Energy and Buildings 74 (2014) 91–100

Contents lists available at ScienceDirect

Energy and Buildings

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

Development and validation of a gray box model to predict thermal behavior of occupied office buildings

Thomas Berthou [a,b,*], Pascal Stabat [a], Raphael Salvazet [b], Dominique Marchio [a]

[a] Mines ParisTech, Centre for Energy efficiency of Systems, ARMINES, Paris, France
[b] Veolia Environment Research and Innovation, Limay, France

ARTICLE INFO

Article history:
Received 6 December 2013
Received in revised form 17 January 2014
Accepted 22 January 2014

Keywords:
Gray-box models
Parameter estimation
Load prediction
Heating and cooling
Sensitivity analysis

ABSTRACT

Due to the development of energy performance contracting and the needs for peak electric demand reduction, the interest for building energy demand prediction is renewed. Gray-box models are a solution for energy demand prediction. However, it is still difficult to find the best level of model complexity and the good practices for the training phase. Since models' order and parameter identification method have a strong impact on the forecasting precision and are not intuitive, a comparative design approach is used to find the best model architecture and an adequate methodology for improving the training phase. The gray box models are compared on their ability to forecast heating and cooling demands and indoor air temperature. An objective function is proposed aiming to minimize both power and indoor temperature prediction errors. Moreover, for each model, several training period durations are tested. First, this study shows that a R6C2 (second order model) model is adapted to predict the building thermal behavior. Furthermore, the best fits are obtained with two weeks of data for the identification process. Second, a sensitivity analysis using total Sobol index calculation leads to validate the objective function and identify the most important parameters for prediction.

1. Introduction

Today's buildings move toward low-energy standards but the buildings' renewal rate is very low (for instance, 1% a year in France [1]) and building annual energy consumption remains high (for instance at 209 kWh/m^2 on average in France [2]). Therefore, it is essential to propose solutions to reduce consumption on medium to high-energy buildings. Potentially, significant consumption reduction can be reached by using smart control strategies of heating and cooling systems.

Another issue is the impact of electric heating systems on the peak demand of the electricity grid. Indeed, electric heaters were supported in France by low electricity prices, and as a result, demand may exceed supply during very cold days. Building's thermal inertia and time-of-use electricity tariffs can be used to reduce the stress levels on the electricity network. But, to improve building control and to anticipate high price periods, precise and robust predictive models are needed.

Multi-zone building simulation tools based on physical knowledge (EnergyPlus or TRNsys) are efficient but they need very detailed data on the building characteristics. For this reason, this type of simulation is time consuming to parameterize [3].

A second type of simulation is the statistical or black box approach. This approach is adapted if limited information is known about the building. It implies to choose a mathematical function (i.e. polynomial, ARMAX, transfer function) which may represent the building thermal behavior and uses measured data to identify its parameters [4–6]. The main difficulty is to represent nonlinear phenomena such as power saturation. In this purpose, nonlinear black box models are available such as artificial neural network (ANN) [7], but their ability to predict building behavior in the case of new control strategies not present during the learning phase is not demonstrated. For example, it has been shown that black box models are not able to forecast the response to load shedding if there is no load shedding in the learning data [8].

Gray box models combine both approaches. A very simple physical model (often mono-zone) is used and its parameters are identified with measured data. Compared to black box models, gray box models better predict the building thermal behavior in the case of new control strategies [8]. Many designs of gray box models are available; however, it remains difficult to choose the best model structure and the methodology for the learning process.

2. State of the art on gray box building modeling

Many papers deal with gray box modeling but only a few of them details the best practices for the learning process. Fux et al. [9] compare four building models (1- to 4-order) on their ability

* Corresponding author at: Mines ParisTech, CES, 60, Boulevard Saint-Michel, Paris 75005, France. Tel.: +33 681 696 607..
E-mail address: t.berthou@gmail.com (T. Berthou).

0378-7788/$ – see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.enbuild.2014.01.038

Table 1
Examples of data size used for the identification process.

Reference	Type of inverse model	Data for identification process	Time step
Madsen and Holst [13]	Gray box (2nd order)	4 days	10 min
Coley and Penman [14]	Gray box (2nd order)	10 days	30 min
Braun and Chaturvedi [15]	Gray box (8th order)	7 to 14 days	1 h
Hazyuk et al. [16]	Gray box (2nd order)	60 days	1 min

to forecast indoor temperature. They use an unoccupied period of 12 days to identify the models' parameters and show that a R1C1 model is efficient to forecast indoor temperature of a residential building. Bacher and Madsen [10] compare 10 gray box models (2- to 5-order) on their ability to predict indoor temperature. In their study, the heat flux is an input chosen to excite the model on a large frequency domain regardless the indoor temperature. The reference data are measured on a single floor unoccupied building. This methodology cannot be applied in most of the real buildings because the heat flux is controlled via the temperature set-point. Palomo et al. [11] use simplified models (2- to 6-order models) to represent a multi-zone individual building. The authors conclude that, for the tested individual building, a second order model allows a suitable prediction of the daily energy consumption but a 4-order model is needed for high quality indoor air temperature and heating power prediction. Mejri et al. [12] compared gray box models from order 1 to 5 and show that increasing the model order beyond 2 does not lead to a significant improvement and could even create unreliable physical results. The tests were executed on a small single-floor office building.

Table 1 [13–16] gives an overview of the diversity found in the literature concerning data size for the identification process. Recommended values vary from few days to several weeks. This review illustrates a diversity of conclusions on gray box models. Nevertheless, the authors globally agreed to choose a second order model as the simplest building model. In this paper, the model structure is discussed using realistic data coming from simulated building with variable occupancy profile and ventilation set-points and using a sensibility analysis method. Moreover, the identification method is rarely discussed in most of these studies and could deeply affect the performances of the models. In this paper, some good practices for identification are proposed.

3. Methodology

In this study we propose to compare four gray box models on their ability to predict both heating and cooling demands and indoor air temperature. All the physical parameters, the occupancy heat gains and the ventilation mass flow rate are identified thanks to data presented in Table 2. Synthetic learning data have been built using a multizone building simulation (TRNsys). This solution has been chosen since it can compare gray box models on their ability to predict building thermal behavior without any noise and measurement uncertainties. Weather data used for testing gray box models in prediction are the same as those used to generate the synthetic data. In practice, these data are weather forecast coming from a complex atmospheric model. The solar radiation from the weather file is pretreated in order to calculate both solar gains on the wall and through the windows to be used directly as input in the gray box models. This pretreatment is presented here after.

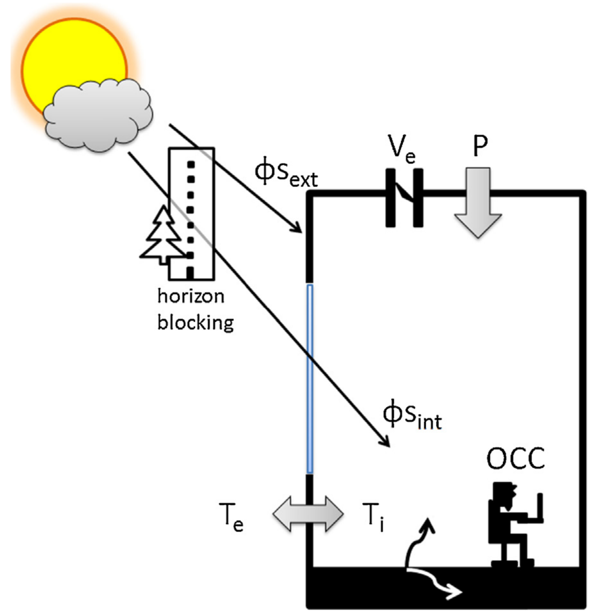

Fig. 1. Physical representation of main inputs in the simplified building model.

3.1. Selection of the semi-physic (gray box) models

Fig. 1 presents the main physical solicitations which impact the energy demand and the indoor temperature. These physical relations can be represented by RxCy networks. The model inputs and outputs are presented in Table 2.

All the gray box building models are presented as thermal network. An RxCy model has *x* resistances and *y* capacitances.

The R3C2 model is one of the simplest physical building models found in the literature adapted to buildings with constant airflow ventilation. The R4C2 model is an extension of the R3C2 model with a supplementary resistance used to characterize variable airflow ventilation. This reference model is used as a base for setting more complex models. This model has the disadvantage not to be able to take into account solar flux coming on external walls. This is not problematic in winter when this flux has a very small impact on the load but it might be in summer when this flux has a real impact on the load.

To better take into account solar gains, a R6C2 model is proposed. Indeed, the addition of two specific nodes (T_s and T_h) enables to split each solar flux (Φs_{int} and Φs_{ext}) in two parts: For the solar flux transmitted through the windows (Φs_{int}), one part hits directly the wall capacitance and the other one impacts directly the air capacitance. This repartition is determined by the value of R_i and R_s. The second part represents the solar flux coming on light furniture which has a fast impact on indoor air temperature compared to solar flux coming on heavy walls. For the solar flux on the external walls (Φs_{ext}), one part hits the wall capacitance through

Table 2
Description of model inputs and outputs.

	Name	Description
Outputs	T_i	Indoor temperature (°C)
	P	Thermal power (W)
Inputs	OCC	Occupancy profile [0 to 1]
	V_e	Ventilation set point [0 or 1]
	T_c	Temperature set-point (°C), the regulation is supposed to be ideal
	Φs_{int}	Solar gain on the walls (W)
	Φs_{ext}	Solar gain transmitted through the windows (W)