



Title	A generalization approach for reduced order modelling of commercial buildings
Authors(s)	Shamsi, Mohammad Haris, Ali, Usman, O'Donnell, James
Publication date	2019-07-26
Publication information	Shamsi, Mohammad Haris, Usman Ali, and James O'Donnell. "A Generalization Approach for Reduced Order Modelling of Commercial Buildings." Informa UK Limited, July 26, 2019. https://doi.org/10.1080/19401493.2019.1641554 .
Publisher	Informa UK Limited
Item record/more information	http://hdl.handle.net/10197/10996
Publisher's statement	This is an electronic version of an article published in Journal of Building Performance Simulation. Journal of Building Performance Simulation is available online at: www.tandfonline.com/doi/10.1080/19401493.2019.1641554
Publisher's version (DOI)	10.1080/19401493.2019.1641554

Downloaded 2026-05-02 00:29:39

The UCD community has made this article openly available. Please share how this access benefits you. Your story matters! (@ucd_oa)



© Some rights reserved. For more information

ORIGINAL ARTICLE

A generalisation approach for reduced order modelling of commercial buildings

Mohammad Haris Shamsi^a, Usman Ali^a and James O'Donnell^a

^a School of Mechanical and Materials Engineering, Energy Institute, University College Dublin, Belfield, Dublin 4, Ireland.

ARTICLE HISTORY

Compiled August 15, 2019

ABSTRACT

Energy-efficient retrofits have become crucial in building sector as approximately 80% of buildings in developed countries are over 10 years old. Building simulation tools are now being used to provide estimates of energy consumption and implement various models which differ on the basis of enclosed details. Not all of these models are effective in terms of computation and the associated computational costs.

This work devises a novel and generalised reduced-order grey-box modelling approach to predict the thermal behaviour of commercial buildings. The generalisation approach reduces the order/complexity of model and lays out a general structure to obtain reduced-order models based on easily identifiable building metrics. We also implemented a forward-selection procedure to compare results obtained using a metrics-based approach. The network order obtained using metrics-based approach matches with the network order predicted by the forward selection procedure. The generalised structure would reduce the complexities involved in the dynamic simulation of urban building stock.

KEYWORDS

Energy Modelling; Reduced-order Models; Commercial Buildings; Thermal network RC models

1. Introduction

Energy efficient retrofits have become crucial in the building sector as approximately 80% of the buildings in developed countries are over 10 years old and consume a major portion of total energy demand (Pérez, Ortiz, and Pout 2008). According to the International Energy Agency, global building sector energy intensity (measured by final energy per square metre) fell by 1.3% per year between 2010 and 2014 due to the continued adoption and enforcement of building energy codes and efficiency standards. However, the progress has not been fast enough to offset growth in floor area (3% per year globally) and increasing demand for energy services in buildings (International Energy Agency 2017). A growing number of countries have devised policies to improve the building energy performance, but the average energy consumption per person in this sector still remains practically unchanged since 1990. To meet the required targets, average building energy use per person globally needs to fall by at least 10% (less than 4.5 MWh) by 2025. Concerted global effort is needed to rapidly expand, strengthen

and enforce building energy policies across all countries to avoid inefficient building investments.

Buildings represent complex systems with high levels of interdependence on many dynamic external sources (weather, occupancy etc.). In addition, the optimisation of building systems requires balancing of sometimes contradictory objectives in terms of energy efficiency and overall performance. Building Energy Performance Simulation (BEPS) tools provide an efficient means of conducting performance-based analysis and optimisation, taking into account the multitude of complex model inter dependencies, internal and external inputs as well as various performance objectives. BEPS tools, such as, EnergyPlus, TRNSYS etc. are now being implemented not only at the design stage, but also in latter post-construction stages of the building life-cycle (BLC), such as commissioning and operational management and control (Crawley et al. 2008). These tools are also used to evaluate the retrofit strategies at the individual building scale or at a district scale (Lin, Middelkoop, and Barooah 2012). Moreover, the existing BEPS tools produce non-scalable and non-generalisable models (Wang et al. 2013; Amara et al. 2015).

The development of detailed building models (using BEPS tools) often requires detailed geometric and non-geometric data and therefore, is not a cost-effective solution (Harb et al. 2016). As each individual building differs in terms of structural parameters and nature of operation, developing a detailed model for each building would be impractical. Furthermore, owing to the existing complex and inter-dependant relationships in the system network, the detailed building models often turn out to be computationally intensive (Coakley, Raftery, and Keane 2014). Therefore, it is crucial to devise modelling approaches that utilise sparse data and represent the actual behaviour of the building under study. Grey box modelling is one such approach, which has been used on a wide scale (Lin, Middelkoop, and Barooah 2012).

Grey box modelling delivers the advantages of data-driven and physical modelling approaches. This approach entails the accuracy levels of physical modelling approaches and the computational efficiency of data-driven modelling approaches. However, the grey box approach often leads to application and stakeholder specific models, for instance, the design approach for grey box modelling of commercial buildings differs on a case by case basis (Harb et al. 2016). A grey box model of one individual building can't be implemented to perform energy simulations for another building (even after the modification of model parameters). Another major drawback of grey box models is that the scalability of these models is limited by the network order. The network order defines the level of complexity incorporated in the model. Reduced order grey box approaches counter these limitations by achieving a trade off between the network order and desired accuracy (Lin, Middelkoop, and Barooah 2012). There is a need for a generalisation framework to address the limitations associated with grey-box networks.

The main objective of this paper is to develop a novel generalisable framework which could be used to identify grey box networks for different types of buildings. Physical parameters and the nature of operation of individual buildings constitute the key elements of the developed framework. The framework also relies on the retrofit history and installed HVAC system to deduce the order of the grey box network. Parameters associated with the network are identified using standard calibration strategies.

The paper is organised as follows: Section 2 states the basics and characterisation of building energy models with a detailed insight into grey-box modelling and reduced-order grey-box models. This section also illustrates the previous research conducted in this domain along with the identified research gaps. Section 3 describes the method-

ology formulated to realise the generalisation of grey-box RC network models. The section relates the concept of generalisation to crucial building metrics that aid in the determination of network order and further states the implemented model parameterisation and selection procedures. Section 4 includes the demonstration cases and discussions followed by conclusions and future work in Section 5.

2. Building Performance Simulation Approaches

The building energy modelling process is typically used to evaluate the implications of different energy design or retrofit implementations. The modelling process is defined as the virtual or computerised simulation of a building or complex that focuses on energy consumption, utility bills and life cycle costs of various energy related items such as air conditioning, lights and hot water (Louis G. Birta Gilbert Arbez 2013). The payback of renewable energy solutions like solar panels and photovoltaics, wind turbines and high efficiency appliances can efficiently be estimated using building energy models. Efficient operation in buildings can be ensured through the process of building energy modelling. Modelling approaches for building performance simulation differ on the basis of the level of detail included in the development of an energy model (Zhou et al. 2008). These approaches are often modified to suit the intended area of application, to achieve a desired level of accuracy or to reduce the computational time (Amara et al. 2015). The different modelling approaches are discussed in Section 2.1.

2.1. Modelling Approaches

The energy modelling approaches are mainly categorised as white box, black box and grey box approaches. Also called as forward modelling approach, the white box approach implements extensive physics based equations for modelling the individual components, sub-systems and systems inside any building. The system dynamics are represented through a detailed description of the associated physical aspects (Li and Wen 2014). The parameters are derived from the physical characteristics, the details of which can be obtained from the design plan, manufacture catalogue or on-site measurements. The derived model inputs are then used to obtain the energy consumption and formulate control and operational schemes (Crawley et al. 2008). Popular white box simulation software include EnergyPlus (EP 2017) and, TRNSYS (TRN 2017). However, the process is usually time consuming and requires large computational effort, which restricts its widespread use in district energy modelling.

Black box approach doesn't require a description of the detailed physical dynamics of the building and can be easily set up. The models derived using the black box approach only represent a relationship between inputs and outputs. The black box models use on-site demand measurements to train predictions of building operation and have been extensively used to design building control strategies (Mustafaraj, Lowry, and Chen 2011). Popular statistical methods, namely, linear regression and self regression models, define the relationships among dependant variables and have been applied in practice to predict the monthly energy consumption of buildings (Ma et al. 2012a). One major concern with black box models is that the performance of these models depends on the input training data set. Also, the parameters often do not imbibe any physical relevance and are often inconsistent with physical reality when applied under hard conditions (sparse building data) (Li, Wen, and Bai 2015).

Grey box approach uses a combination of physical and measured data. This approach represents an amalgamation of white box and black box approach and hence, offer a trade off between the level of computational speed and accuracy. Grey box models integrate the simplified physical model of the building with a model identification process to identify the design parameters. Using simplified physical models reduces the requirement of training data sets and hence, the calculation time. Grey box approach utilises trained simplified physical models, for instance RC networks, to represent the building physics.

A significant number of studies in literature have identified the importance of implementing grey box models in behaviour prediction and control of a building or a group of buildings. A recent study by Belić, Hocenski, and Slišković (2016) implemented RC networks to represent the different walls inside a two-storey family house. The study analysed the effect of RC network complexity on the accuracy level. Another interesting study by Ji et al. (2016) considered a lumped parameter model to estimate the hourly cooling demand of a shopping mall in Shanghai. Both the studies focused on finding the appropriate order of the network to maintain the energy prediction error below 10%.

A detailed study by Harb et al. (2016) investigated the implementation of four grey box models in their ability to forecast the building indoor temperature. The study concluded that a two capacitor model structure with an additional consideration of the indoor air as a mass-less node enables the most accurate qualitative prediction of the indoor temperature. The models were treated with measurements from the buildings in normal operation and the case study included two residential buildings and one office building. Fux et al. (2014) used a series of RC network models (orders 1 to 4) to forecast the indoor temperature. The results identified that a first order model is able to predict the indoor temperature with a root mean square error of $0.5^{\circ}C$. However, the models were treated with data during unoccupied periods. A study by Mejri, Palomo Del Barrio, and Ghrab-Morcos (2011) formulated the size of the model by testing a number of different orders for a small single floor office building and compared the resulting models in terms of their ability to simulate the thermal response. The study concluded that model order greater than two doesn't lead to any significant improvement in simulation results and over-parameterisation was observed when increasing the model size.

Reduced order models have proved to be quite efficient in representing the building dynamics accurately. A study by Goyal and Barooah (2012) investigated several model order reduction techniques. The study implemented balanced truncation methods to compute the sparsity patterns for the non-linear part of the building dynamics. However, the study was not able to identify the applicability of these reduction techniques. Deng et al. (2010) proposed a methodology for model order reduction using aggregation of states using Markov chains. The simplified models were able to retain the physical intuition of the original model. However, the techniques did not take into account any scalability perspectives to allow for modelling a building cluster. Remmen et al. (2017) developed a tool, called TEASER, to address the challenge of data acquisition required for dynamic modelling of buildings at an urban scale. The tool uses a data enrichment process for identification of the network model for various building archetypes. However, the tool implements standard grey box networks that could represent only a certain level of complexity (maximum network order of 4). Also, the grey box networks in TEASER have been developed according to German standards and hence, still need to be integrated with international statistical building stock data to extend its applicability to other countries.

Different buildings have different usage characteristics, which deems it necessary to model the heat dynamics associated with each building differently. As the dynamics are directly related to the network order, a network identification procedure is required to represent the dynamics efficiently (Reynders, Diriken, and Saelens 2014). As evident from the literature, the existing grey-box network development methodologies use statistical techniques for grey-box order identification, which is a long iterative procedure and often leads to over-fitting of the model. Besides, the procedure needs to be repeated every time for each building. Furthermore, only a few studies are able to tackle the non-scalable nature of reduced-order grey-box models. The methodology proposed in this study introduces a generalisable mechanism to address the aforementioned gaps.

3. Methodology

The grey box network development tools and methodologies as identified in the literature tend to be case specific and need to be modified according to the building in consideration. To cater to the varied nature of building operation, we identified a generalisation approach to identify the order of any grey box network. The approach utilises easily identifiable building metrics, such as, gross floor area, window area, etc. to identify the network order. Different metrics that could be directly or indirectly linked to the parameters of the grey box network are included in the framework of the generalisation approach.

This work implements a generalised grey-box modelling approach (Fig. 1) applicable for all types of commercial buildings. The methodology consists of three crucial steps, namely, 1) identification of building metrics, 2) analysis of building metrics and 3) subsequent development of grey-box networks. The identification and analysis of building metrics identify and analyse the factors influencing the order and parameters of grey-box networks. The grey-box model development process involves model input identification, model calibration and identification, and model evaluation. The model calibration and identification process is further sub divided into network parameterisation and model selection process. This is the most crucial step as the order of the developed grey-box network is decided using the information gathered in the first two steps.

As commercial buildings exhibit varied nature of operation, a generalised approach would reduce the complexities in identifying an accurate model for each individual building. The models designed using the approach should be able to predict the thermal behaviour of a building or a stock of buildings.

The key aspect of generalisation lies in the identification of network order, which is linked to the accuracy and complexity of grey-box models. Previous research developed techniques to identify the optimised network order, although the optimisation is conducted on pure statistical grounds (Perera, Pfeiffer, and Skeie 2014). For instance, a forward selection statistical procedure starts with the smallest feasible model and then extends the model by analysing the p-values associated with each extension. The p-value is the level of marginal significance within a statistical hypothesis test representing the probability of the occurrence of a given event (De Smith 2018). A tolerance limit is set corresponding to the p-values; the violation of which terminates the iterations.

The network order identification technique devised in this paper is based on the analysis of pre-defined building metrics. These metrics include the building's physical

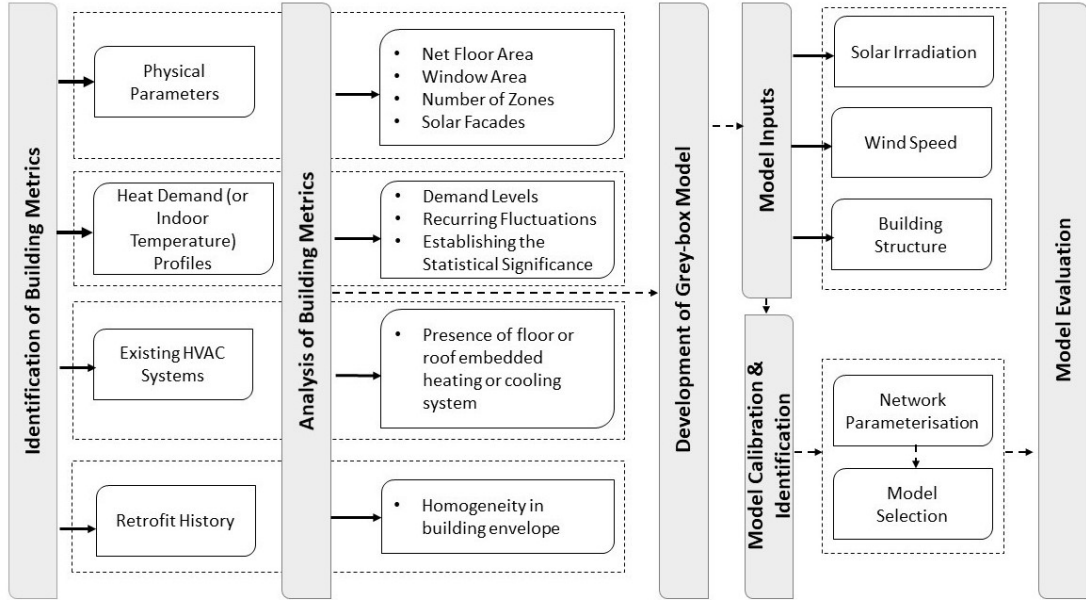


Figure 1. Devised methodology for generalisation. The methodology is based on the analysis of easily identifiable building metrics. All the metrics are tested for their statistical significance and then, used to generate the grey-box networks.

parameters, building type and associated functions. The following sections describe the processes followed in the identification and analysis of building metrics that affect the heat dynamics of a building.

3.1. Identification of Building Metrics

Grey-box models use a combination of prior physical knowledge and statistics (Coakley, Raftery, and Keane 2014). The identification of building metrics goes one step back before the development of any grey-box model. This step involves the identification of metrics that directly influence the grey-box model parameters (or the heat dynamics). The modelling of heat dynamics can be related to two fundamental aspects, namely, the behaviour of heat transfer within a building and the subsequent interactions with different entities (De Coninck et al. 2016).

A grey-box network includes resistance and capacitance elements representing the heat transfer behaviour of the system. The accumulated thermal energy by a physical medium (the air inside a room) or object (walls in a building) is described by its heat capacity C in MJ/K (the capacitance in the network). The heat of any object is a function of temperature and can be defined by eq.(1).

$$C(T) = mc(T) = \frac{dQ}{dT} \quad (1)$$

where m is the mass of the object in kg and $c(T)$ is the specific heat capacity (MJ/kg.K) of the material in the object. The above equation represents a linear approximation where the dependency of T is marginal. Heat transfer mechanisms, such as conduction and convection are represented as resistances against the heat transfer.

We identified various building metrics that are directly or indirectly related to the

grey box network parameters, which constitutes the initial step in the devised methodology. The metrics are identified using previously concluded studies and reports on building energy simulation (Coakley, Raftery, and Keane 2014). These metrics will eventually account for the dynamics of heat transfer inside a building. The metric identification process involves the use of different datasets containing measured and simulated data.

Physical parameters, such as net floor area, number of windows, building envelope, number of zones etc., have a direct influence on the buildings dynamics and hence, play an important role in deciding the network order (Clarke 2001). The parameters and subsequent relationships are identified based on previous literature (Wang and Xu 2006).

The ongoing activities inside a building are usually reflected in the building’s heat demand (Braun and Chaturvedi 2002) and indoor temperature profiles (Zhao and Magoulès 2012). These profiles then relate to the complexity required in representing the dynamics. As the heat demand data was readily available, we used the measured heat consumption to capture the link between the demand patterns and network order. It is crucial to compare a multitude of demand patterns to establish appropriate linkages. Therefore, we used the heat demand data for different commercial buildings at the University College Dublin’s campus as these buildings represent varied nature of building function. Detailed white-box models are also available for different buildings on the campus, which will eventually help in validation of the developed grey-box networks. Furthermore, to investigate the relationship between various physical parameters and the network order, we used the commercial building benchmark model dataset, developed and published by the U.S. Department of Energy (Deru et al. 2011). There are 16 building types in the dataset that represent approximately 70% of the commercial buildings in the U.S. As this dataset consists of white box model representation of the typical commercial building stock, we were able to devise detailed relationships between the building parameters and grey box network order.

Any recent retrofits performed on any building structure will change the dynamics and hence, the retrofit history of a building can be a crucial metric in deciding the desired complexity (Heo, Choudhary, and Augenbroe 2012). These parameters are detailed in Section 3.2.

3.2. *Analysis of Building Metrics*

After the identification process, the metrics are analysed to determine the extent of their influence on the network order.

3.2.1. Physical Parameters

The various physical parameters that have a direct or indirect influence on the grey-box network order are outlined below.

Gross Interior Floor Area (m^2): The gross interior floor area is defined as the total interior floor area of a buildings spaces, measured from the inside surface of the exterior walls or from the interior surface of walls in common with adjoining buildings (Deru and Torcellini 2005). This parameter is used together with other metrics, for instance heat demand profiles, to decide the desired complexity.

Window Area (m^2): The presence of different network elements in the grey-box model is directly influenced by the window area. Windows do not possess any heat retaining capacity and hence are modelled as resistances to heat transfer. This param-

eter is crucial in defining the network order from two aspects. The first aspect relates to windows being a part of the building envelope, which eventually reduces the network complexity due to reduction in the number of capacitance in the network. The second aspects relates to the presence of windows in the internal environment of the building. A large number of windows inside any building will reduce the heterogeneity in the temperature profiles, thus, reducing the complexity in the grey box network. This parameter is used to obtain an approximation of the window to wall ratio existing in the external envelope and hence, the methodology doesn't use the exact values in any calculations.

Number of Floors: This parameter is often used together with the floor area to determine the complexities associated with heat demand profiles. For instance, a high-rise office building might have a high energy consumption although each floor represents a similar demand usage profile with limited complexity.

Number of Zones: This parameter is particularly specific to high rise buildings with large number of floors provided the inside architecture space of each floor is similar. The spaces inside a commercial building are often divided into zones for optimised utilisation of HVAC systems. A single zone represents an area possessing similar thermal characteristics. Hence, with the knowledge of this parameter, the process of deciding the order and hence the grey-box network could be simplified. This parameter is used to get an approximation of the existing zones in order to facilitate the process of network identification. Exact zoning knowledge is not required for any calculations in the methodology.

Solar Facades: Large commercial building envelopes present a great potential of utilizing solar radiation, especially in climate zones with rich solar resources. Solar PV integrated into the building's envelope has a great potential, and far wider application prospects, than rooftop solar power can offer alone. Solar panels can be integrated into the building envelope through placing them on windows or roofing tiles. The heat transfer dynamics are modified through the integration of these panels, which adds up to the complexity of the grey-box networks (Biyik et al. 2017). For instance, a study by Fung and Yang (2008) developed a one dimensional transient heat transfer model for building integrated PV modules. The study investigated the effect of different parameters and found that the change in heat gain with and without the integrated facades is quite significant.

3.2.2. Heat Demand Profile

To develop a generalised approach to decide the RC network order, we analysed the heat demand profiles of different buildings as these profiles indicate the nature of building operation. Furthermore, the implemented methodology considers the thermal zoning of spaces inside the buildings. Thermal zones are assumed to possess similar thermal characteristics and hence can be represented by a single RC network. We identified the grey-box network order using the following procedure (Fig. 2):

Inspection of yearly heat demand profiles and subsequent occurrence of peaks - a high order network (greater than two) is assigned to a building with frequent and recurring peaks (peaks that occur often or at short intervals and repeat time after time). The frequent and recurring peaks relate to the complex nature of operation of the building in consideration. The demand profiles only give an indication of the network order. The variations in the profiles are tested using tests of statistical significance such as ANOVA (Analysis of Variance). ANOVA is a statistical technique used to check if the means of two or more groups are significantly different from each

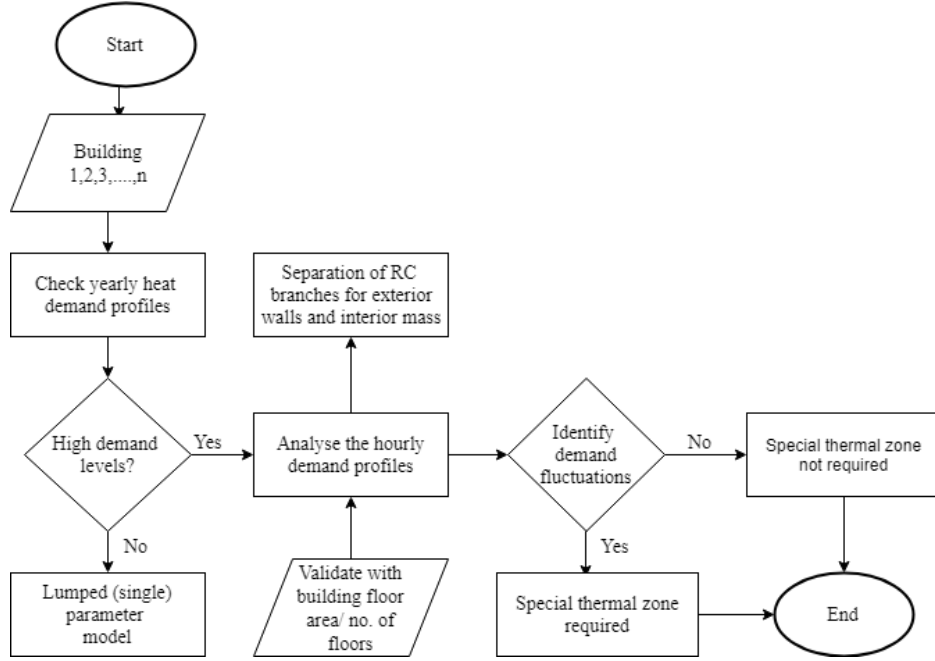


Figure 2. The process of RC network selection. The process is used to identify the network order based on heat demand levels and the presence of demand fluctuations.

other (Crawley 2007). ANOVA checks the impact of one or more factors by comparing the means of different samples. The different groups in ANOVA relate to the different buildings in the cluster. Each day is then considered as a sample. Performing an ANOVA test on the heat demand (W) data confirms the existence of demand variations in the community of buildings. F-test statistic (ratio of two quantities that are expected to be roughly equal under the null hypothesis) is used to compare the results from ANOVA. A large value of F-test statistic signifies that the variability of group means is large relative to the within group variability (required for rejecting the null hypothesis).

Validation of network order is then performed by combining heat demand profiles with net floor area of building and number of floors. Heat demand levels are defined as the magnitude of heat demand at a particular point in time. Higher heat demand levels might not always relate to the heat demand profile of the building as large buildings often have higher demand levels, which suggests that large buildings can be represented using low order networks. Large buildings often represent varying energy dynamics in the inside spaces and the exterior walls. Hence, it is crucial to represent both using separate RC networks.

Hourly periodic fluctuations in heat demand profiles are analysed against a base demand to identify the presence of special thermal zones. Base demand is defined as the minimum heat demand level existing in a heat demand profile during occupied hours. Again, ANOVA is used to analyse the fluctuations in heat demand, the samples being the hourly demand profiles. This test establishes the presence of peaks in the profiles that can be termed as statistically significant.

Heat demand fluctuations can be directly linked to a space inside the building that needs to be assigned a separate special thermal zone, thereby requiring a new RC branch altogether. In case of the presence of statistically significant peaks, the network is extended by addition of additional grey box elements. This process requires the heat

demand data with resolutions ranging between 15 and 60 minutes.

3.2.3. Existing HVAC Systems

The HVAC systems mainly influence the network order of a grey-box model in two ways, namely, through ventilation and through embedded heating or cooling systems in the building fabric. Ventilation causes heat transfer due to the movement of conditioned air across different spaces. The indoor temperature is very much dependent on the flow of air. Ventilation can be modelled as resistances to heat transfer for wind speeds up to 5 m/s (Chen 2009). Beyond this value, the heat transfer becomes non-linear leading to an increased complexity in the system.

Embedded HVAC systems alter the dynamics of the building by increasing the heat capacity associated with the area where the systems are fitted. For instance, radiant heating or cooling systems use piping filled with water or a glycol solution and can heat any kind of floor/ceiling (Djuric et al. 2007). These systems alter the heat capacity of the floor, thus, increasing the complexity associated with the grey box model.

3.2.4. Retrofit History

The retrofit history of the building is another crucial parameter that is directly linked to the complexity of the grey-box network. The most common retrofit technologies for commercial buildings consist of the strategies to reduce building heating and cooling demand, and the use of energy efficient equipment. The heating and cooling demand of a building can be reduced through retrofitting the building fabric or through technologies that improve air tightness (Ma et al. 2012b). All these measures, especially the improvement of building fabric, alter the building dynamics and therefore, increase the complexity associated with the grey-box model (to account for the loss of homogeneity in building envelope). Simply changing the network parameter values might lead to unrealistic description of the grey box network. The retrofit history relates to the increased heterogeneity in the building envelope, due to which, a higher order network might be required to represent the building.

3.3. Development of Grey-box Model

Grey-box model development initiates with the specification of model inputs followed by model calibration and identification and terminates with model evaluation. The model calibration and identification is further divided into network parameterisation and network selection processes. The details of the aforementioned processes are listed in the following sections.

3.3.1. Model Inputs

The inputs to the model include solar irradiation, wind speed, and building structure. Solar irradiation and wind speed inputs are extracted from the weather data. Building structure includes geometric information regarding the building.

3.3.2. Model Calibration and Identification

Model calibration and identification is the most crucial step in grey-box model development and follows the identification and analysis of building metrics affecting the

network order. This step consists of network parameterisation procedure and model selection (or selection of network order).

Network Parameterisation Also referred to as the model calibration procedure, this step involves the estimation of network parameters, i.e. thermal resistance and capacities, and the calculation of the associated time constants. For existing buildings with available monitoring data, the grey-box approach is considered to combine the best of two worlds: physical insight and model structure from the white-box paradigm and parameter estimation and statistical framework from the black-box paradigm. In the absence of monitoring data, the grey-box network is trained using the simulation data produced by the white-box model so that the grey-box network could accurately replicate the results. White-box energy simulation can't be used as the simulation engine as that would result in larger calculation times. Also, a fundamental advantage of this procedure is that it can compare grey-box models on their ability to predict building thermal behaviour without any noise and measurement uncertainties. The methodology would ideally compare Reduced Order Model (ROM) outputs with measured data. In the case of our representative example, measurements were not available at the required resolution, so we used a calibrated white box model instead of measured data.

A number of studies have used statistical estimation procedures, for instance, maximum likelihood technique to compute the model parameters (Bacher and Madsen 2011). Brastein et al. (2018) used Monte Carlo sampling of parameter space to gain cognitive insight into the expected behaviour of estimation algorithms based on numerical optimisation of the mean square error between simulation and measurement data. Another study by Berthou et al. (2014) used the Interior Point algorithm that can handle large-scale non-linear optimisation problems with inequality constraints to identify the network parameters. Few other studies have implemented genetic algorithm to minimise the difference between predicted and measured consumption and estimate the network parameter values (Rouchier, Rabouille, and Oberlé 2018). None of the studies listed offer an automated procedure to identify the parameters. Hence, we used a parameter estimation tool to effectively parameterise the formulated grey box networks.

We estimated the parameters of the developed model using Continuous Time Stochastic Modelling-R (CTSM-R) tool, which uses the maximum likelihood technique and automates the estimation procedure. CTSM-R is a free, open source and cross platform tool for identifying physical models using real time series data, which can estimate embedded parameters in a continuous time stochastic state space model (CTSM-R n.d.). By using a continuous time formulation of the dynamics and discrete time measurements, the tool bridges the gap between physical and statistical modelling. It is possible to generate both pure simulation and k-step prediction estimates of the states and the outputs, filtered estimates of the states and, for nonlinear models, smoothed estimates of the states.

Model Selection The model selection procedure is based on the above identified crucial building metrics, which are used to then decide the resulting network order. For instance, a simple lumped parameter model would be sufficient in describing the system dynamics of a building which has a stable heat demand without any recurring fluctuations and a homogeneous construction. However, a building with a variable heat demand pattern and non-homogeneous construction will require separate RC networks

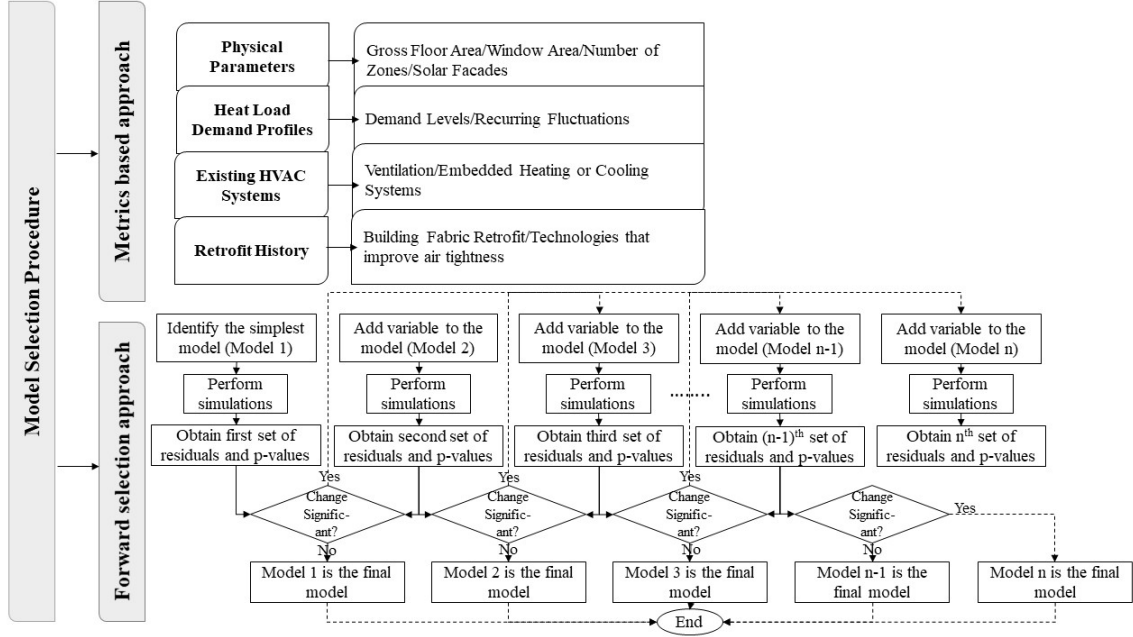


Figure 3. The process of model selection. Model selection is performed using a metrics based approach and a forward selection strategy.

for exterior and interior walls. Separate branches for exterior and interior will be able to capture the varied system dynamics. Also, a building with a floor radiant heating system would require a separate RC branch to model the floor.

After the identification of the network order using the above specified approach, we used a statistical procedure, known as forward selection strategy, to validate the network order (Fig. 3). The forward selection strategy is only used to validate the devised methodology. The forward selection process is initiated with the simplest model and subsequent iterations are performed to select more complex models (Hutter, Hoos, and Leyton-Brown 2013). This implies fitting a set of models from the simplest model to the most feasible complex model, denoted as the full model. In this approach, one adds variables to the model one at a time. At each step, each variable that is not already in the model is tested for inclusion in the model. The most significant of these variables is added to the model, as long as its p-value is below some pre-set level.

3.3.3. Model Evaluation

Model evaluation follows the process of calibration and identification and is used as a measure of the goodness of the model. This process is a means to check if a model satisfies the assumptions and gives reasonable estimates from a physical point of view. Furthermore, model deficiencies can also be identified using the evaluation process.

We implemented the likelihood ratio tests in the forward selection procedure to determine if the larger model performs significantly better than the sub-model. Alongside, we also used the residual plots to identify the suitability of each model and understand the effects not properly described by the model. Models can be compared using the residual plots of both internal temperature profiles and space heating demand. In this study, we compared the models based on the residual plots of internal temperature profiles as internal temperature (a state variable) forms a fundamental

part of the state equation. Comparison of residual plots of space heating demand would yield a similar model identification result.

4. Demonstration Cases and Discussion

Commercial buildings differ significantly in terms of the individual building function. The varied nature of operation makes it difficult to identify a suitable approach that could be used to model the commercial building stock. Moreover, detailed modelling of each and every commercial building is impractical and therefore, approaches are required that could tackle the resource intensive nature of the previous modelling techniques. The approach developed in this paper builds upon the grey-box networks and uses easily accessible building metrics to identify linkages between those metrics and grey box model parameters.

We applied the devised methodology to identify a suitable model for two buildings (a regular office building and a full service restaurant) with different building functions. The buildings selected are connected to a district heating network, which caters to the space heating demand. We selected the use case of a typical office building to reflect the influence of a stable operation schedule on the network order. To account for added complexities, we selected a restaurant as our second use case. The second use case takes into consideration the presence of fluctuations in heat demand and special thermal zones. The demonstrated use cases indicate that although the proposed methodology introduces simplifications in the modelling process, the method can produce significantly accurate results in accordance with the established standards (Coakley, Raftery, and Keane 2014). The proposed approach will find suitable applications in heat demand forecasting of buildings when the detailed geometric and non-geometric information is not available. The following sections describe the model development process for the considered case studies.

4.1. Case Study 1: Model Identification for an Office Building

The Tierney Building is part of the district heating network at University College Dublin in Ireland. This building consists of administrative offices following a 9 AM to 6 PM schedule (a typical office building). Measurements and synthetic data consisting of 15 min values over a fixed time period, are used. Synthetic data is obtained by performing simulations using EnergyPlus software. The synthetic data is required to identify the parameters of the grey box network associated with the building in consideration. The methodology would ideally compare ROM outputs with measured data. In the case of our representative example, measurements were not available at the required resolution, so we used a calibrated white box model instead of measured data. The following sections describe the building considered for model identification. The sections further provide an outline of the considered models followed on by model identification and evaluation.

4.1.1. Description of the building

The outer walls of the building consist of brick blocks insulated according to Building Regulations 2000 (R2 2000). The building rests on an insulated floor slab consisting of 30 mm insulation and 150 mm of cast concrete. The roof is flat and insulated according to Building Regulations 2000. The building spans across an area of 3400

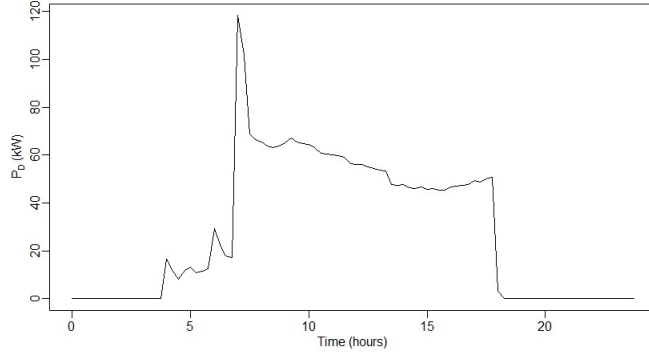


Figure 4. Heat demand data for the Tierney building.

m² approximately and is connected to a central district heating network. Measured energy consumption data is available for the entire UCD campus on a 15 min basis. The building is simulated using EnergyPlus to obtain the internal temperature profile, which is then used to calibrate the grey-box network. The temperature profile is needed to identify the parameters of the grey-box network.

4.1.2. Models identified using parametric and statistical methodologies

Models for the building are identified using the proposed methodology in this study and a statistical forward selection procedure. The results are then compared for the two procedures to map the similarities and identify any discrepancies.

Model identification is first performed using the metrics identified in the proposed methodology. Metrics associated with the building are described below:

Physical Parameters: The building occupies a gross interior floor area of 3400 m², which relates to a big number (compared to the average size of building on the UCD campus) and eventually, a complex RC network. A large area also reflects the presence of a large number of zones. Windows constitute of approximately 42% of the external wall area. The building interiors consist of a large office area with small cubicles on the ground floor while the first and second floors consist of individual offices and meeting rooms. The building doesn't have any integrated solar PV panels.

Heat Demand Profiles: The concerned building is only for administrative purposes and follows a set schedule of HVAC operation. The building is analysed based on the algorithm presented in Fig. 2. This analysis illustrates the importance of combining net floor area with peak heat demand. The heat demand (kW) is analysed at fifteen minute intervals. The yearly maximum heat demand for the building is 250 kW. Hourly heat demand patterns are analysed using the ANOVA test to establish the significance of variations in the usage patterns. No significant variations are found in the heat demand pattern of the Tierney building (low F value and high p value of 0.409 corresponding to a significance level of 0.05).

Owing to the flat power demand curve, as depicted in Fig. 4, it can be inferred that a second order RC network would accurately represent the dynamics inside the building (following the process introduced in Section 3.2.2). The building, in consideration, is a multi-floor building and each floor represents a similar usage profile. As the demand profile is stable (absence of demand fluctuations) during the day, it would be enough to separate the exterior walls from the internal environment using two heat capacities, C_i and C_e (Fig. 5). C_i represents the internal heat capacity while C_e represents the

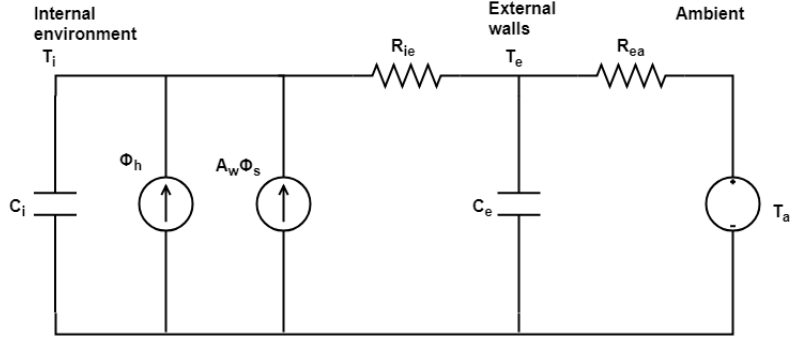


Figure 5. The implemented second order network for Tierney building and the restaurant. T_i , T_e and T_a represent the states of the internal, special zone, heater, exterior and ambient elements in the network. R_{ie} and R_{ea} represent the thermal resistances of the network. C_i and C_e represent the thermal capacities of the network. ϕ_h and ϕ_s are the heater and solar radiation flux elements. A_w represents the effective window area.

external heat capacity. Separation of interior environment from external envelope is crucial as buildings with large number of enclosed spaces (for instance, offices) often possess varying dynamics inside and around the external envelope.

Existing HVAC System: The building HVAC system consists of a simple 4-pipe fan coil unit with an air cooled chiller. No embedded HVAC systems are present in the building. Also, the building relies on natural ventilation for fresh air requirements.

Retrofit History: The building hasn't been renovated and as such, the building fabric is homogeneous across the entire building, thus, removing the possibility of any added complexities.

Based on the above metric analysis, a second order RC network should be able to express the heat dynamics associated with the building. The two capacitance in the network would separate the exterior walls from the internal mass. The formulated second order network is shown in Fig. 5.

The second procedure of model identification involves a statistical forward selection strategy. The process is initiated with the simplest model and subsequent iterations are performed to select more complex models. This implies fitting a set of models from the simplest model to the most feasible complex model, denoted as the full model. The simplest model considered for the analysis consists of only a single state variable T_i , internal temperature profile and a single thermal capacitance, representing the entire thermal mass of the building. More complex model involve the representation of heat dynamics using more than one state variable and thermal capacitance.

4.1.3. Model Evaluation

We applied the identification procedure to find a sufficient model in the set of models ranging from only one state variable and then increasing the variables one at a time. Inputs to the model are depicted in Fig. 6. T_i is the internal temperature profile in $^{\circ}C$, T_A is the ambient temperature in $^{\circ}C$, G_s is the global solar radiation in kWh/m^2 and P_D is the heat demand in kW .

The data is selected for a duration of 24 days, starting from 15th February 2017 and ending on 09th March 2017. The district heating network operates at full potential during this time period. The procedure begins with the simplest model with only one state variable. The log-likelihood values are computed at each iteration. For the Tierney building, a higher log likelihood value of 3682 is obtained for a second order model as compared to a value of 2785 for a first order model.

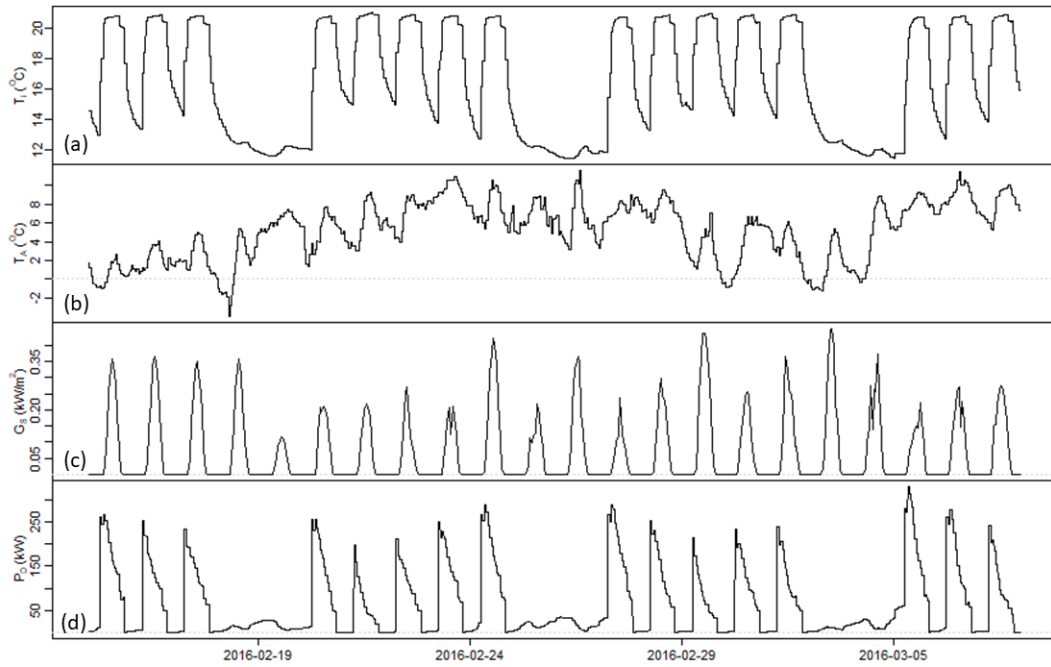


Figure 6. Plots of the input data for Tierney building. (a) T_i is the internal temperature profile in $^{\circ}C$, (b) T_A is the ambient temperature in $^{\circ}C$, (c) G_s is the global solar radiation in kWh/m^2 and (d) P_D is the heat demand in kW .

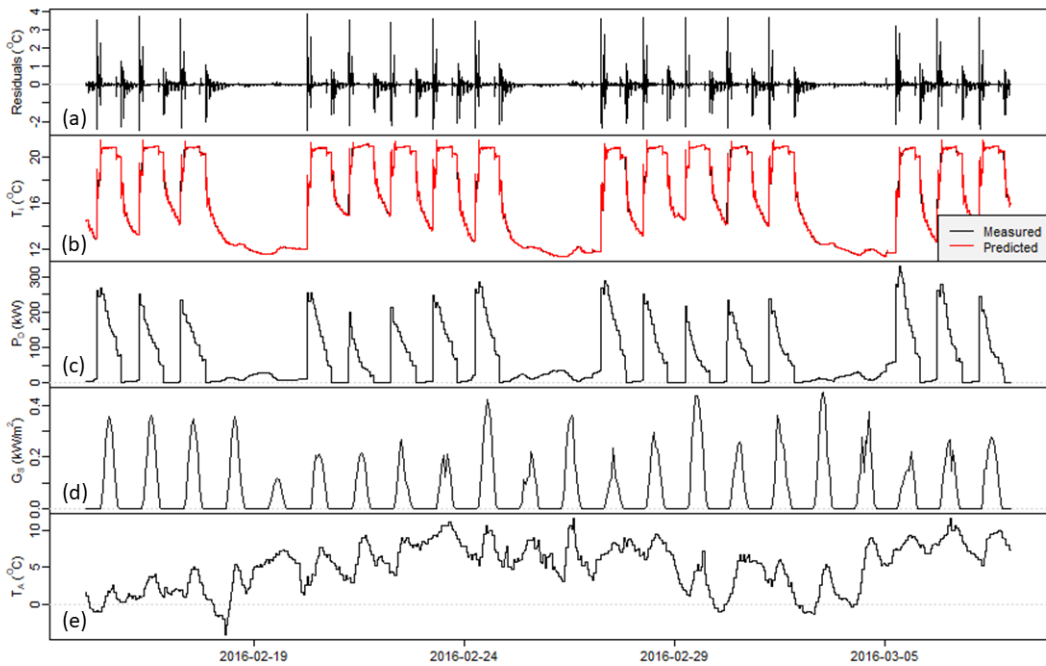


Figure 7. Residual and prediction plots for the second order model for Tierney building. (a) Residuals represent the difference in measured and predicted internal temperature profile in $^{\circ}C$, (b) T_i is the internal temperature profile in $^{\circ}C$ (measured and predicted), (c) P_D is the heat demand in kW , (d) G_s is the global solar radiation in kWh/m^2 and (e) T_A is the ambient temperature in $^{\circ}C$.

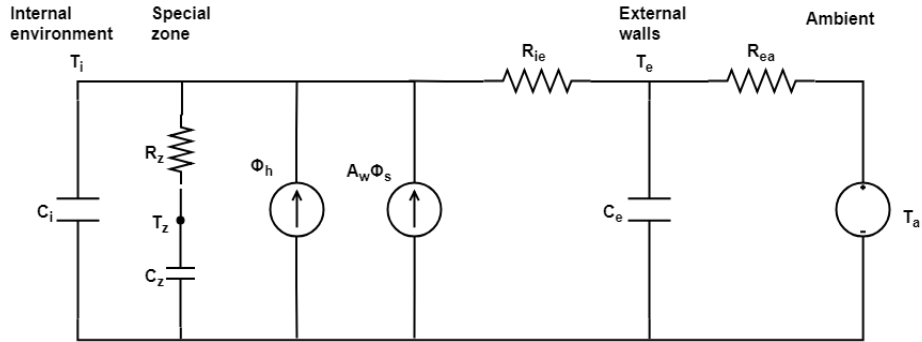


Figure 8. The implemented third order network for Tierney building and the restaurant. T_i , T_z , T_e and T_a represent the states of the internal, special zone, exterior and ambient elements in the network. R_z , R_{ie} and R_{ea} represent the thermal resistances of the network. C_i , C_z , and C_e represent the thermal capacities of the network. ϕ_h and ϕ_s are the heater and solar radiation flux elements. A_w represents the effective window area.

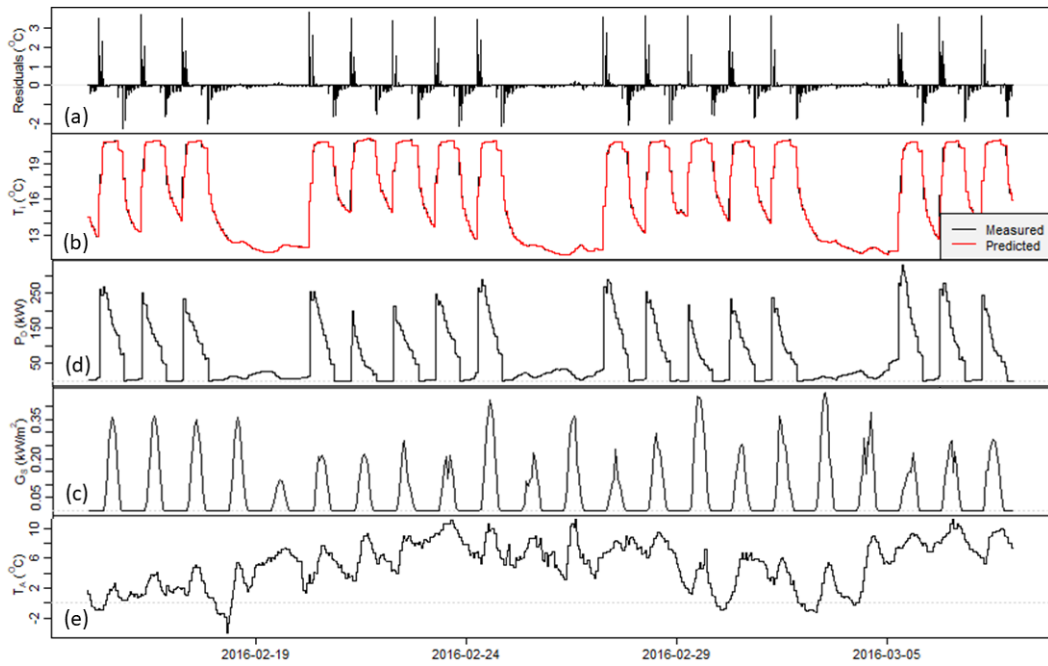


Figure 9. Residual and prediction plots for the third order model for Tierney building. (a) Residuals represent the difference in measured and predicted internal temperature profile in $^{\circ}C$, (b) T_i is the internal temperature profile in $^{\circ}C$ (measured and predicted), (c) P_D is the heat demand in kW , (d) G_s is the global solar radiation in kWh/m^2 and (e) T_A is the ambient temperature in $^{\circ}C$.

The second order model is further analysed using residual plots. Parameter estimation is performed using the input data, which is then used to predict the internal temperature profile. Estimation is done using the CTSM-R stochastic modelling tool that allows to control the range of parameter values. The residual and prediction plots along with the input data for the second order model are depicted in Fig. 7. Residuals represent the difference in measured and predicted internal temperature profile in $^{\circ}C$, T_i is the internal temperature profile in $^{\circ}C$ (measured and predicted), P_D is the heat demand in kW , G_s is the global solar radiation in kWh/m^2 and T_A is the ambient temperature in $^{\circ}C$.

The selection procedure is carried out further considering a third order model for the same building, depicted in Fig 8. The residual and prediction plots along with the input data for the second order model are depicted in Fig. 9.

Comparing the residual plots of the second and third order models, it can be seen that the residuals are not affected dramatically and can be considered to be only slightly improved, which denotes that a second order model is sufficient to represent the building. Model evaluation of the statistically chosen model relate to the fact that the considered building could be represented using a model of second order, which is similar as the order identified using parametric approach. Furthermore, the highest level of error is observed when the solar irradiance is high, which relates to the constant spikes in the residuals of zone temperature (Fig. 7 and Fig. 9). Further improvement of the model should be concentrated on the part in which solar radiation enters the building.

4.2. Case Study 2: Model Identification for a full service restaurant

The restaurant is also a part of the district heating network at University College Dublin. This building consists of a cafeteria with a large seating area. There are also a few retail shops that operate during the semester period. Measurements and synthetic data consisting of 15 min values over a fixed time period, are used. The methodology would ideally compare ROM outputs with measured data. In the case of our representative example, measurements were not available at the required resolution, so we used a calibrated white box model instead of measured data. Synthetic data is obtained by performing simulations using EnergyPlus software. The synthetic data is required to identify the parameters of the grey box network associated with the building in consideration. The following sections describe the building considered for model identification.

4.2.1. Description of the building

The outer walls of the building are constructed of brick blocks insulated according to Building Regulations 2000 (R2 2000). The building rests on an insulated floor slab consisting of 30 mm insulation and 150 mm of cast concrete. The roof is flat and insulated according to Building Regulations 2000. The building spans across an area of 1689 m^2 approximately. The building is connected to a central district heating network. The measured energy consumption data is available for the entire UCD campus on a 15 min basis. The building is simulated using EnergyPlus to obtain the internal temperature profile which is then used to calibrate the grey-box network.

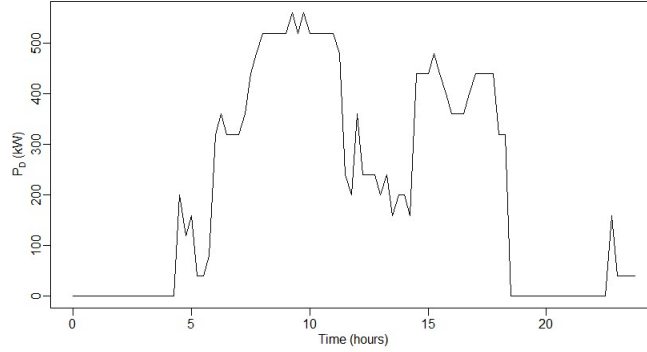


Figure 10. Heat demand data for the restaurant.

4.2.2. Models identified using parametric and statistical methodologies

Models for the building are identified using the proposed methodology in this study and a statistical forward selection procedure. The results are then compared for the two procedures to map the similarities and identify any discrepancies.

Model identification is first performed using the metrics identified in the proposed methodology. Metrics associated with the building are described below:

Physical Parameters: The building occupies a gross interior floor area of 1689 m². Windows constitute of approximately 30% of the external wall area. The building interiors consist of a large seating area on the first floor with small retail shops on the ground floor. There is also a lounge in the basement of the building where a number of international events are held all throughout the year. The building doesn't have any integrated solar PV panels.

Heat Demand Profiles: The concerned building is used for various purposes and follows a variable schedule of the HVAC operation. The building is analysed based on the algorithm presented in Fig. 2. The heat demand (kW) is analysed at fifteen minute intervals. The yearly maximum heat demand for the building is 640 kW. The ANOVA test is performed using the demand patterns of the full service restaurant. This test establishes the significance of existing variations in the heat demand patterns as the obtained F-statistic value is high (low p value of 0.034 corresponding to a significance level of 0.05). Moreover, the heat dynamics existing inside the kitchen of a restaurant are different than the dynamics prevailing in the seating spaces.

Owing to the variable demand and presence of a special zone, as depicted in Fig. 10, it can be inferred that a third order RC network would be required to accurately represent the dynamics inside the building, particularly because of the varied nature of operation (following the process introduced in Section 3.2.2). The exterior walls are separated from the internal environment using two separate thermal capacities, C_i and C_e . A special zone is added to the network with a thermal capacity of C_z . C_i , C_e and C_z are the thermal capacities of internal environment, exterior walls and the special zone respectively.

Existing HVAC System: The building HVAC system consists of a simple 4-pipe fan coil unit with an air cooled chiller. No embedded HVAC systems are present in the building. Also, the building relies on natural ventilation for fresh air requirements.

Retrofit History: The building hasn't been renovated and as such, the building fabric is homogeneous across the entire building, thus, removing the possibility of any added complexities.

Based on the above metric analysis, a third order RC network should be able to

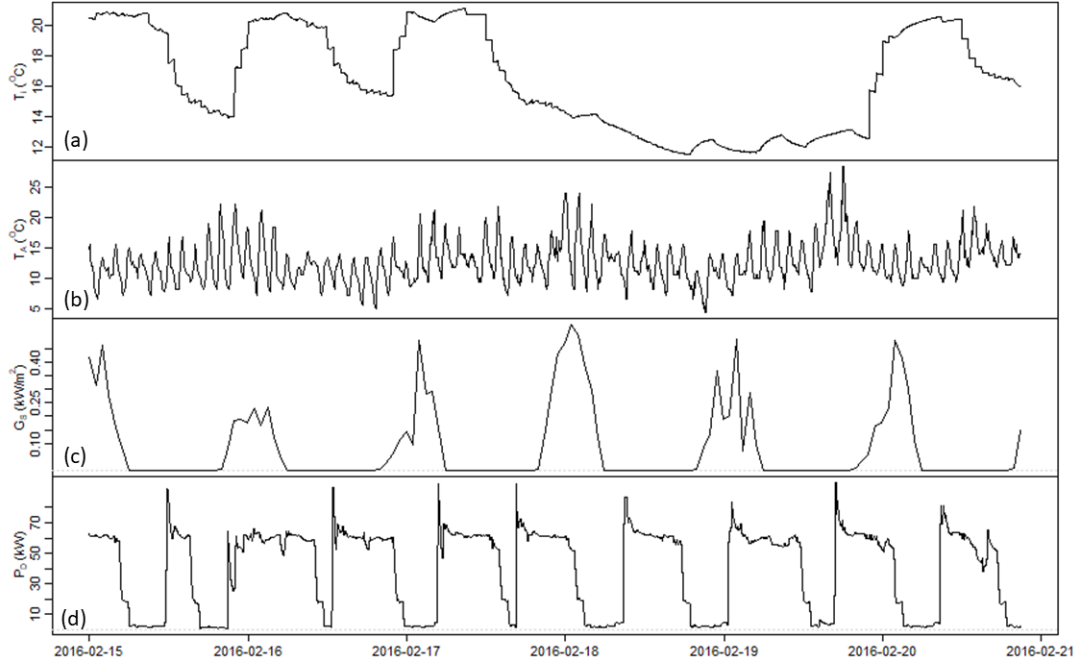


Figure 11. Plots of the input data for the full-service restaurant. (a) T_i is the internal temperature profile in $^{\circ}C$, (b) T_A is the ambient temperature in $^{\circ}C$, (c) G_s is the global solar radiation in kWh/m^2 and (d) P_D is the heat demand in kW .

express the heat dynamics associated with the building. The two capacitance in the network would separate the exterior walls from the internal mass while a third capacitance would represent the varied nature of operation.

The second procedure of model identification involves a statistical forward selection strategy, as demonstrated for the Tierney building. The process is initiated with the simplest model and subsequent iterations are performed to select more complex models. This implies fitting a set of models from the simplest model to the most feasible complex model, denoted as the full model.

4.2.3. Model Evaluation

The identification procedure is applied to find a sufficient model in the set of models ranging from only one state variable and then increasing the variables one at a time. Inputs to the model are depicted in Fig. 11. T_i is the internal temperature profile in $^{\circ}C$, T_A is the ambient temperature in $^{\circ}C$, G_s is the global solar radiation in kWh/m^2 and P_D is the heat demand in kW . The data is selected for a duration of 7 days, starting from 15th February 2017 and ending on 21st February 2017. The district heating network operates at full potential during this time period. The procedure begins with the simplest model with only one state variable. The log-likelihood values are computed at each iteration. For the restaurant, a higher log likelihood value of 4085 is obtained for a second order model as compared to a value of 2561 for a first order model.

The second order model is further analysed using residual plots. Parameter estimation is performed using the input data, which is then used to predict the internal temperature profile. The residual and prediction plots along with the input data for the second order model are depicted in Fig. 12. Residuals represent the difference in

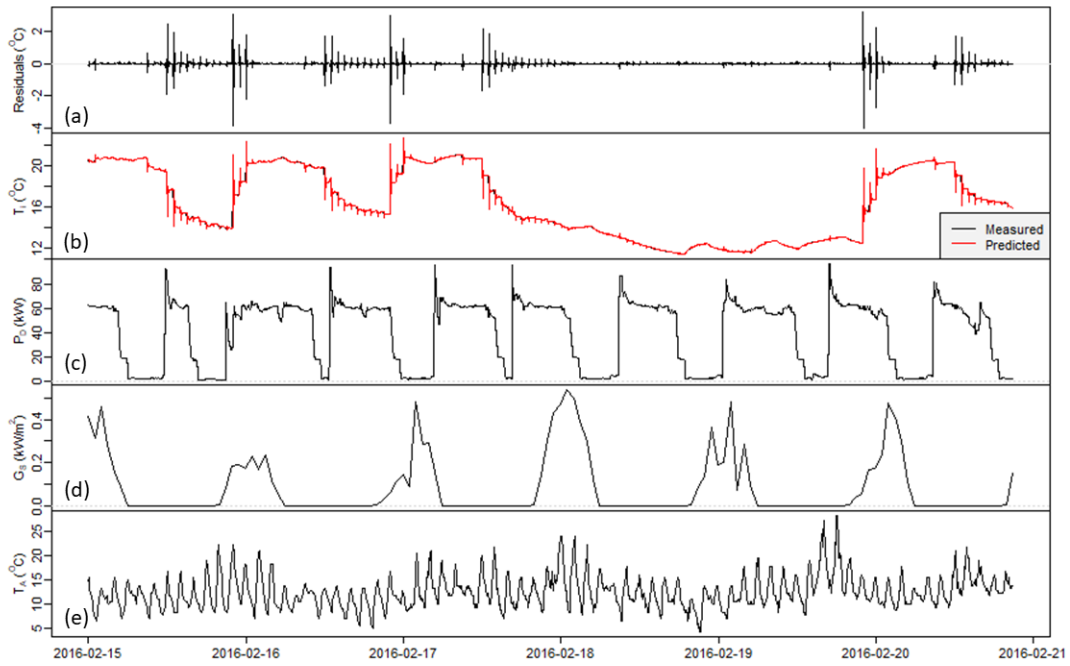


Figure 12. Residual and prediction plots for the second order model for the full-service restaurant. (a) Residuals represent the difference in measured and predicted internal temperature profile in $^{\circ}C$, (b) T_i is the internal temperature profile in $^{\circ}C$ (measured and predicted), (c) P_D is the heat demand in kW , (d) G_s is the global solar radiation in kWh/m^2 and (e) T_A is the ambient temperature in $^{\circ}C$.

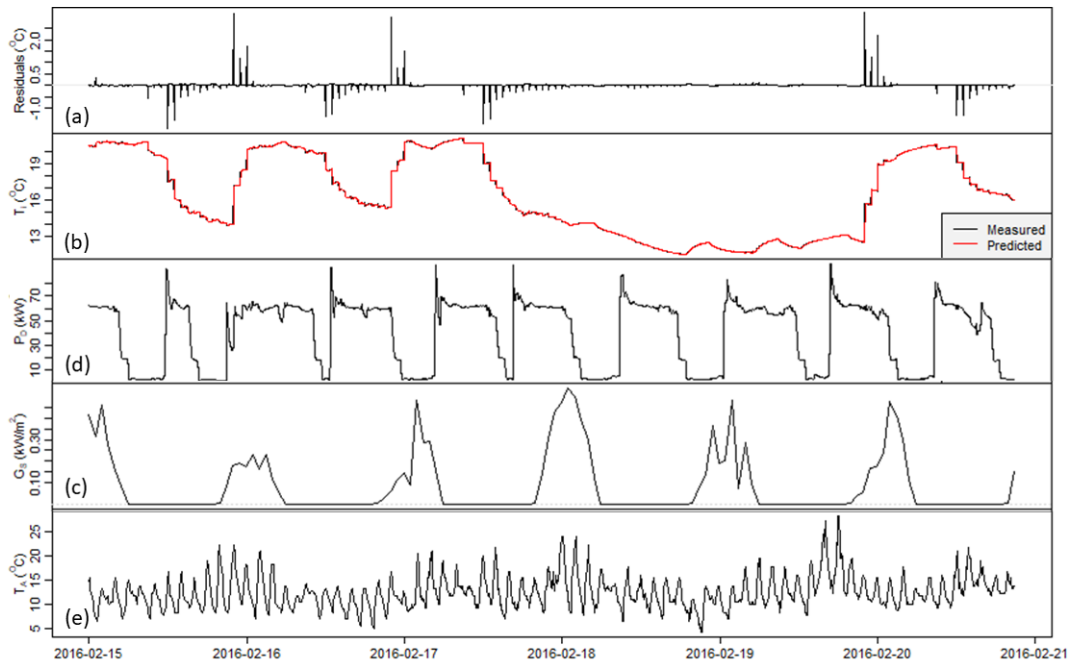


Figure 13. Residual and prediction plots for the third order model for the full-service restaurant. (a) Residuals represent the difference in measured and predicted internal temperature profile in $^{\circ}C$, (b) T_i is the internal temperature profile in $^{\circ}C$ (measured and predicted), (c) P_D is the heat demand in kW , (d) G_s is the global solar radiation in kWh/m^2 and (e) T_A is the ambient temperature in $^{\circ}C$.

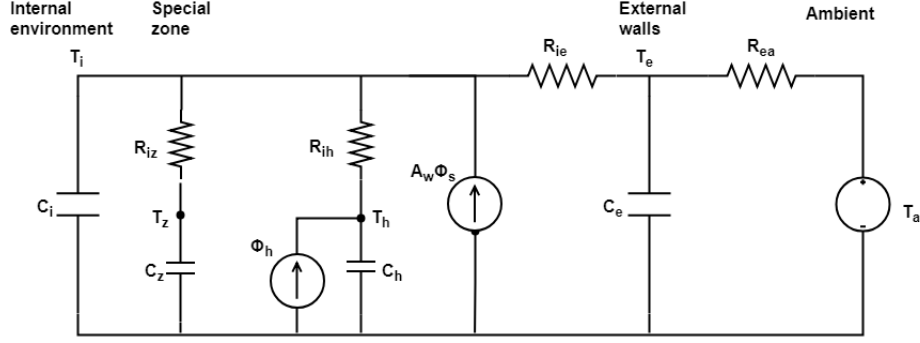


Figure 14. The implemented fourth order network for the restaurant. T_i , T_z , T_h , T_e and T_a represent the states of the internal, special zone, heater, exterior and ambient elements in the network. R_z , R_{ih} , R_{ie} and R_{ea} represent the thermal resistances of the network. C_i , C_z , C_h and C_e represent the thermal capacities of the network. ϕ_h and ϕ_s are the heater and solar radiation flux elements. A_w represents the effective window area.

measured and predicted internal temperature profile in $^{\circ}C$, T_i is the internal temperature profile in $^{\circ}C$ (measured and predicted), P_D is the heat demand in kW , G_s is the global solar radiation in kWh/m^2 and T_A is the ambient temperature in $^{\circ}C$.

The selection procedure is carried out further considering a third order model for the same building. The residual and prediction plots along with the input data for the second order model are depicted in Fig. 13.

Comparing the residual plots of the second and third order models, it can be seen that the residuals are significantly improved, which denotes that a third order model

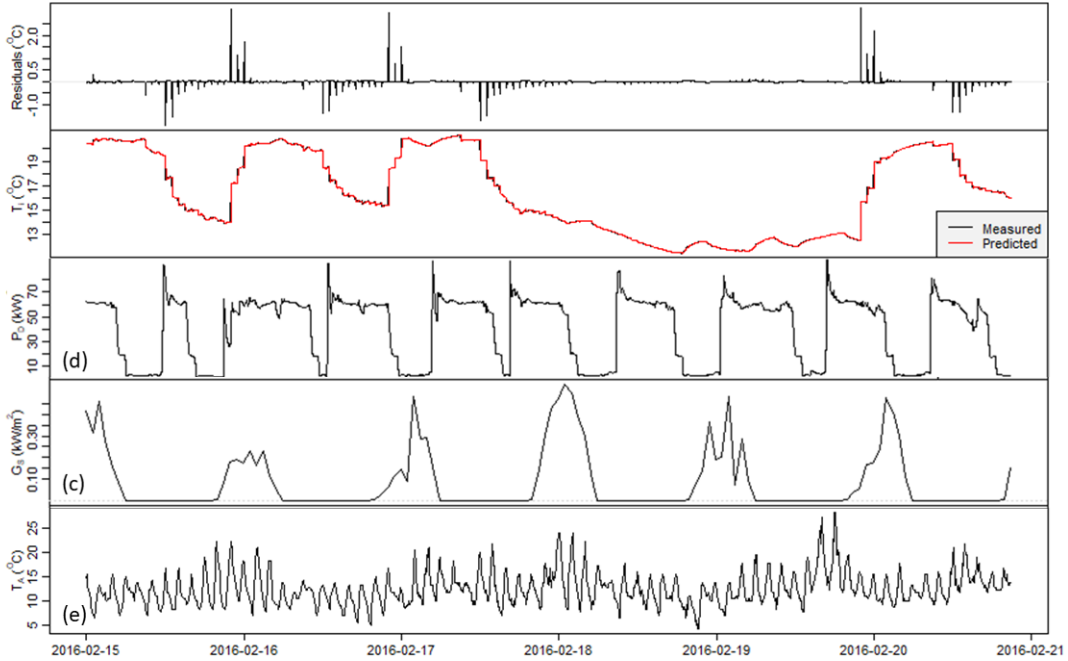


Figure 15. Residual and prediction plots for the fourth order model for the full-service restaurant. (a) Residuals represent the difference in measured and predicted internal temperature profile in $^{\circ}C$, (b) T_i is the internal temperature profile in $^{\circ}C$ (measured and predicted), (c) P_D is the heat demand in kW , (d) G_s is the global solar radiation in kWh/m^2 and (e) T_A is the ambient temperature in $^{\circ}C$.

is required to represent the building dynamics. Model evaluation of the statistically chosen model relate to the fact that the considered building could be represented using a model of third order, which is similar as the order identified using parametric approach. A further extension of the model to fourth order model, depicted in Fig. 14, is formulated for the building. The residuals are slightly improved as depicted in Fig. 15. However, the change in residuals is found to be insignificant. Furthermore, the highest level of error is observed when the solar irradiance is high, which relate to the constant spikes in the residuals of zone temperature (Fig. 12, Fig. 13 and Fig. 15). Further improvement of the model should be concentrated on the part in which solar radiation enters the building.

5. Conclusions and Future Work

Identification of a suitable model to represent a building's dynamics is often challenging. Furthermore, the model development process is resource and time intensive. Simulation tools play a significant role to effectively find the optimum balance between comfort, cost and efficiency. In order for BEPS models to be used with any degree of confidence, it is necessary that the existing model closely represent the actual behaviour of the building under study. Grey-box models offer an advantage to balance complexity and accuracy according to the application. However, due to the varied nature of operation of buildings, these networks have to be modified to represent different buildings. This paper introduces a generalisation approach to grey box model development using easily identifiable building metrics. The devised approach would reduce the complexities associated with the identification of the network order while maintaining the desire level of accuracy. In addition, the approach can provide additional insight and information to designers considering novel alternative design approaches, where prior information may not be readily available.

The task of modelling energy in buildings becomes more and more challenging as the associated scale (city/district/urban) changes. It is often difficult to gather geometric and non-geometric data associated with each and every building at the urban scale. Detailed building models often fail to reproduce the model structure and thus lack scalability. The approach devised in this study will aid in energy modelling of buildings at an urban scale as the approach makes use of easily identifiable building metrics. This approach is based on a generalised structure and hence, will be able to reproduce models for different buildings.

We identified a procedure for identification of the most suitable models for the heat dynamics of a building. The procedure is devised on the basis of measured and simulated data for two buildings with different operational characteristics. The model identification procedure is done using two methods, namely, the building metric approach as devised in this study and a statistical forward selection procedure. The metrics used to develop the generalisation framework includes physical parameters, operational characteristics, existing HVAC systems and renovation history. The implemented statistical procedure is based on likelihood-ratio testing combined with a forward selection strategy. The results of the identification procedure are evaluated and compared for both approaches. The evaluation reveals that the network order obtained using the building metric approach is similar to the one obtained using statistical procedure. This indicates that the proposed methodology can be used to generalise the idea of obtaining grey box networks for different commercial buildings. Furthermore model deficiencies are pointed out, from which further advancement of the model

should be pursued.

It has been shown that the method is able to provide rather detailed knowledge of the heat dynamics of the building. The calculated parameters of the network, such as thermal resistances and capacitance, give an indication of the heat storage capabilities of the building. The developed generalisation approach has the potential applications in several domains of energy research for the building sector. In the absence of the detailed information associated with a building, the building is often reduced to a black box and there is no information of the system dynamics. The developed approach uses grey box networks to identify the energy signature of buildings that also includes details about the energy dynamics inside the building. The knowledge of dynamics is essential in the identification of cost effective energy retrofits for any building in consideration. Alongside, simplified networks identified using the devised approach are also crucial in the integration of thermal networks with other energy vectors, for instance, electricity.

Whole building energy modelling and simulation tools are increasingly being used for detailed performance analysis and for evaluation of multiple retrofit design options. However, the models typically involve several hundreds of input parameters and processes that are uncertain in early stages of design, and are not fully understood until after retrofit installation and commissioning. A future work to this study could involve the quantification of uncertainties associated with the parameters of developed grey box networks.

Funding

This publication has emanated from research conducted with the financial support of Science Foundation Ireland under the SFI Strategic Partnership Programme Grant Number SFI/15/SPP/E3125. The opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of Science Foundation Ireland.

Disclosure statement

There are no known conflicts of interest associated with this publication.

Nomenclature

C	Capacitance of a grey box thermal network in MJ/K
$c(T)$	Specific heat capacity in MJ/K.kg
G_S	Global solar radiation in kWh/sq. m
m	Mass of an object in kg
p	p-values representing the statistical significance
P_D	Heat demand in kW
Q	Heat transfer across a medium in W
R	Resistance of a grey box thermal network in K/W
RC	Resistance-Capacitance thermal grey box network
T_A	Ambient temperature in °C
T_i	Internal temperature profile in °C

References

2000. *The Building Regulations 2000*. Technical Report. <http://www.housing.gov.ie/>.
2017. “Office of energy efficiency and renewable energy.” Date last accessed 24-04-2017, <http://energyplus.net>.
2017. “Transition System Simulation Tool.” Date last accessed 24-04-2017, <http://trnsys.com>.
- Amara, Fatima, Kodjo Agbossou, Alben Cardenas, Yves Dubé, and Souso Kelouwani. 2015. “Comparison and Simulation of Building Thermal Models for Effective Energy Management.” *Smart Grid and Renewable Energy* 6 (April): 95–112.
- Bacher, Peder, and Henrik Madsen. 2011. “Identifying suitable models for the heat dynamics of buildings.” *Energy and Buildings* 43 (7): 1511–1522. <https://www.sciencedirect.com/science/article/pii/S0378778811000491>.
- Belić, Filip, Željko Hocenski, and Dražen Slišković. 2016. “Thermal Modeling of Buildings with RC Method and Parameter Estimation.” In *International Conference on Smart Systems and Technologies*, 19–25. IEEE.
- Berthou, Thomas, Pascal Stabat, Raphael Salvazet, and Dominique Marchio. 2014. “Development and validation of a gray box model to predict thermal behavior of occupied office buildings.” *Energy and Buildings* 74: 91–100. <https://www.sciencedirect.com/science/article/pii/S0378778814000760>.
- Biyik, Emrah, Mustafa Araz, Arif Hepbasli, Mehdi Shahrestani, Runming Yao, Li Shao, Emmanuel Essah, et al. 2017. “A key review of building integrated photovoltaic (BIPV) systems.” *Engineering Science and Technology, an International Journal* 20 (3): 833–858. <https://www.sciencedirect.com/science/article/pii/S2215098616309326#!>
- Brastein, O. M., D. W.U. Perera, C. Pfeifer, and N. O. Skeie. 2018. “Parameter estimation for grey-box models of building thermal behaviour.” *Energy and Buildings* 169: 58–68. <https://www.sciencedirect.com/science/article/pii/S0378778817331791>.
- Braun, James E., and Nitin Chaturvedi. 2002. “An inverse gray-box model for transient building load prediction.” *HVAC and R Research* 8 (1): 73–99. <http://www.tandfonline.com/doi/abs/10.1080/10789669.2002.10391290>.
- Chen, Qingyan. 2009. “Ventilation performance prediction for buildings: A method overview and recent applications.” *Building and Environment* 44 (4): 848–858. <https://www.sciencedirect.com/science/article/pii/S0360132308001510>.
- Clarke, J.A. 2001. *Energy Simulation in Building Design*. Routledge. <https://www.taylorfrancis.com/books/9780080505640>.
- Coakley, Daniel, Paul Raftery, and Marcus Keane. 2014. “A review of methods to match building energy simulation models to measured data.” *Renewable and Sustainable Energy Reviews* 37: 123–141. <https://www.sciencedirect.com/science/article/pii/S1364032114003232#!>
- Crawley, Drury B., Jon W. Hand, Michaël Kummert, and Brent T. Griffith. 2008. “Contrasting the capabilities of building energy performance simulation programs.” *Building and Environment* 43 (4): 661–673. <http://www.sciencedirect.com/science/article/pii/S0360132306003234>.
- Crawley, Michael J. 2007. *The R book*.
- CTSM-R. n.d. “CTSM-R - Continuous Time Stochastic Modelling for R.” Accessed 2019-01-24. <http://ctsm.info/>.
- De Coninck, Roel, Fredrik Magnusson, Johan Åkesson, and Lieve Helsen. 2016. “Toolbox for development and validation of grey-box building models for forecasting and control.” *Journal of Building Performance Simulation* 9 (3): 288–303. <http://www.tandfonline.com/doi/full/10.1080/19401493.2015.1046933>.
- De Smith, Michael. 2018. *Statistical Analysis Handbook A Comprehensive Handbook of Statistical Concepts, Techniques and Software Tools 2018 Edition*. <https://www.statsref.com/StatsRefSample.pdf>.
- Deng, K., P. Barooah, P. G. Mehta, and S. P. Meyn. 2010. “Building thermal model reduction via aggregation of states.” In *Proceedings of the 2010 American Control Conference*, June, 5118–5123.

- Deru, M., and P. Torcellini. 2005. *Standard Definitions of Building Geometry for Energy Evaluation*. Technical Report October.
- Deru, Michael, Kristin Field, Daniel Studer, Kyle Benne, Brent Griffith, Paul Torcellini, Bing Liu, et al. 2011. *U.S. Department of Energy commercial reference building models of the national building stock*. Technical Report. <https://www.nrel.gov/docs/fy11osti/46861.pdf> http://digitalscholarship.unlv.edu/renew_pubs/44.
- Djuric, Natasa, Vojislav Novakovic, Johnny Holst, and Zoran Mitrovic. 2007. "Optimization of energy consumption in buildings with hydronic heating systems considering thermal comfort by use of computer-based tools." *Energy and Buildings* 39 (4): 471–477. <https://www.sciencedirect.com/science/article/pii/S0378778806002210>.
- Fung, Tady Y Y, and H. Yang. 2008. "Study on thermal performance of semi-transparent building-integrated photovoltaic glazings." *Energy and Buildings* 40 (3): 341–350. <https://www.sciencedirect.com/science/article/pii/S0378778807000941>.
- Fux, Samuel F., Araz Ashouri, Michael J. Benz, and Lino Guzzella. 2014. "EKF based self-adaptive thermal model for a passive house." *Energy and Buildings* 68: 811–817.
- Goyal, Siddharth, and Prabir Barooah. 2012. "A method for model-reduction of non-linear thermal dynamics of multi-zone buildings." *Energy and Buildings* 47: 332–340.
- Harb, Hassan, Neven Boyanov, Luis Hernandez, Rita Streblov, and Dirk Müller. 2016. "Development and validation of grey-box models for forecasting the thermal response of occupied buildings." *Energy and Buildings* 117: 199 – 207.
- Heo, Y., R. Choudhary, and G. A. Augenbroe. 2012. "Calibration of building energy models for retrofit analysis under uncertainty." *Energy and Buildings* 47: 550–560.
- Hutter, Frank, Holger H. Hoos, and Kevin Leyton-Brown. 2013. "Identifying key algorithm parameters and instance features using forward selection." In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Vol. 7997 LNCS, 364–381. Springer, Berlin, Heidelberg. http://link.springer.com/10.1007/978-3-642-44973-4_40.
- International Energy Agency. 2017. *Tracking Clean Energy Progress 2017*. Technical Report.
- Ji, Ying, Peng Xu, Pengfei Duan, and Xing Lu. 2016. "Estimating hourly cooling load in commercial buildings using a thermal network model and electricity submetering data." *Applied Energy* 169: 309–323. <https://www.sciencedirect.com/science/article/pii/S0306261916301593>.
- Li, X., J. Wen, and E. W. Bai. 2015. "Building energy forecasting using system identification based on system characteristics test." In *2015 Workshop on Modeling and Simulation of Cyber-Physical Energy Systems (MSCPES)*, April, 1–6.
- Li, Xiwang, and Jin Wen. 2014. "Review of building energy modeling for control and operation." *Renewable and Sustainable Energy Reviews* 37: 517 – 537.
- Lin, Yashen, Timothy Middelkoop, and Prabir Barooah. 2012. "Issues in identification of control-oriented thermal models of zones in multi-zone buildings." In *Proceedings of the IEEE Conference on Decision and Control*, 6932–6937.
- Louis G. Birta Gilbert Arbez. 2013. *Modelling and simulation*. Simulation Foundations, Methods and Applications. London: Springer London. <http://link.springer.com/10.1007/978-1-4471-2783-3> <http://linkinghub.elsevier.com/retrieve/pii/S0304380096019163>.
- Ma, Jingran, Joe Qin, Timothy Salsbury, and Peng Xu. 2012a. "Demand reduction in building energy systems based on economic model predictive control." *Chemical Engineering Science* 67 (1): 92–100.
- Ma, Zhenjun, Paul Cooper, Daniel Daly, and Laia Ledo. 2012b. "Existing building retrofits: Methodology and state-of-the-art." *Energy and Buildings* 55: 889–902. <https://www.sciencedirect.com/science/article/pii/S0378778812004227>.
- Mejri, Olfa, Elena Palomo Del Barrio, and Nadia Ghrab-Morcos. 2011. "Energy performance assessment of occupied buildings using model identification techniques." *Energy and Buildings* 43 (2): 285–299.
- Mustafaraj, G., G. Lowry, and J. Chen. 2011. "Prediction of room temperature and relative humidity by autoregressive linear and nonlinear neural network models for an open office."

- Energy and Buildings* 43 (6): 1452–1460.
- Perera, D. W.U., Carlos F. Pfeiffer, and N. O. Skeie. 2014. “Modelling the heat dynamics of a residential building unit: Application to Norwegian buildings.” *Modeling, Identification and Control* 35 (1): 43–57. <http://www.mic-journal.no/ABS/MIC-2014-1-4.asp>.
- Pérez, Luis, Jos Ortiz, and Christine Pout. 2008. “A review on buildings energy consumption information.” *Energy and Buildings* 40 (3): 394 – 398.
- Remmen, Peter, Moritz Lauster, Michael Mans, Marcus Fuchs, and Dirk Müller. 2017. “TEASER : an open tool for urban energy modelling of building stocks.” *Building Performance Simulation* 1493.
- Reynders, G., J. Diriken, and D. Saelens. 2014. “Quality of grey-box models and identified parameters as function of the accuracy of input and observation signals.” *Energy and Buildings* 82: 263–274. <https://www.sciencedirect.com/science/article/pii/S0378778814005623>.
- Rouchier, Simon, Mickaël Rabouille, and Pierre Oberlé. 2018. “Calibration of simplified building energy models for parameter estimation and forecasting: Stochastic versus deterministic modelling.” *Building and Environment* 134: 181–190. <https://hal.archives-ouvertes.fr/hal-01739625>.
- Wang, Shengwei, and Xinhua Xu. 2006. “Parameter estimation of internal thermal mass of building dynamic models using genetic algorithm.” *Energy Conversion and Management* 47 (13-14): 1927–1941. <https://www.sciencedirect.com/science/article/pii/S0196890405002311>.
- Wang, Sixuan, Gabriel Wainer, Rhys Goldstein, and Azam Khan. 2013. “Solutions for Scalability in Building Information Modeling and Simulation-Based Design.” In *SimAUD 2013 Conference Proceedings: Symposium on Simulation for Architecture and Urban Design*, 1–8.
- Zhao, Hai Xiang, and Frédéric Magoulès. 2012. “A review on the prediction of building energy consumption.” *Renewable and Sustainable Energy Reviews* 16 (6): 3586–3592. <https://www.sciencedirect.com/science/article/pii/S1364032112001438>.
- Zhou, Qiang, Shengwei Wang, Xinhua Xu, and Fu Xiao. 2008. “A grey-box model of next-day building thermal load prediction for energy-efficient control.” *International Journal of Energy Research* .