiding the edges. A detailed description of the hot box facility and the specimen can also be found in a previous study by Lu and Memari (2018), which has a specific focus on the experimental setup of hot box tests. By performing new tests covering a range of temperature differences, this paper aims to provide a complete understanding of the performances of the explored dynamic models.

The measurement uncertainty is calculated according to ISO/IEC Guide (GUM) taking into consideration the equipment accuracy with a coverage factor of 2. The uncertainty for the measured R-value (u_R) is obtained based on the following expression:

$$u_R^2 = \left(\frac{\partial R}{\partial q}\right)^2 u_q^2 + \left(\frac{\partial R}{\partial T_{int}}\right)^2 u_{T_{int}}^2 + \left(\frac{\partial R}{\partial T_{ext}}\right)^2 u_{T_{ext}}^2 \quad (8)$$

where u_q is the uncertainty of the heat flux measurement and $u_{T_{int}}$ and $u_{T_{ext}}$ are the uncertainties of the interior and exterior surface temperature measurements? The uncertainties for each test are summarized in Table 3.

TABLE 2. Specimen Construction.

Layers	Thickness (mm)	Conductance (W/m^2K)
Parging	9.5	75.59
Concrete Masonry Unit	192.5	1.74
Fiberglass	38.1	0.89
Overall Design R-value: 1.528 m^2K/W		

TABLE 3. Measurement Uncertainty

Test Series	Temperature Difference (°C)	Uncertainty (m^2K/W)
1	10	±0.095
2	10	±0.075
3	15	±0.084
4	15	±0.102
5	20	±0.094
6	20	±0.101
7	25	±0.090
8	25	±0.103
9	30	±0.096
10	30	±0.097
11	35	±0.096
12	35	±0.100
13	40	±0.095
14	40	±0.101

4. RESULT AND DISCUSSION

A series of 14 tests were performed, with a temperature difference ranging from 10°C to 40°C. For each test, the Average Model, the Pentaur Model, and the R-C Network Models of different orders are used for analysis and the results are summarized in Table 44.

The results in Table 4 suggest that all types of models studied provide noticeable differences compared to the design R-value (1.528 m^2K/W), although the variation among models is relatively small. Furthermore, the second to the last row of 4 shows that the standard deviation of the Average Model with respect to changes of temperature differences is the highest among all models, which is expected as the Average Model is sensitive to temperature difference. The dynamic models show lower standard deviations than the Average Model, meaning that the dynamic models do not have a strong dependence on the temperature difference as in the case of the Average Model. Therefore, the dynamic models are more applicable for in-situ measurement especially when it is difficult to create a large indoor-outdoor temperature difference. It should also be noted that while all dynamic models show improved performance, the 3R2C model shows the lowest standard deviation, meaning that the 3R2C model turns out to be the most "stable" one with respect to changes of temperature difference. The result also validates that higher orders of R-C Network Models do not necessarily lead to better performance, as the 4R3C model has a higher standard deviation than the 3R2C model with respect to changes of temperature difference.

TABLE 4. R-value Calculated by Each Model.

Test Series	Temperature Difference (°C)	R-value, m^2K/W (Average Model)	R-value, m^2K/W (Pentaur Model)	R-value, m^2K/W (2R1C Model)	R-value, m^2K/W (3R2C Model)	R-value, m^2K/W (4R3C Model)
1	10	2.173	1.678	1.627	1.627	1.708
2	10	1.840	1.711	1.696	1.764	1.770
3	15	1.858	1.808	1.864	1.689	1.696
4	15	1.917	1.649	1.654	1.669	1.652
5	20	1.764	1.742	1.790	1.671	1.638
6	20	1.826	1.795	1.701	1.701	1.706
7	25	1.638	1.803	1.724	1.731	1.735
8	25	1.760	1.908	1.553	1.708	1.707
9	30	1.730	1.858	1.922	1.734	1.739
10	30	1.742	1.820	1.764	1.716	1.865
11	35	1.710	1.808	1.727	1.693	1.698
12	35	1.767	1.800	1.863	1.717	1.721
13	40	1.743	1.812	1.742	1.726	1.730
14	40	1.817	1.848	1.807	1.780	1.785
Standard Deviation:		0.1266	0.0708	0.1000	0.0392	0.0560
R-value Averaged Percentage Difference		0 (baseline)	0.96%	3.37%	5.38%	4.49%

By setting the R-value calculated from the Average Model as the baseline, the mean R-values calculated from the dynamic models are expressed as percentage difference in the last row of Table 4. It can be seen that the mean R-values calculated from the dynamic models are reasonably close to the Average Model, with percentage differences ranging from 0.96% to 5.38%. It should be noted that a higher percentage value here does not necessarily mean lower accuracy, as the baseline for comparison is the mean R-value from Average Model. Instead, it shows how the dynamic model output differs from the output of the Average Model in an average sense. The percentage difference with respect to a certain temperature difference level can be calculated according to the data provided in Table 44 based on one's need. The model accuracy is further elaborated in the following discussion.

To have a better understanding of the results, the R-values calculated based on different models with respect to changes of temperature difference are plotted in Figure 2. It can be observed from Figure 2 that when the temperature difference is relatively small (10°C), the R-values calculated from different models show large difference compared with the Average

Model. As the temperature difference increases up to 40°C, all types of models give close results. It should also be noted that while the R-C Network Models turn out to give very close results to the Average Model when the temperature difference is large (40°C), the Pentaur Model generally provides slightly higher R-value estimates compared to the R-C Network Models, which has also been found in the study by Deconinck and Roel (2016). Therefore, for in-situ measurement when the temperature difference is large, using dynamic models and the Average Model do not lead to considerably different results. However, when the temperature difference is relatively small, dynamic models are recommended (as shown in Figure 2, a temperature difference of 10°C is enough to be considered as "small" since the results show noticeable variations). It is also clearly shown in Figure 2 that compared with the 2R1C model, the 3R2C model and 4R3C model show better stability, as the results do not fluctuate much with respect to changes of temperature difference as in the 2R1C model.

To better evaluate the accuracy of the dynamic models, identification parameters obtained from the MATLAB IDENT toolbox for grey-box modeling are used for comparison. RMSE is the Root-Mean-Square Error that describes how well the model response fits the estimation data. In other words, as mentioned in Section 2.3 that the heat flux has been selected as the output variable, RMSE describes how the "predicted heat flux" given by the model fits the "measured heat flux." Therefore, the lower the RMSE value, the better the model performs. For the R-C Network Models, two other identification parameters are used to make accurate conclusions on the model performance: FPE and FIT. FPE is the Akaike's Final Prediction error that measures the accuracy of the model outputs as well. FIT (Level of Fit) is the normalized root mean squared

FIGURE 2. R-value versus Temperature Difference.

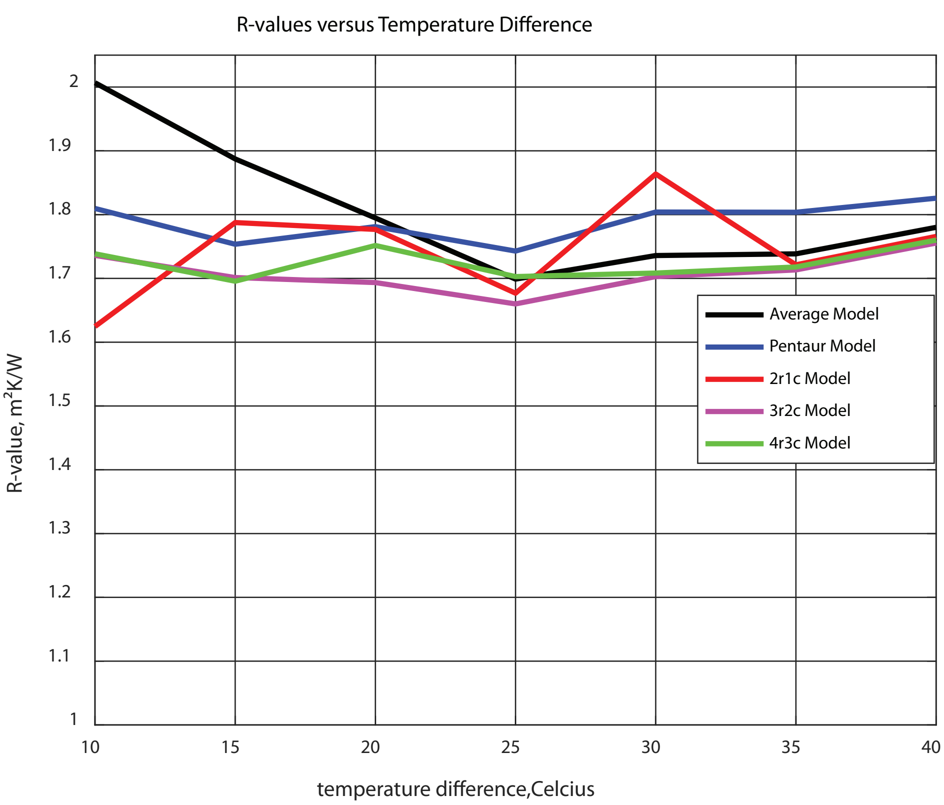


TABLE 5. RMSE of Dynamic Models.

Temperature Difference (°C)	Pentaur Model	2R1C Model	3R2C Model	4R3C Model
10	0.574	0.487	0.407	0.693
15	1.131	2.162	0.512	0.547
20	1.477	1.545	0.540	0.607
25	1.064	1.365	0.462	0.439
30	0.979	1.228	0.497	0.867
35	1.488	1.744	0.507	0.479
40	1.689	1.451	0.563	0.506
Column Average	1.200	1.426	0.498	0.591

error expressed as a percentage that summarizes the model goodness of fit (similar to RMSE). A higher value of FIT means better model accuracy. The three identification parameters are averaged for each temperature difference value, and summarized in Table 5.

At first by comparing the identification parameters for each row, one can find that the 3R2C model always has good accuracy while other models show some variations. For example, the FIT of the 3R2C model always remains above 80% while the FIT of the 2R1C Model varies from 30.39% to 76.13% and that of 4R3C Model from 69.37% to 90%. In general, the 3R2C model has the best accuracy with an averaged RMSE of 0.498, an averaged FPE of 0.263 and FIT of 84.74%, as indicated by the last rows of Table 5 To have a more direct illustration and comparison on the accuracy of the dynamic models, the data recorded in Table 5 is further plotted in Figure 3 and Figure 4.

TABLE 6. FPE of R-C Network Models.

Temperature Difference (°C)	2R1C Model	3R2C Model	4R3C Model
10	0.246	0.173	0.606
15	4.828	0.281	0.318
20	2.449	0.302	0.409
25	1.943	0.226	0.205
30	1.557	0.259	0.947
35	3.273	0.269	0.244
40	2.156	0.330	0.271
Column Average	2.350	0.263	0.428

TABLE 7. FIT of R-C Network Models.

Temperature Difference (°C)	2R1C Model (in %)	3R2C Model (in %)	4R3C Model (in %)
10	76.13	80.59	69.37
15	30.39	83.92	82.12
20	56.58	84.81	82.77
25	52.74	84.44	85.01
30	57.67	82.61	71.42
35	57.99	87.96	88.61
40	71.24	88.86	90.00
Column Average	57.53	84.74	81.33

FIGURE 3. RMSE of Dynamic Models.

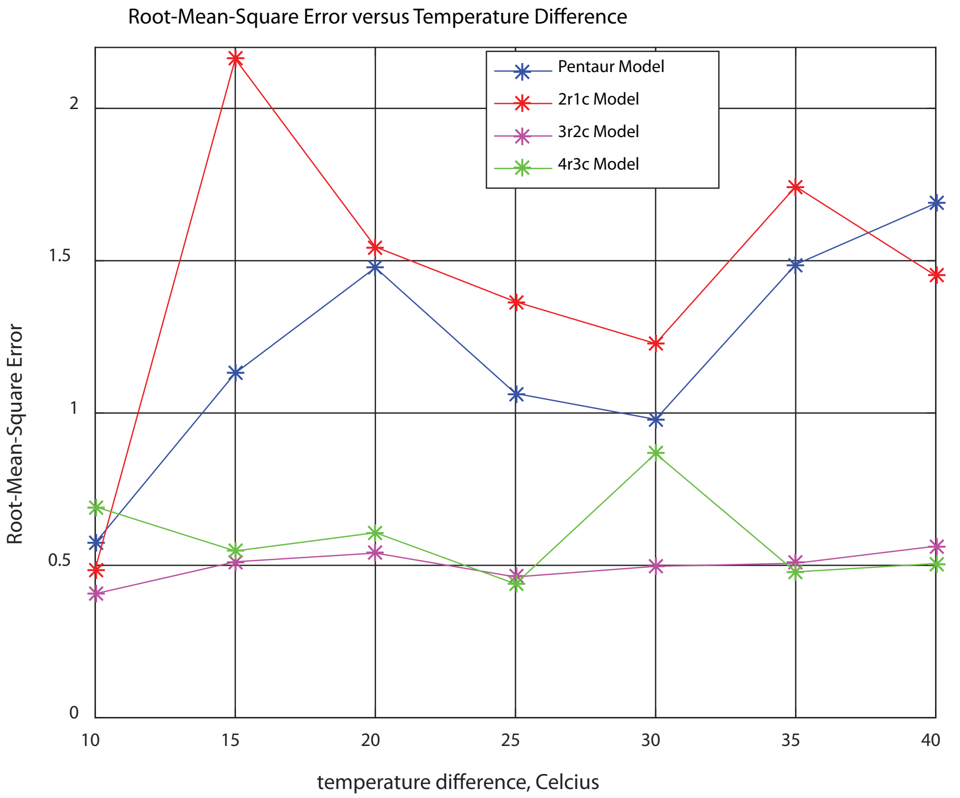
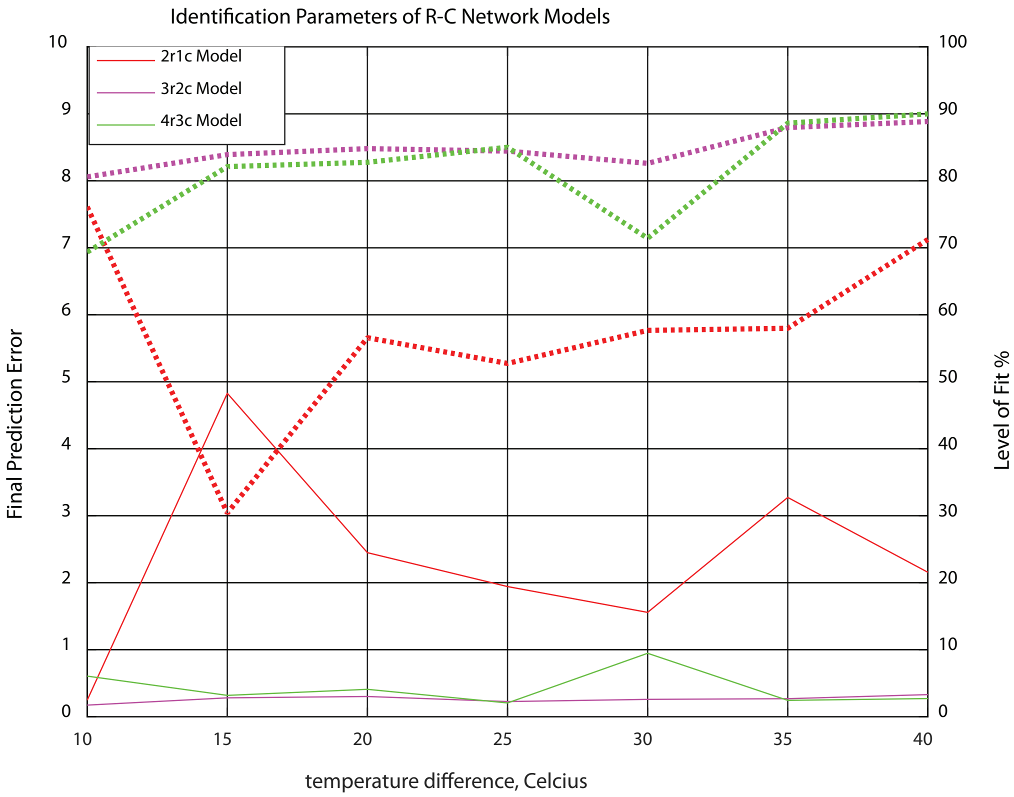


FIGURE 4. Identification Parameters of R-C Network Models. The solid lines represent the Final Prediction Error (left) and the dash lines represent the Level of Fit (right).



From Figure 3 it can be observed that the 2R1C Model generally shows higher RMSE than the other dynamic models for almost all levels of temperature difference, meaning that it is the least accurate model among all the dynamic models explored in this study. The Pentaur Model has slightly better accuracy than the 2R1C Model with a lower RMSE value. The accuracies of the 2R1C model and Pentaur Model also vary noticeably with respect to the changes of temperature difference. The 3R2C and 4R3C models show similar accuracy with lower RMSE values than the other two models. However, the 3R2C model has a better "stability" than the 4R3C model, meaning that the accuracy level of the 3R2C model remains relatively stable regardless of the changes of temperature difference as can be clearly observed from Figure 3. It should also be noticed that unlike the Average Model that has been widely verified to have an improved accuracy with the increase of temperature difference, the accuracy of dynamic models does not have a direct dependence on the level of temperature difference. As shown in Figure 3, larger temperature differences do not necessarily lead to better accuracy.

The discussion above can be further validated by Figure 4. The FPE parameters (solid line) of all dynamic models do not show a direct relationship with the level of temperature difference. While the 3R2C and 4R3C models show a similar level of accuracy better than the 2R1C model, the 3R2C model shows an apparently better stability with respect to the level of temperature difference considering both FPE and FIT parameters. Regarding all the identification parameters, the 2R1C model shows large variations with respect to temperature difference, meaning that even though the 2R1C model may perform well for certain cases, its accuracy cannot be guaranteed at a constant level. Therefore, for measurement under uncontrolled environment, the 3R2C model is the preferred one to use. It is worth mentioning here that all these models need to be validated in real world building conditions in follow-up research.

5. SUMMARY AND CONCLUSIONS

In this paper, the application of several dynamic models for measuring the thermal properties of building components were studied and compared with the standard Average Model. A series of hot box tests were performed in the Building Enclosure Testing Laboratory and the data was collected for analysis and validation of the studied models. The results show that R-values calculated from the dynamic models show percentage differences ranging from 0.96% to 5.38% compared to the Average Model, all indicating a noticeable difference with respect to the R-value. Therefore, for more accurate evaluation of existing buildings and whole-building energy simulation, thermal properties should be measured on-site to obtain the real building performance. When the temperature difference is relatively small (10°C in this study), the Average Model gives a R-value noticeably different from the ones calculated from the dynamic models. When the temperature difference increases, however, all models give close results. The Pentaur Model generally leads to a slightly higher R-value estimate compared with the R-C Network Models. The model accuracy is also studied using MATLAB IDENT toolbox with three identification parameters: RMSE, FPE and FIT. The results show that the 3R2C model has the best performance, showing a slightly better accuracy compared with the 4R3C model and the best stability. The 2R1C model has the lowest accuracy and relatively large instability with respect to temperature difference, and thus is not recommended. For all the dynamic models, unlike the Average Model, the level of accuracy does not depend on the temperature difference. Considering the model accuracy as well as the stability under all temperature differences, the 3R2C model is suggested for practical measurement. It is worth mentioning that the uncertainty level for this type of experiment is also dependent on the choice of sensors, and therefore if different sensors are used, the uncertainty level needs to be recalculated. While the results in this study indicate that the 3R2C model has the best performance, it may not be the most "practical" method for in-situ measurement. As the R-C network model requires more background knowledge in programming and statistics, the Pentaur Model may be a better choice due to its simplicity and convenience. It should also be noted that all the conclusions made in this study are only based on the selected type of building envelope system. For other types of building envelope systems, such as masonry or concrete block, more tests are needed to validate the results. Future work will be focused on the application of these dynamic models on different types of building envelope systems and comparison of more sophisticated methods, such as the transfer function method and the equivalent wall method. The dynamic models explored in this study can be used to measure building thermal properties on site and provide accurate inputs for building energy simulation.

FUNDING

This research was supported by the Pennsylvania Housing Research Center (PHRC). This support is gratefully acknowledged. The opinions expressed in the paper are those of the authors only and do not necessarily reflect those of PHRC

REFERENCE

Ahmad, A., Maslehuddin, M. and Al-Hadhrami, L. M. (2014) 'In situ measurement of thermal transmittance and thermal resistance of hollow reinforced precast concrete walls', *Energy and Buildings*. Elsevier B.V., 84, pp. 132–141. doi: 10.1016/j.enbuild.2014.07.048.

Anderlind, G. (1992) 'Multiple Regression Analysis of Thermal Measurements—Study of an Attic Insulated with 800 mm Loose', *Journal of Building Physics*. doi: doi.org/10.1177/109719639201600109.

ASTM C1363 Standard Test Method for Thermal Performance of Building Materials and Envelope Assemblies by Means of a Hot Box Apparatus (2014) *American Society for Testing Materials*. doi: 10.1520/C1363-11.2.

Atsonios, I. A. *et al.* (2017) 'A comparative assessment of the standardized methods for the in–situ measurement of the thermal resistance of building walls', *Energy and Buildings*. Elsevier B.V., 154, pp. 198–206. doi: 10.1016/j.enbuild.2017.08.064.

Biddulph, P. *et al.* (2014) 'Inferring the thermal resistance and effective thermal mass of a wall using frequent temperature and heat flux measurements', *Energy and Buildings*. Elsevier B.V., 78, pp. 10–16. doi: 10.1016/j.enbuild.2014.04.004.

Cesaratto, P. G. and De Carli, M. (2013) 'A measuring campaign of thermal conductance in situ and possible impacts on net energy demand in buildings', *Energy and Buildings*. Elsevier B.V., 59, pp. 29–36. doi: 10.1016/j.enbuild.2012.08.036.

Deconinck, A. H. and Roels, S. (2016) 'Comparison of characterisation methods determining the thermal resistance of building components from onsite measurements', *Energy and Buildings*. Elsevier B.V., 130, pp. 309–320. doi: 10.1016/j.enbuild.2016.08.061.

Desogus, G., Mura, S. and Ricciu, R. (2011) 'Comparing different approaches to in situ measurement of building components thermal resistance', *Energy and Buildings*. Elsevier B.V., 43(10), pp. 2613–2620. doi: 10.1016/j.enbuild.2011.05.025.

Fonti, A. *et al.* (2017) 'Low Order Grey-box Models for Short-term Thermal Behavior Prediction in Buildings', *Energy Procedia*. Elsevier B.V., 105(May), pp. 2107–2112. doi: 10.1016/j.egypro.2017.03.592.

Gori, V. and Elwell, C. A. (2018) 'Estimation of thermophysical properties from in-situ measurements in all seasons: Quantifying and reducing errors using dynamic grey-box methods', *Energy and Buildings*. Elsevier B.V., 167, pp. 290–300. doi: 10.1016/j.enbuild.2018.02.048.

Jiménez, M. J., Madsen, H. and Andersen, K. K. (2008) 'Identification of the main thermal characteristics of building components using MATLAB', *Building and Environment*, 43(2), pp. 170–180. doi: 10.1016/j.buildenv.2006.10.030.

Lu, X. and Memari, A. M. (2018) 'Comparative study of Hot Box Test Method using laboratory evaluation of thermal properties of a given building envelope system type', *Energy and Buildings*. Elsevier B.V., 178, pp. 130–139. doi: 10.1016/j.enbuild.2018.08.044.

Ma, Z. *et al.* (2012) 'Existing building retrofits: Methodology and state-of-the-art', *Energy and Buildings*. Elsevier B.V., 55, pp. 889–902. doi: 10.1016/j.enbuild.2012.08.018.

Meng, X. *et al.* (2017) 'A new simple method to measure wall thermal transmittance in situ and its adaptability analysis', *Applied Thermal Engineering*. Elsevier Ltd, 122, pp. 747–757. doi: 10.1016/j.applthermaleng.2017.05.074.

Naveros, I. *et al.* (2014) 'Setting up and validating a complex model for a simple homogeneous wall', *Energy and Buildings*. Elsevier B.V., 70, pp. 303–317. doi: 10.1016/j.enbuild.2013.11.076.

Peng, C. and Wu, Z. (2008) 'In situ measuring and evaluating the thermal resistance of building construction', *Energy and Buildings*, 40(11), pp. 2076–2082. doi: 10.1016/j.enbuild.2008.05.012.

Wolfgang, B. M. (2010) *Hygro-thermal Performance of Imperfectly Protected Below-Grade Walls with Interior Insulation*.