Methodologies for building envelope and whole building performance assessment

Hans Bloem[1], Maria-Jose Jimenez[2], Irati Uriarte[3] and Paul Baker[4]

[1] INIVE-DYNASTEE, Belgium,
[2] CIEMAT PSA, Spain,
[3] UPV/EHU Bilbao, Spain,
[4] GCU Glasgow, United Kingdom.

Abstract. The paper will present the analysis of high quality experimental data from a test device by means of two different dynamic methodologies. The aim of this testing and analysis approach is to train people in performing the experimental work on-site for the development of a skill that gives confidence in reporting results for building energy performance assessment. The expertise, that forms the basis for this paper comes from several EU research projects on building energy performance evaluation. A building physics perspective based on physical knowledge of heat transfer combined with advanced statistical techniques are essential for the success. The building physicist needs the statistician and vice versa. The work supports the auditor's need in the framework of building energy labelling. The increased complexity of the analysis steps demonstrates clearly that a proper assessment needs a well-defined approach that offers the application of different analysis techniques, such as discrete or continuous time methods. These steps will be briefly presented in the paper. For doing so, several high quality data series are made available. The paper will discuss the test environment, the creation of high quality data series that represents the increased complexity of going from assessing wall thermal properties to a whole building structure heat loss coefficient. Two methods are discussed and applied to experimental data in more detail. The lumped parameter method (LORD) and the continuous time method (CTSM-R), will be demonstrated on excited indoor temperature and co-heating data. Uncertainty and error sources in the whole evaluation process are presented as well as differences between these two methods which are explained on the same data. The experience gained during almost ten years of training to about 200 students, researchers and building energy professionals has convinced the authors that the approach presented is highly recommended to develop the necessary skill.

1. INTRODUCTION

Careful examination of energy consumption in the building sector, which is about 39% (2019) of the final energy consumption in EU-28 is needed in order to identify the specific areas for energy savings. Due to improved insulation levels in buildings the potential for energy savings shifts to the more dynamic energy use sectors such as gains from appliances, high energy demand and consumer behaviour. Today, more and more data related to building and building components originate from outdoor testing under time-varying and dynamic conditions, or from real life use of buildings. Dynamic evaluation methods are techniques to analyse time series of data related to dynamic processes and to identify typical parameters of the physical processes for evaluation.

The objective of training is to develop a common methodology to assess the thermal characteristics of building components and to assess whole building energy performance. In addition training should bridge the gap between expertise from both physical and mathematical/statistical analysis and modelling practice. The aim is also to gain a better understanding of the possible approaches to achieve a satisfactory analysis for the building energy performance indicator. The goal is not to promote a specific analysis or simulation tool but rather to transfer the knowledge on a common methodology. Quality benchmark data, as a reliable reference and presented below, is required to develop methods and models.

[1] hans.bloem@inive.org

Research on the energy performance of buildings is concerned with various levels of complexity, ranging from the impact of weather and climate to user behaviour in occupied buildings. In this general context, it can be divided in to three major areas: 1) building components, 2) test cells and unoccupied buildings in real climate and 3) occupied buildings. This paper deals with high quality data from outdoor experiments of a test box, addressing 2 out of these 3 areas.

The development of dedicated software tools to identify thermal parameters from physical systems has progressed hand in hand with the increasing processing speed of computing hardware. Software tools like CTSM-R [1], LORD [2] or the System Identification Toolbox in the MATLAB environment [3] are good examples of the developments over the past 25 years.

1.1 How to obtain results using different models and methods from experimental data?

System identification is a well-developed discipline, and is considered as a systems science discipline. Many textbooks are available; recommended is [4]. Here it is used to estimate certain physical parameters of a dynamic system. A wide range of system identification techniques is now being applied to the analysis problems involved with estimation of thermal characteristic properties of buildings and building components, which is an important and interesting area [5]. Specific parameter identification techniques enable the assessment of unknown thermal parameters in building physical systems and in particular those systems that have dynamic or non-linear behaviour (such as ventilated facades). Applying system identification techniques on physical systems requires throughout knowledge of the physical system [6]. For buildings it is important to know what the impact is of cold-bridges, corner effects, etc. The researchers goal is to estimate physical parameters by using mathematical models. In most cases the calculation from mathematical parameters, which derive from the chosen model, to physical parameters, in this case the thermal transfer coefficient and solar aperture, introduces another point for discussion between physicists and mathematicians. Physicists like to compare the obtained values of the estimates from different methods, however they do not always realise that the way they are obtained from mathematics might be different. On the other hand, for the determination of the thermal and solar characteristics the knowledge of the heat flow through the test room envelope is an absolute must, in order to be able to obtain the properties of the test component decoupled from the device under test.

The here presented analysis and validation approaches will be illustrated step by step using simple and well documented case studies dealing with increased complexity. The following approaches will be considered: average and linear regression methods, transfer function models and continuous-time state space models. The software tool LORD is applied as well as CTSM-R and routines in the R-environment [7].

The considered case studies include a wide representation of the physical phenomena that are present in actual buildings. The aim is to focus on how to transfer the available information of the physical systems to different mathematical models, e.g. the importance of model simplification of building physics represented by measured signals. The different approaches will be presented “bottom up”, starting from the simplest and gradually increasing complexity. This complexity will be introduced by using data-series for analysis of physical processes that take place in a test box under real conditions. In particular the variability of the environments and the uncertainty of data will be discussed, e.g. how to deal with measured data and unmeasured phenomena and how to build a mathematical model based on the available input.

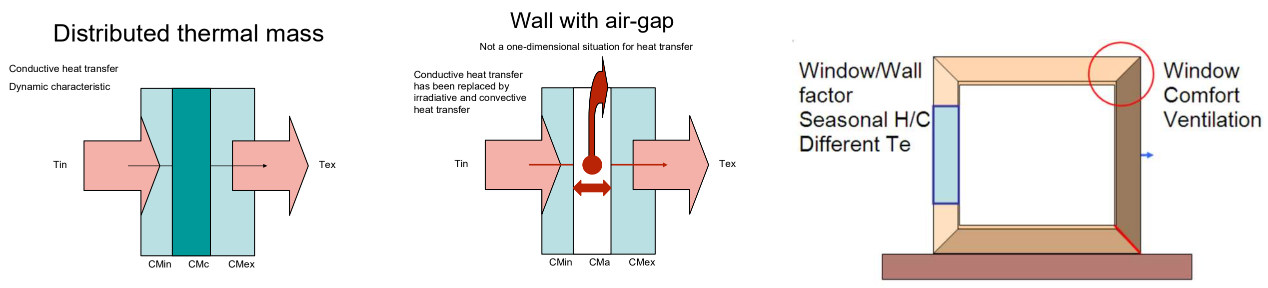


Figure 1a (conduction only), 1b (air-gap case) and 1c (enclosed volume).

The training exercises focuses on energy flows through walls (mainly by conduction; Figure 1a). Initial assessment proposed is based on physical processes that are presented by mathematical models. Thermal characteristic parameters are the thermal resistance (R), capacitance (C) and time constant (τ). See below.

Impact of solar radiation, convection and variable climate conditions will be presented. Analysis methods that will be used have an increased complexity and will focus on the energy balance while dealing with all physical processes, e.g. conduction, convection and radiation. However real applications are more complex systems such as facades with air gap (Figure 1b) as well as whole buildings that will be considered as closed volume/surface environments (Figure 1c).

A mathematical model represents reality, however, by definition, it is always a simplification of the true physical system. The analyst is responsible for defining the model and hence the simplification of it. Therefore benchmark tests should reveal the ability of the final method and model to analyse data from in-situ measurements correctly, within defined uncertainty limits.

Discrete Time Method

The software tool LORD builds on a lumped parameter (RC-network) model. The original method implemented in LORD (a user-friendly software tool with graphical user interface) is the Output Error Method (OEM). A new quality in the analysis of outdoor experiments can be achieved by using advanced statistical methods: e.g. LORD also includes the Prediction Error Method (PEM) to make it more powerful. This tool is easy to use and specially adapted to the requirements of testing under real climate. The selection and creation of models is one of the items which is simplified in a graphical way. LORD provide all the available statistical methods in an easily applicable way and includes instructions and data for self-training. The development of the software tool LORD has involved close cooperation of mathematicians and the intended users, building physicists. It is available on request at www.dynastee.info

Continuous Time Method

CTSM-R (2015) is an R package [7] providing a framework for identifying and estimating stochastic grey-box models. A grey-box model consists of a set of stochastic differential equations coupled with a set of discrete time observation equations, which describe the dynamics of a physical system and how it is observed. The grey-box models can include both system and measurement noise, and both nonlinear and nonstationary systems can be modelled using CTSM-R. It has been successfully applied to a wide range of data-driven modelling applications: heat dynamics of walls and buildings, dynamics of heat exchangers, radiators and thermostats, solar thermal collectors, building integrated photovoltaic systems and more. 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. By using a continuous time formulation of the dynamics and discrete time measurements the framework bridges the gap between physical and statistical modelling. Continuous Time Stochastic Modelling for R or CTSM-R is a free, open source and cross platform tool for identifying physical models using real time series data. How to use CTSM-R is described in detail in the CTSM-R user's guide and reference manual [1] (2018).

Both methods mentioned above are useful for most standard buildings and building components and are applied in this paper. The difference between discrete and continuous time method can easily be derived from comparing the mathematical expressions for the increment of the timestep.

$$\text{Discrete time; } \frac{X_{k+1} - X_k}{\Delta T} = AX_k \qquad \text{eq.1}$$

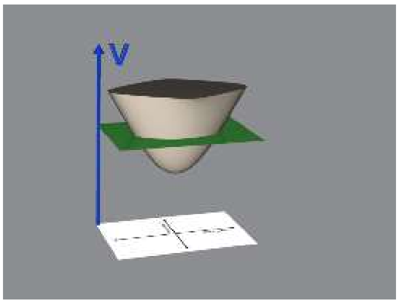
$$\text{Continuous time; } X_{k+1} = e^{A\Delta T} X_k \qquad \text{eq.2}$$

In general two types of criteria for parameter identification can be distinguished in: the Prediction Error Method (PEM) and the Output Error Method (OEM). The OEM is a special case of the PEM for the condition that H(q) = 1, when the following formula is taken into consideration:

$$Q(t) = G(q)u(t) + H(q)e(t) \qquad \text{eq.3}$$

The Prediction Error Method (PEM, applied in CTSM-R) is based on statistical models assessing parameters by minimising the error between a k-step (usually k=1) ahead prediction and the measured output. Some characteristics of PEM are that it is more sensitive to high frequency parameters and will be disturbed if residuals are auto correlated. Obviously, PEM requires more calculation time.
The Output Error Method (OEM) or Simulation Method (like in LORD) is based on deterministic models and assesses parameters by minimising the error between simulation and measurement over a whole test period.

$$V = \sqrt{\left(\frac{1}{N} * \sum (T_{meas} - T_{calc})^2\right)} \qquad \text{eq.4}$$



$$V = \sqrt{\frac{1}{N} \cdot \sum (T_{meas} - T_{calc})^2}$$

Figure 2. Minimisation by the Least Squares Method.

Some characteristics of the method are that it is more sensitive to low frequency parameters and gives too optimistic confidence intervals if residuals (here simulation errors) are auto correlated.

1.2 GENERAL CONCEPT

In brief, the general approach for analysis is to plot the available data (see figure 6) and decide which data could be used for analysis. Some basic knowledge of the physical processes of heat transfer is required to do so. Apply the average method to get a first impression of the expected result. In a further step apply a linear regression method and get an idea of the uncertainty. Note that NO dynamic information can be obtained from either average or linear regression method. For assessing dynamic characteristics the analysis requires dedicated methods for which two are presented below.
The measured data leads to ‘raw data’ with a certain accuracy. It should be stored and documented for later analysis or modelling work. It leads to the following overview of the interaction between experimental work and analysis of the obtained measured data. It can be applied to heat transfer through the building envelope, either a wall or the whole building.

General analysis approach.

The objective is the identification of a mathematical modelling with application to energy performance assessment in the built environment. The problem is stated as follow: how to go from many measured data to a few estimated values and try to get confidence in the whole process of identifying the searched parameters of the system. This can be achieved by:

- Plot the data and try to understand what it represents. Are the signals as you expect them to be, in particular temperatures. EXCEL is often used but any software package that you are acquainted with, will do.
- Apply the average method (several ways to do so). No dynamic but steady state conditions
- Apply a regression method; it is a bit more dynamic and linear.
- Any other method, such as ARX and grey box models; dynamic methods for which dedicated routines should be used

Averaging and filtering

How to get from many observations as input for the calculation process to one or a few limited output values for reporting? In that process the accuracy of input data, the propagation of the errors in the calculation process and the required accuracy of the reported value are of high importance. Once data has been produced (raw data), from a dedicated experiment, it is assumed that these data contain all information describing the physical processes that a mathematical model is supposed to analyze. Treatment of raw data is therefore crucial and should be performed by someone who has knowledge about the physical processes as well as the experimental set-up. In fact, by averaging and filtering in the space and time domain, thus by reducing the data, one has to understand that dynamic information will be lost.

Table 1: Process to obtain a few estimated values (characteristic parameters) from many measured data. Main elements of the methodology.

INPUT	METHODOLOGY	OUTPUT
Many observations from time and space ; raw data Physical processes Literature General knowledge	Description of physical processes into mathematical equations. Method should fulfil the aim taking into account the searched output	Limited value(s) Period; annual, daily, hourly Performance Efficiency; reference value Data for simulation
	Pre-processing, Model choice Iteration process, Post-processing Statistical tests, Model validation External tests	

Averaging is frequently used as filtering and also as resampling technique. Other filtering and resampling techniques can be applied. Discussing filtering and resampling techniques is out of the scope of this document, however some relevant issues regarding their application are discussed in the following.

- Averaging techniques to obtain an idea about the thermal resistance
- Apply different length for period. Split the whole period in 3 or more shorter periods (up to a day)
- Apply increasing length and investigate the result
- Standard deviation of the different average periods.

$$\hat{R}_{avg} = \frac{\sum_{k=1}^{600} \left[\theta_i(k) - \theta_e(k)\right]}{\sum_{k=1}^{600} q_i(k)} \qquad \sigma^2 = \frac{\sum_{k=1}^{600} (R_k - \hat{R}_{avg})^2}{599}$$

In general any averaging and filtering carried out must be justified, explaining its interpretation, which are the beneficial performances that are expected applying it, why improvements are expected, etc. The figures below show some results of averaging data.

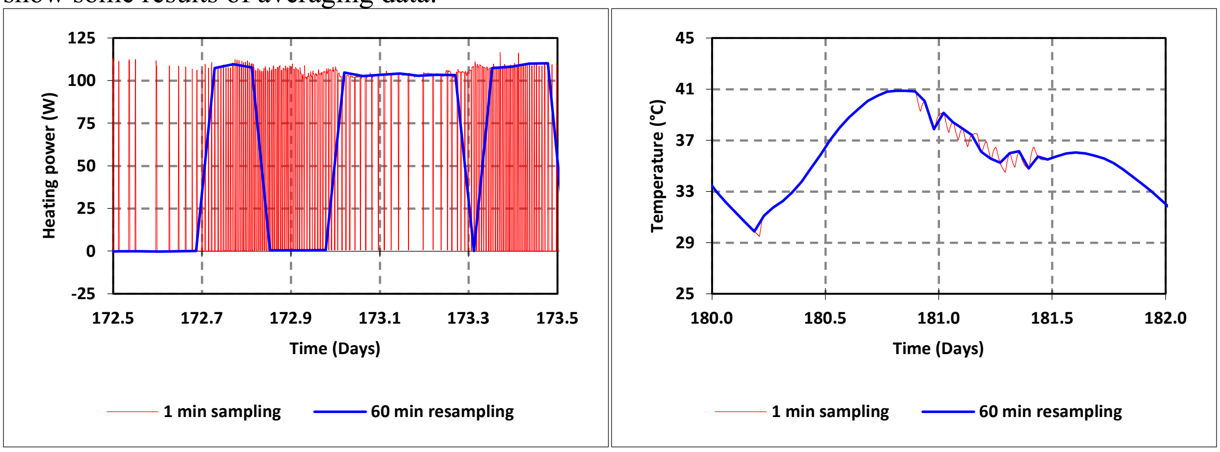


Figure 3a and 3b

Dynamics of physical processes are supposed to be represented in the data. Sometimes it should be carefully considered as contributing to the physical process that is being studied. In general solar radiation is giving problems to the correct interpretation of sensor observations. In the figure 4, an example is given of the effect of solar radiation on sensors. The heatflux given by the sensors not hit by solar radiation is lower than the heatflux given by sensors hit by solar radiation. Compare the heatflux of the ceiling (yellow) with that of the floor (blue) or right wall (red). One important question here is: does the sensor installation guarantee that the heatfluxes of the sensing elements are the same as the fluxes through the measured surfaces, that are much larger than the sensor? In this case averaging would contribute to an error in the analysis process. Good documentation of experiment set up including detailed description of sensor installation is an indispensable complement for high quality data sets, and would be very helpful to answer this kind of questions.

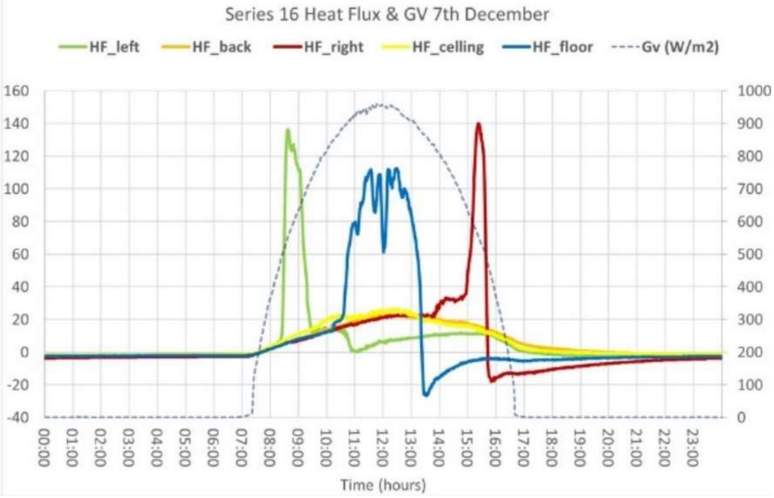


Figure 4. Heat flux and solar radiation signals for 7 December.

As an example the raw data from period 16 can be considered; see figure 5. There are 17279 records, each containing 45 observations that easily can be placed in Excel. To plot all sensors in one graph would give too much and too complicated information to draw useful conclusions. As a first attempt, it has been decided to take the Indoor Surface Temperatures of each wall, as well as the air temperatures in the center of the box. Clearly can be recognized that the whole period is composed of 12 days and a day-night sequence can be seen.

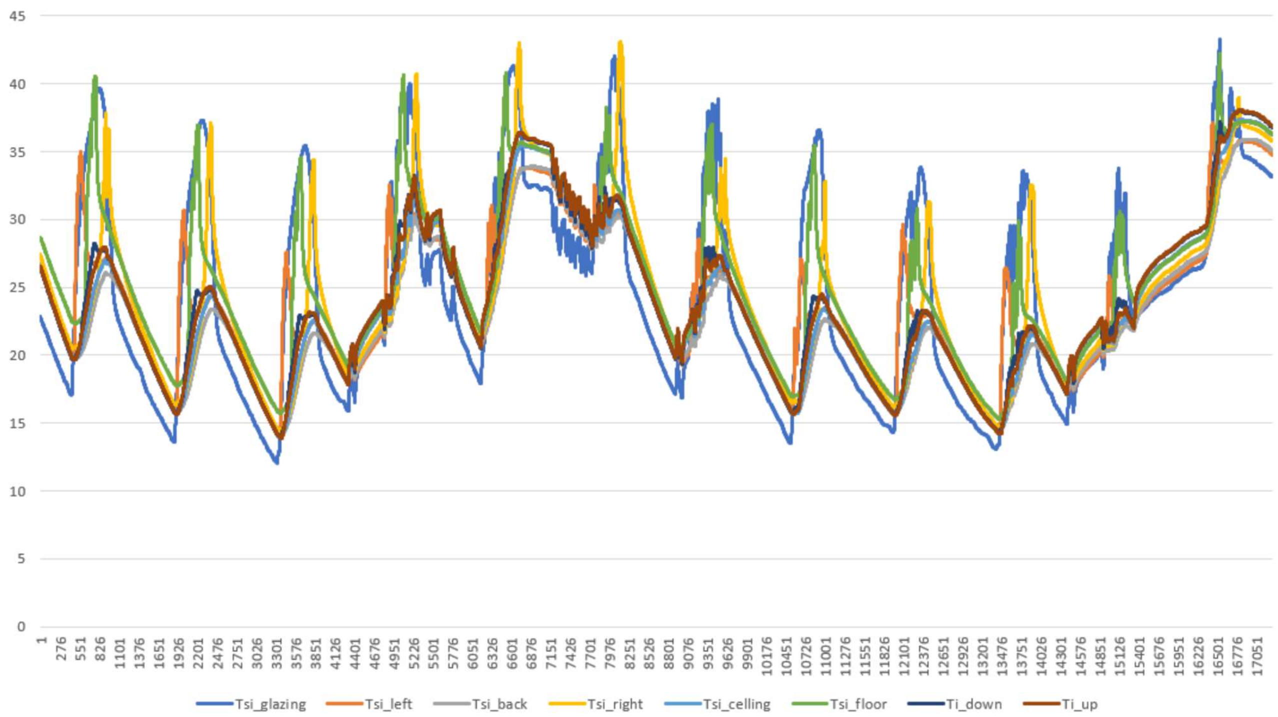


Figure 5. Indoor Surface Temperatures (6+2 sensors) for the whole data series 16.

The signals differ in amplitude and the data require a more closer look. So, one day is selected, 7 December, to study the details of the signals as can be found in figure 6. Almost all 8 signals behave different which is not a problem if the reason for the differences can be traced. First of all, there are 6 surface temperatures placed in the center of each wall. Note that one of the sensors (blue) is on the glazing of the wall oriented to the South and is hit by solar radiation. It indicates when more or less the night is finished and the sun comes up. All temperatures start to rise. The red signal corresponds with the sensor on the West wall and rises immediately when the solar radiation hits the sensor. During the morning the sun goes further up and at a certain moment hits the sensor on the floor with the maximum around noon. Note that the radiation is not hitting the sensors on the West and East wall. What also can be seen is that both air temperature sensors (dark blue and brown) tend to increase and reach their maximum, late in the afternoon. Meanwhile the solar radiation hits the East wall and hence the sensor. Once the sun disappears, the walls and window glazing cool down, a process that continues until the morning hours. Note that the window glazing is cooler than the walls and the floor is a bit warmer; all according to the physical processes of heat transfer.

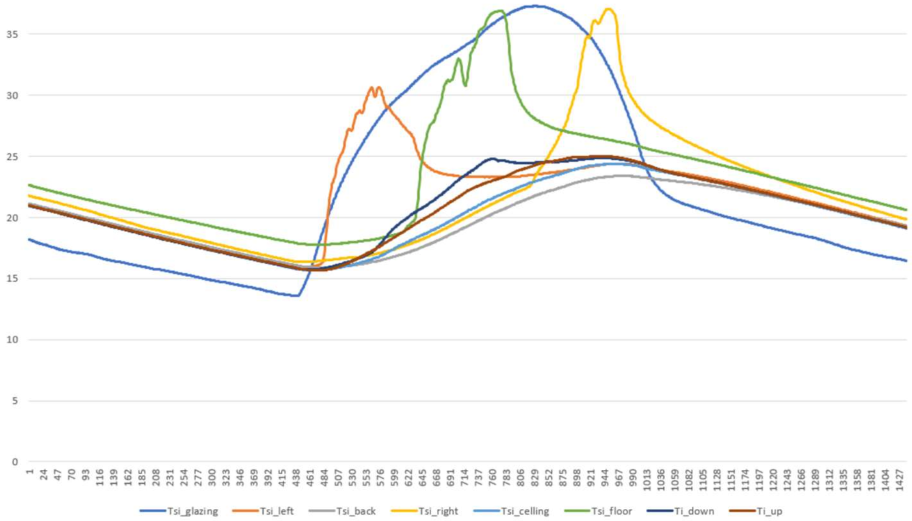


Figure 6. One day; 7 December shows details.

Now the question is if the data can be used for the analysis? Should the disturbed sensor readings be excluded or can it be used in an intelligent way? One of the steps forward could be to analyze the rear wall since it is not exposed to solar radiation whereas the composition of the wall is similar to the other walls, except for the window part. Since all walls are measured in detail, results can be used for more complex models.
The application of filtering techniques is useful when there is certainty that their effect is removing information in the data that doesn't correspond to the phenomena that we are studying in the building or building component. However filtering and averaging could have harmful effects if it removes relevant information to the process under study. A next example is the study of the heatloss coefficient of the glazing part of the window façade.

Table 2. Daily averaged values for the glazing part.

	'Tsi_glazing'	'Ti'	'Tse_glazing'	'Te'	DT	HF_glazing Series 16'	
06/12/2013	19.75	22.90	19.56	5.25	17.65	18.34	
07/12/2013	16.30	18.99	17.32	4.75	14.23	15.05	
08/12/2013	14.90	17.33	17.07	4.88	12.45	13.31	
09/12/2013	21.04	23.51	19.40	6.76	16.74	18.66	
10/12/2013	24.05	27.10	19.95	7.37	19.73	21.25	
11/12/2013	26.33	29.75	19.68	7.47	22.28	25.47	
12/12/2013	18.70	21.42	17.55	6.05	15.37	17.01	
13/12/2013	16.52	18.86	17.81	6.38	12.49	13.01	
14/12/2013	15.94	17.82	20.24	8.11	9.71	10.16	
15/12/2013	14.84	16.67	20.09	6.96	9.70	9.73	
16/12/2013	20.24	22.57	18.61	5.78	16.79	20.47	
17/12/2013	28.17	31.14	21.83	7.23	23.91	27.62	
Average					15.92	17.51	
					U-value =	*1.10*	*W/m^{2}K*

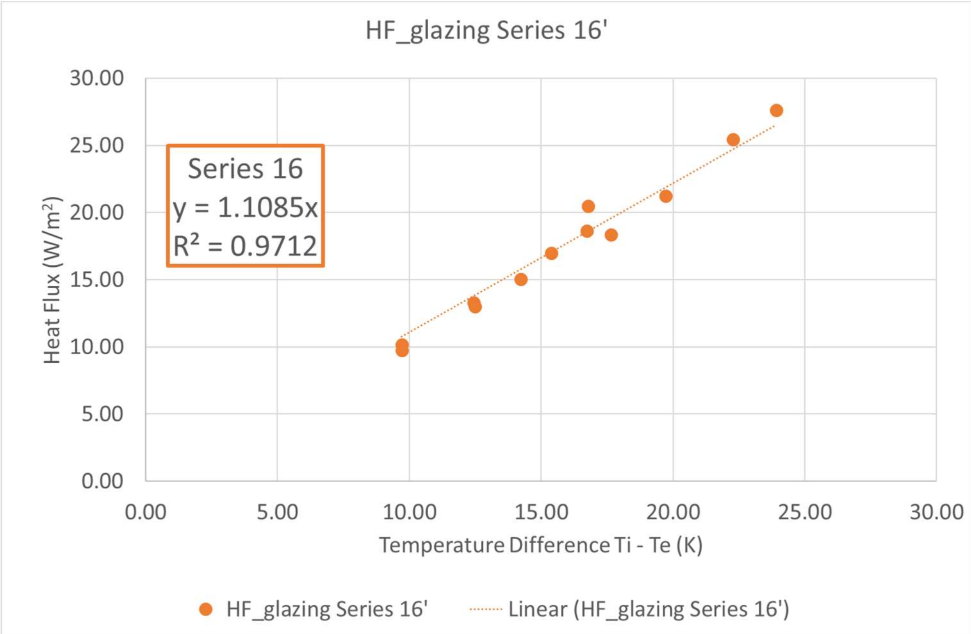


Figure 7. Plotted daily averaged values for the heatflux and temperature difference.

By the averaging method, a value for the glazing U-value of 1.10 (W/m^{2}K) is derived. Daily averaged values for the relation between the heat flux and the temperature difference gives 1.11 (W/m^{2}K) whereas the linear regression method gives a value for the U-value of 1.11 ±0.04 (W/m^{2}K).

Definitions of used expressions.

Definition given by International standards are used in this document - particularly the following included in ISO 7345:1987 (Thermal Insulation – Physical Quantities and Definitions):

Thermal resistance, R: Temperature difference divided by the density of heat flow rate in the steady state condition. Units: m^{2}K/W.

Thermal conductance, Λ: Reciprocal of thermal resistance from surface to surface under conditions of uniform density of heat flow rate. Units: W/(m^{2}K).

Thermal transmittance, U: Heat flow rate in the steady state divided by area and by the temperature difference between the surroundings on each side of a system. Units: W/(m^{2}K).

The following are defined by the ISO 13790:2008(E) (Energy performance of buildings - Calculation of energy use for space heating and cooling).

Heat transfer coefficient: Heat flow rate divided by the temperature difference between two environments; specifically used for heat transfer coefficient by transmission or ventilation. Units: W/K.

Transmission heat transfer coefficient: Heat flow rate due to thermal transmission through the fabric of a building, divided by the difference between the environment temperatures on either side of the construction. Units: W/K.

Ventilation heat transfer coefficient: Heat flow rate due to air entering an enclosed space, either by infiltration or ventilation, divided by the difference between the internal air temperature and the supply air temperature. Units: W/K.

Characterization by system identification techniques requires a lumped representation of building fabric. The following parameters are usually considered in such lumped representation of a given building envelope:

Often the following parameters have to be identified

Overall thermal transmittance coefficient UA: the heat flow rate in the steady state divided by the temperature difference between the surroundings on each side of the system or component, in W/K. For the 1-D case the U-value, in W/m^{2}K.

Total solar energy transmittance or solar aperture gA: the heat flow rate leaving the component at the inside surface, under steady state conditions, caused by solar radiation incident at the outside surface, divided by the intensity of incident solar radiation on the component, in m^2. For the 1-D case the g-value [-].

The main conclusion from this introductory section is that in general one has to understand what a measured signal represents. What information is available from the single sensor signal or from a group of signals (such as the average from 7 indoor air temperatures that are supposed to represent one indoor air temperature)? To get more knowledge about measurements and what information is contained in the observations a pragmatic approach for checking is proposed:

- Graphical plot of signals; it may indicate outliers, sudden changes as function of expected time constants.
- Statistical methods; average and variance of group of sensors; check of expected limits.
- Pre-processing of data for the purpose of mathematical modeling is therefore important and should be carried out with caution. Reduction of observations and signals on the input side implies the examination of the uncertainty of the input data to the calculation model. More about this aspect can be found below and illustrated in the figures.

Data analysis must consider at least the following steps:
Pre-processing: Any pre-processing carried out must be reported. Participants in the analysis exercises were encouraged to report data overview based on plots, discussion about quality of data and their suitability to fit objectives, etc.
Modelling approach: The methods and models used must be described. The hypotheses and approximation about the physics behind the considered candidate models must be justified. Schematic representations of heat flows in the building are recommended to support explanations. The process of model selection and the decisions made in this process must be explained. The software tools used to identify the parameters must be mentioned.
Validation: The validity of the results must be demonstrated. Statistical criteria are very useful in this process. Results must not contradict physical consistency. The process followed to demonstrate the validity of the results must be explained.
Results using different data must be compared. Since the data comes from the same physical system the best model should give similar results for two (or more) data series.
Results: A value estimated for each parameter and its corresponding uncertainty must be clearly marked as the final result. A list of the hypotheses and approximation of the physics behind the model finally selected to give the final results, must be given together with the final result.
Reporting: Reports must include at least a section devoted to error analysis
Conclusions: Any relevant finding resulting from the analysis, about the results, about the experiment set up and measurement campaign, etc., must be summarised.

Feedback among the different points should be made in every phase of the process. Is the model accepted? It is advisable to apply more than one method to get a better understanding of the whole problem. Common sense should always be used and all the available physical and statistical knowledge should be used whenever possible. A more detailed article "Guidelines for data analysis from dynamic experimental campaigns", is available from www.dynastee.info

2. MEASUREMENT DATA FROM THE PSA TEST BOX

2.1. Experiment set up

The high quality data series considered here were obtained from a test campaign carried out on a test box (simplified building) as it was constructed within the framework of IEA EBC Annex 58 and used for previous round robin activities in this context [8]. These experiments and data are considered a very useful benchmark for the development, application and validation of data analysis procedures for building dynamic modelling and energy performance assessment and also to support training initiatives. The main issues that make the test set up and data relevant in this context are:

- The incorporation of features which are present in real-life buildings, bringing physical effects of increasing complexity to the analysis, but in a somewhat simplified framework, as the test box is relatively simple.
- Tests with different conditions are useful to try and analyse the robustness of different analysis approaches, validate the results, and evaluate the impact of the different conditions on the accuracy of the results.
- The availability of the geometrical details and the thermal properties of its materials, providing theoretical target values that are useful to support validation of the experimental results obtained from data analysis.

Figure 8. Test box tested at PSA, Spain.

2.1.1 Description of the test box

The test box consists of a cubic structure that can be seen as a simplified building. Its exterior dimensions are $120x120x120cm^3$. All the walls floor and ceiling are identical and 12cm thick. One of the walls has a window with a wooden frame. The size of this window is $71x71cm^2$, and its glazed area is $52x52cm^2$. The box is suspended over a structure that avoids the contact with the ground. Figure 8 shows a view of the test box and more detailed descriptions are included in [10, chapter 3 and 5].

Design values of the characteristic parameters of the box are summarised in Table 1. This table also includes estimates of the deviations of these parameters from constant values due to wind speed and surface temperature. The values were obtained as:

- Target value: taking into account the properties of the materials provided by the manufacturers and the geometrical details of the construction.
- Target value assuming some imperfections as a thin air layer, 2.0 mm thick, at the interface between each two layers. Slight differences have been calculated in U and HLC values caused by this issue.
- Maximum deviation: as the standard deviation among all the average values for winter and summer data series due to variations of surface temperature and wind speed. Minor deviations have been estimated for all the relevant parameters (see Table 3 below).

Table 3: Target values and their deviations from constant values due to wind speed and surface temperature.

Parameter	Target value	Target value including imperfections	Maximum deviation
U opaque walls (W/m^2K)	0.476	0.413	0.003 (0.7%)
HLC whole box (W/K)	4.08	3.75	0.03 (0.7%)
gA whole box (m^2)	0.162	n/a	0.00002 (0.01%)
g opaque walls (-)	0.0038	0.0033	0.0004 (7%)

It is noted that the theoretical g-values estimated for the opaque walls are very small, therefore candidate models can be considered which neglect the solar gains through the opaque walls. Observation of the interior surface temperatures and heat fluxes are also relevant to discuss the suitability of such an assumption.

2.1.2. Boundary conditions

A test campaign was conducted starting on the 28th of May 2013 and ended on the 10th of January 2014. A selection of data recorded from the 06/12/2013 to the 07/01/2014 have been considered in this paper. The outdoor tests were carried out at CIEMAT's Plataforma Solar de Almeria in Tabernas (Almería), in the south east of Spain (37.1°N, 2.4°W). The local climate is characterised by dry and very hot summers and cold winters, large temperature swings between day and night, strong global solar radiation on the south vertical surfaces in winter, strong global solar radiation on the horizontal surfaces in summer, and clear sky. The data files contain detailed measurements regarding these boundary conditions.

2.1.3. Measurement devices

The following outdoor climate measurements were recorded (Figures 9 and 10): air temperature, vertical global solar radiation (parallel and next to the glazing), wind speed and direction, horizontal global solar radiation, beam solar radiation, diffuse solar radiation, vertical long-wave radiation, relative humidity, horizontal long-wave radiation from the sky, vertical global solar radiation facing north.

The following variables were measured on the test box: indoor air temperature, heat-flux density and temperature on the interior surfaces of each wall (including floor and ceiling), exterior surface temperature of each wall (including floor and ceiling) and heating power.

Device accuracy as well as its correct integration in the measured system has been carefully implemented in order to record reliable representation of each measured variable. Particularly, shielding of air temperature sensors (figure 9a, b and c), integration of surface temperature sensors and heat flux meters gluing them to the interior surfaces and protecting them with tapes of the same colour of the surface (figure 10a and b). A comprehensive description of the sensors and data-acquisition system is reported in [10].

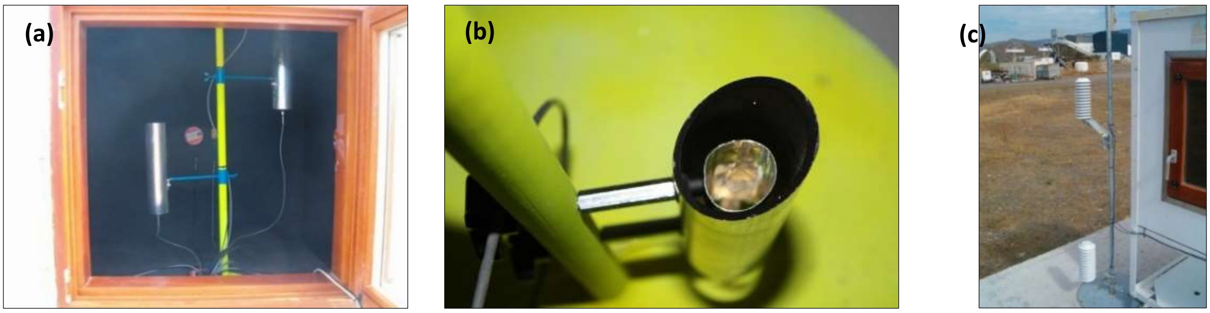


Figure 9. Test set up in Almería. Temperature measurement devices: (a) Indoor air temperature, (b) detail of indoors shielding devices, (c) outdoor air temperature.

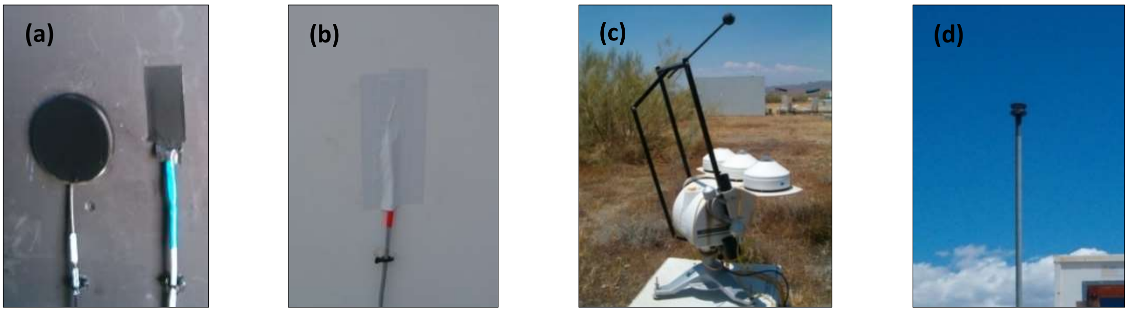


Figure 10. Test set up in Almería. Other measurement devices: (a) heat flux and internal surface temperature, (b) external surface temperature, (c) beam, diffuse and global solar radiation, (d) wind speed and direction.

2.2. Conducted experiments and data series

Several experiments have been carried out under different test conditions to obtain different data series which are suitable to analyse different aspects of the modelling methods. The heating power is provided by a shielded 100W incandescent lamp and the indoor air temperatures have been controlled in order to set the different test conditions in the different data series. The experiments have been designed with the aim of producing data series containing the information necessary to apply system identification techniques. These techniques require that the phenomena to be characterised are manifested and strong enough for the analysis process. Phenomena are strong enough in this context, when the amplitude of the corresponding driving variable is significantly higher than the uncertainty in its measurement. Otherwise signal-to-noise ratio is poor leading to inaccurate parameter estimates. Accordingly:

- To identify the heat loss coefficient, the experiment set up must ensure strong enough heat loss through the building envelope. This is achieved maximising the temperature difference between the indoor and outdoor air, which is the driving variable in this case. This difference is limited according to the safety recommendations of the construction materials and their usual operational conditions.
- To identify the overall gA-value, solar gains must be strong enough during the test. This is achieved when the experiment contains sunny days, with high solar radiation, which is the driving variable in this case.

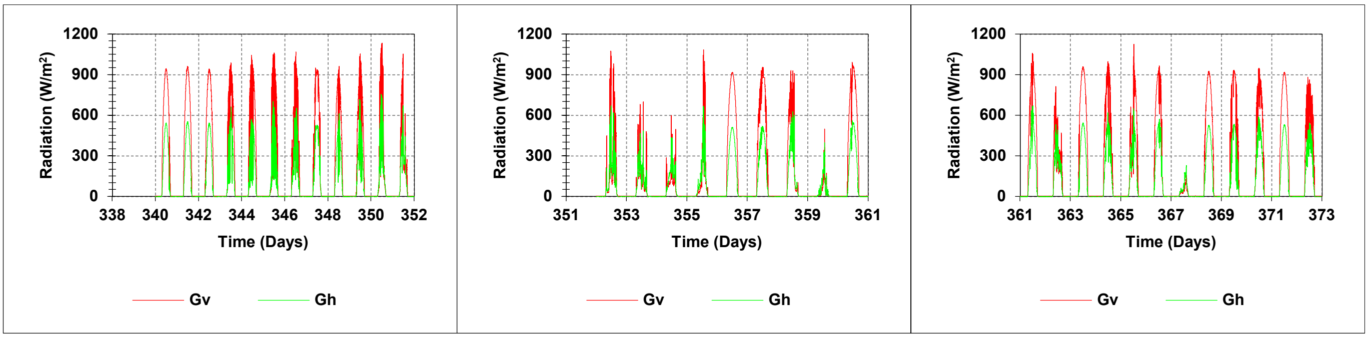
- To identify dynamic models, the system must be excited by dynamic input signals in a wide range of frequencies covering the characteristic time constants of the system. This is achieved e.g. by applying a PRBS or ROLBS power sequence. See ref. [11] for further information regarding the ROLBS sequence.

The following issues have also been considered relevant regarding the experiment set up:
- Heating Power: heating in the indoor air during the experiment is necessary to maximise temperature differences between indoor and outdoor air. Free running tests may lead to poor signal to noise ratios in the temperature difference measurements and problems with identifiability. Also, the heating power is a relevant variable to the energy balance equation used in the analysis, so it must be strong enough during the test.
- Homogeneity of the indoor air temperature: the different sources of heat such as heating devices and solar radiation can lead to some inhomogeneity of the indoor air temperature contributing to the uncertainty budget of the parameter estimates. A mini fan has been used to avoid stratification. A small device with very small ventilation power has been used to avoid perturbations in the interior convection coefficients.
- Sampling frequency: the sampling theorem must be fulfilled. The sampling frequency must be at least twice the frequency of the variable which is being measured.

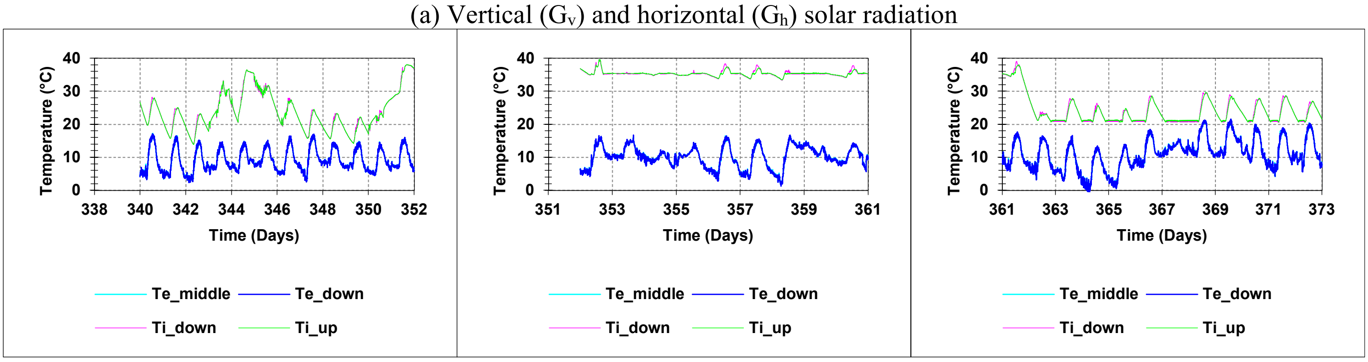
According to all these issues three PSA data series corresponding to different periods have been created. Each of the data series has 46 signals. A one-minute sampling interval has been set. This sampling interval guarantees that the raw data are fulfilling the sampling theorem. Note that any resampling applied in the pre-processing phase must also fit this theorem. The three periods have different characteristics concerning to the auxiliary heating and indoor temperature setting:
- <u>Dataset 16</u>: 06/12/2013 to 17/12/2013 (12 days). One ROLBS power sequence. This sequence was mainly designed to optimise the test conditions regarding the application of the system identifications techniques.
- <u>Dataset 17</u>: 18/12/2013 to 26/12/2013 (9 days). Aiming to replicate a co-heating test, but setting the indoor air temperature set point to 35°C with a dead band of 0.5°C. This test sequence was designed in order to have a reference analysis corresponding as much as possible to the traditional co-heating test, also to explore the application of steady-estate approaches and to analyse the capability of applying the system identification techniques to this type of test. Regarding these techniques, this series is also interesting to analyse causality issues and different variables as input signal. It must be observed that this test isn`t identical to the co-heating test as described in literature [12]. It is modified taking into account the given boundary conditions: in winter the sun position is low, so that strong global solar radiation is incident on the window (Figure 9a) and therefore indoor air temperatures cannot be maintained constant at 25°C. Even when the set point is raised to 35°C, the indoor air temperature shows some dynamic behaviour with overheating intervals (Figure 9b, centre). This problem isn't seen in summer, but then the outdoor temperature would very high, and consequently a key requisite to be able to identify the HLC from an outdoor test is to have an indoor air temperature significantly above 25°C. Otherwise, the heat flux through the building envelope is too weak, compromising the accuracy of the estimated value.
- <u>Dataset 18</u>: 27/12/2013 to 7/01/2014 (12 days): Also aiming to replicate a co-heating test, but in this case setting the indoor air temperature set point to 21°C, which could be in the comfort zone in an in-use building. The dead band in this test was set to 0.8°C. These more realistic conditions are set in order to explore the possibilities to apply these identification techniques to in-use buildings using data recorded by on-board monitoring systems. This series is also interesting to analyse causality issues and different variables as input signal. In this case, the difference between indoor and outdoor air temperature is small which increases the difficulty of identifying the HLC. Additionally, the contribution of the solar radiation to the space heating makes it impossible to maintain a constant indoor air temperature (Figure 11b, right) stressing the need to apply techniques for dynamic analysis.

2.3. Data overview
A preliminary observation of the graphical presented measurements gives key information to construct candidate models. Figure 11 includes some selected graphs that have been considered relevant after such preliminary observation of all the recorded data. Equivalent graphs of the three data series are included for being considered useful to illustrate the building response under the different test and boundary conditions. Some of the issues that could be addressed using these plots are discussed below. As a first analysis step the individual walls can be studied taking into consideration the impact of solar radiation. In fact six different conditions can be noted at the same moment of measurement whereas the composition of the walls (apart from the window) is the same. The impact of the solar radiation on the indoor energy balance, hitting the opaque

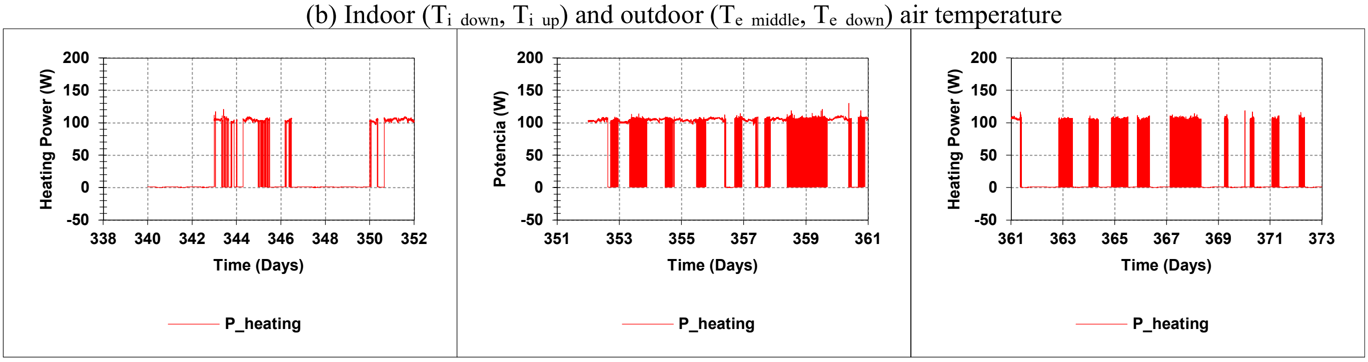
walls can be discussed from these plots. The levels of solar radiation are high and differ according to the orientation (Figure 11a), generating different temperatures on the exterior surfaces of the walls. However, the interior surface temperatures of the ceiling and back wall are similar (Figure 11d) which is an indication of the negligible effect on the heat flux through the opaque walls due to the solar radiation incident on the external surfaces. This makes sense considering the low g-value theoretically calculated as the value reference for the opaque walls. However, the interior surface temperature of the floor and right and left walls, presents some peaks (Figure 11d) attributed to the solar radiation transmitted through the window and incident on the interior surfaces.



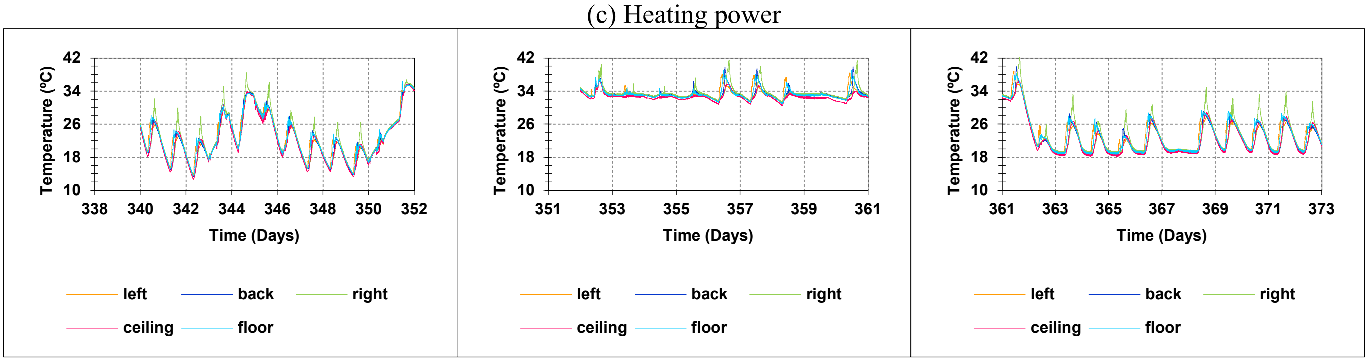
(a) Vertical (G_v) and horizontal (G_h) solar radiation



(b) Indoor ($T_{i\ down}$, $T_{i\ up}$) and outdoor ($T_{e\ middle}$, $T_{e\ down}$) air temperature



(c) Heating power



(d) Interior surface temperatures

Figure 11. Data overview: Dataset 16 from 6 to 17/12/2013 (left), Dataset 17 from 18 to 26/12/2013 (centre), and Dataset 18 from the 27/12/2013 to the 07/01/2014 (right).

3. APPLICATION OF LORD TO PSA DATA SERIES

LORD is user friendly software which enables a transient mathematical model of a component or building to be constructed. The parameters of the model (e.g. resistances, capacitances and heat flow admittances) essentially define the dynamic and steady-state thermal and solar properties of the system. Using LORD, the user describes the component or building as a series of nodes with a network of conductances and capacitances and measured input values of the indoor and outdoor temperature, the heating power, heat flux, the solar radiation, etc. The full dynamic information in the data is retained. Using the OEM, the output of the actual test (for instance, the room temperature as a function of time) is compared with the output which the model produces for the same input conditions. LORD also has a PEM facility.

Figure 12 shows a typical 4-node model which would be used to analyze heat flux measurements through a component such as the envelope of the PSA test box. Figure 13 shows a 5-node model used to estimate the Heat Loss Coefficient (HLC) and Solar Aperture (gA) of the PSA test box from the internal and external temperatures (Tint and Text), the heating power (Ph) and the solar radiation in the vertical plane (Gv). For estimation of the gA, two admittances (A2 and A4) are used to describe the solar radiation partly absorbed by the exterior of the test box and partly transmitted through the window, represented by conductance H2-4, and absorbed within the interior test room. A variant of this model has been used previously for co-heating tests [13].

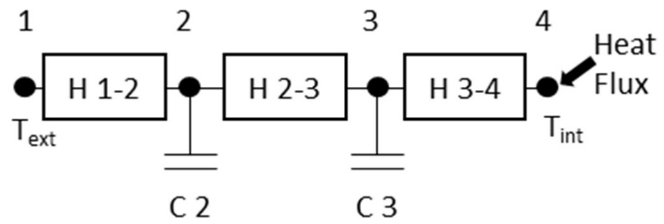


Figure 12. 4-node model of a component with measured external and internal temperatures and heat flux.

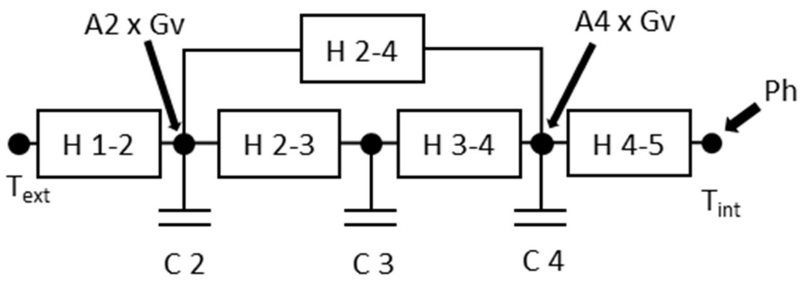


Figure 13. 5-node model of PSA test box.

Focusing on the whole test box HLC and gA, it is useful to use a steady state method, Siviour Analysis [14], to make a first estimate of HLC & gA. The steady state heat balance is described by Equation 5:

$$Ph = HLC \times \Delta T - gA \times Gv \qquad \text{eq. 5}$$

Dividing by ΔT gives:

$$\frac{Ph}{\Delta T} = HLC - gA \times \frac{Gv}{\Delta T} \qquad \text{eq. 6}$$

By calculating daily average values of the main parameters, the data can be represented graphically [ref 13, 14] in an X-Y plot with Gv/ΔT on the X-axis and Ph/ΔT on the Y-axis. HLC and gA can then be estimated by linear regression, where the intercept on the Y-axis is HLC and the slope is gA. The analysis was applied to each of the three data series (Figure 14). Series 17 gives the best fit: HLC = 4.06 ± 0.10 W/K and gA= 0.15 ± 0.02 m^2. Series 17 is the best 'steady state' data series with high ΔT and also variation in Gv data over the period.

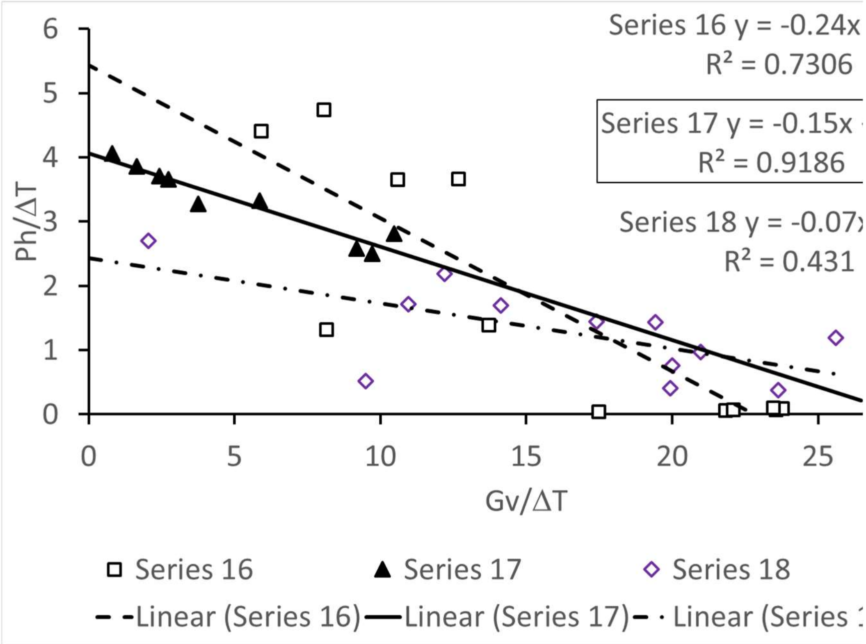


Figure 14. Siviour analysis applied to Datasets 16, 17 and 18.

The LORD analysis was carried out with 10-minute average data which retains most of the dynamic data, particularly in Series 16 with the ROLBS sequence, whilst reducing the size of the data set. The 5-node model (Figure 13) was used using the internal temperature, Ti, as the output variable. OEM was used for each of the data series 16-18 and all data; an additional analysis was performed on series 16 with PEM. The results, including the Confidence Interval (CI) are given in Table 4.

Table 4 indicates that similar results are obtained from the three individual series and all data for gA, however there is more variation in HLC. There is higher uncertainty for the series 17 results. Comparing OEM and PEM with series 16 data shows that PEM gives a better fit to all data, particularly series 17 (Figure 15).

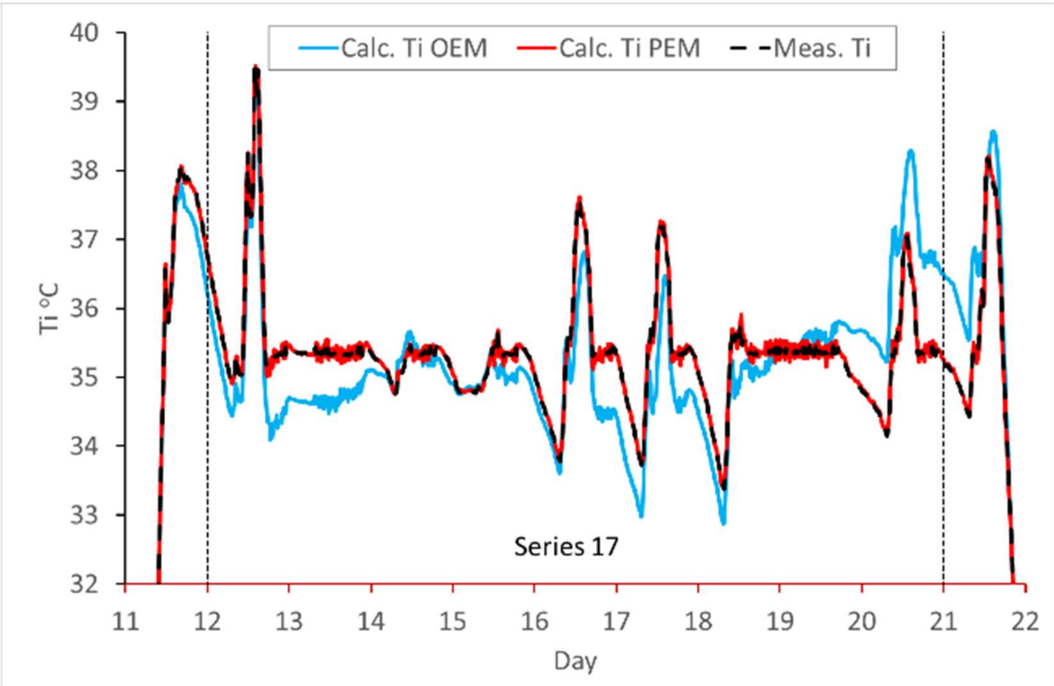


Figure 15. Comparing OEM and PEM: applying the results of Dataset 16 to Dataset 17 using the 5-node model.

Table 4. LORD results for PSA test box HLC and gA values using the 5-node model.

	HLC W/K	HLC: CI of the estimation method W/K	gA m^2	gA: CI of the estimation method m^2
Series 16	4.16	0.02	0.15	0.00
Series 16 PEM	4.09	0.03	0.15	0.00
Series 17	4.18	0.16	0.17	0.04
Series 18	4.30	0.03	0.17	0.00
All data	4.15	0.00	0.17	0.00

A simplified model with four nodes without a parallel conductance representing a window was also tried. PEM was also used to analyse the series 16 data. The results are shown in Table 5.

Table 5. LORD results for PSA test box HLC and gA values using a 4-node model.

	HLC W/K	HLC: CI of the estimation method W/K	gA m^2	gA: CI of the estimation method m^2
Series 16	4.15	0.00	0.15	0.00
Series 16 PEM	4.06	0.00	0.15	0.00

Table 5 shows that the simple 4-node model produces similar results to the 5-node model (Figure 13). Using PEM also produces a better fit to all data, particularly series 17 (Figure 16).

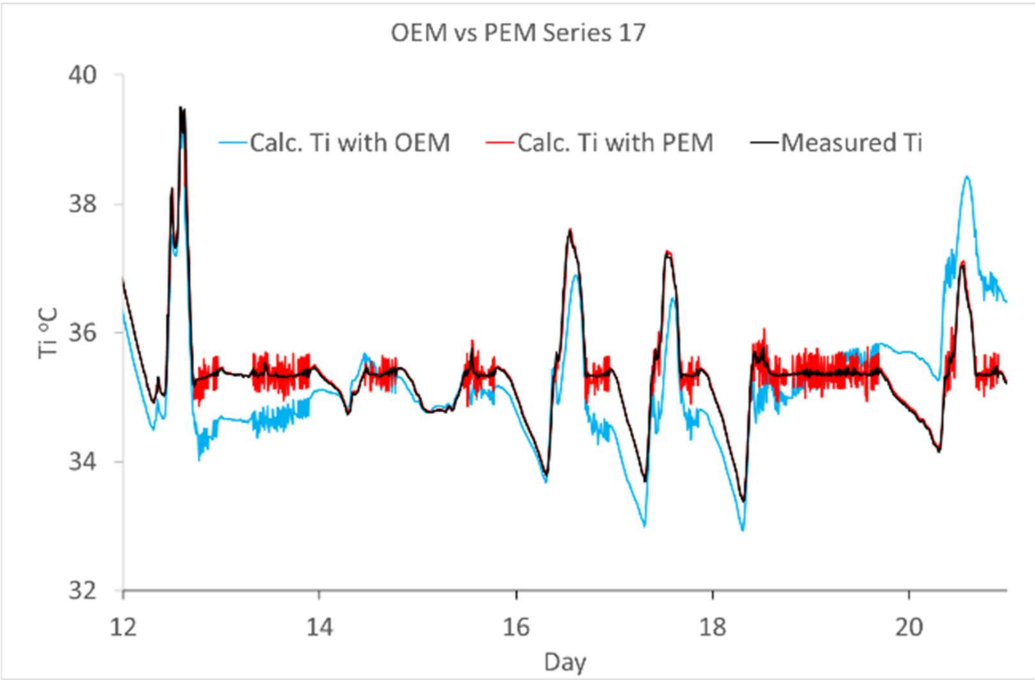


Figure 16. Comparing OEM and PEM: applying the results of Dataset 16 to Dataset 17 using the 5-node model.

3.1 Conclusions of the LORD analysis

Series 17 gave good results using Siviour analysis but was less satisfactory using LORD. Series 16 with ROLBS is better than the 'steady state' series 17 for identification of HLC and gA using a dynamic analysis. The more dynamic Series 16 produced satisfactory results, particularly using PEM which gave a good fit to series 17 & 18 data. This validates both the 5-node and simplified 4-node models for series 16 models. PEM gives more random residuals but takes significantly longer to run a calculation than OEM. LORD has the advantage that models can be quickly produced using the user-friendly interface and results can be quickly obtained using OEM. This enables students to gain confidence using identification techniques.

4. APPLICATION OF CTSM-R TO PSA DATA SERIES

This section describes the application of the grey-box modelling method CTSM-R [1] to obtain the Heat Loss Coefficient (HLC) results for the provided data. This method has been applied with the objective of comparing of the results with those previously obtained using LORD. Both methods should be able to provide similar results since they are applied for the same datasets. The similarity between the results will demonstrate the robustness of the methods.

In order to carry out the robustness analysis of the mentioned methods two different datasets will be considered: dataset 16 and dataset 17. However, to estimate the HLC of the test box, only selected signals such as the external temperature, internal temperature, heating power and the solar radiation have been used. The one minute data was converted into hourly averages due to some difficulties found when working with not relevant, minutely information.

Before testing the dataset with the proposed method, a simple visual input data analysis has been developed plotting the input data. Since in dataset 16, the heater was switched off during several days, it can be assumed that the grey-box model will have difficulties when trying to estimate the HLC due to its irregular behaviour compared with the rest of the input variables. Despite that, the analysis has been developed for the whole dataset 16. As expected from the visual check of the data, the analysis has to be limited to a shorter period in order to obtain better results, since some of the models were not able to converge and others were providing very unsuitable residuals. See also [16] for selecting suitable models. However, due to the limited period of time where the heater was working, it was not possible to limit the rest of the parameters. Unlike dataset 16, dataset 17 provides a wide range of records where the heater was switching ON/OFF resulting in more dynamic behaviour of the measurements. However, when analyzing the whole dataset 17 with CTSM-R, the obtained results were not accurate enough. Therefore, a shorter period is also selected in this dataset, where apart from the heater also the solar radiation effect is considered when selecting it as an input signal. For this selection, the period when less solar radiation was observed has been considered.

This method works for the case where each model includes the internal parameterization of the model. In other words, the models need to have some initial physical knowledge in order to make proper estimations. Thus, during the grey-box model analysis, the simplest one will be the first model to be studied and the models will increasingly become more complex. A set from the simplest to the most complex model will be fitted and the results will be described. All of the used models are plotted in Annex A for which the most complex is given in figure 16:

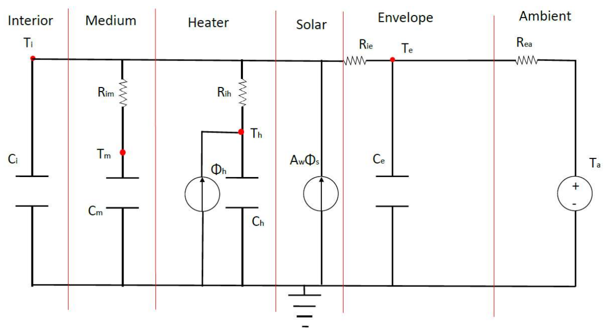


Figure 17. The most complex tested model is TiTmTeTh.

The shown model (Figure 17) contains seven parts that are combined in order to estimate the rest of the models, as shown in Annex A (models a, b, c, d, e, f, g, h) and the obtained results are shown in Table 6. The parts considered are the interior, the medium, the heater, the solar radiation, the envelope and the ambient. As seen, this model includes four state variables that represent the temperature in each part of the building; the interior temperature (Ti), the medium temperature (Tm), the heater temperature (Th) and the building exterior temperature (Te).

4.1 Validation and results for dataset 16:

The analysis has been performed between the 9th and the 13th of December 2013 (4 days), since it is the only period in the whole dataset 16 where the heater is switched ON. All the used input data is plotted in figure 18.

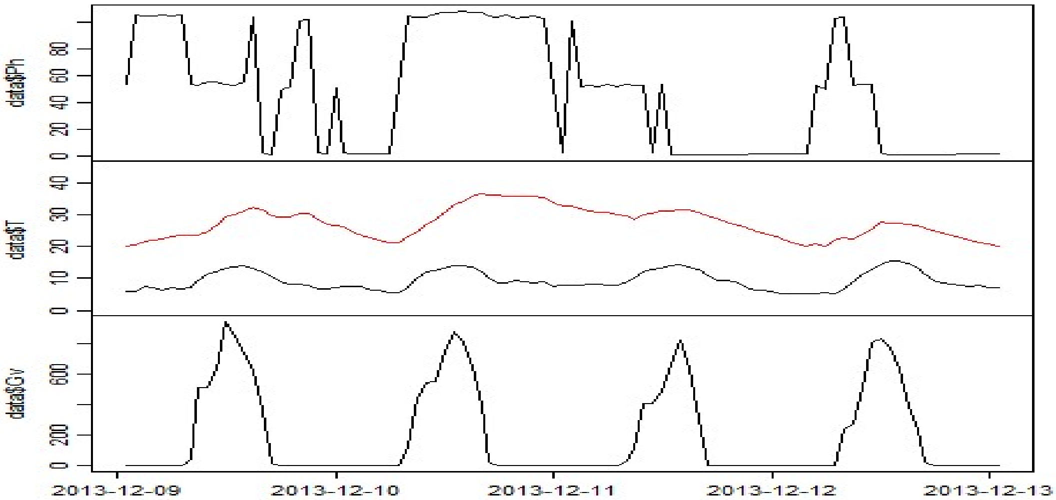


Figure 18. Input data for dataset 16

The selected period has been tested from the simplest model (Ti) until the most complex model (TiTmTeTh) where the internal (Ti), medium (Tm), external (Te) and heater (Th) temperatures have been used as state variables. The first step that needs to be followed when applying any modelling approach is common sense and physical criteria. In a first overview of the results model TiTeTh can be disregarded considering the large value of the error estimated for the HLC and the result obtained for the solar aperture that is significantly larger than the window area which is physically impossible. Large disagreement regarding the results obtained from other data series or other approaches which cannot be justified from the physical point of view, are also relevant criteria to disregard results from some models. Afterwards the validity of the grey-box model obtained using CTSM-R, is analysed from the statistical point of view through the Likelihood Ratio Test.

Applying this test, it will be possible to compare the different models and select the ones that best fit the data. In order to carry out the test, the data is analysed with each of the models, where each of them will provide a Log-likelihood value.

Table 6 . Results for the period in dataset 16

	HLC [W/K]	Error [W/K] (HTC 95% Confidence Interval of the estimation method)	Aperture (A_2) [m^2]	Error [m^2]	Log-likelihood	P value
Ti	3.80	±0.48	0.110	±0.040	-71.0	
TiTe	**4.10**	**±0.11**	**0.150**	**±0.010**	**-1.5**	**0 (Ti ->TiTe)**
TiTeRia	4.10	±0.11	0.150	±0.010	-1.5	0 (Ti ->TiTeRia)
TiTh	-	-	-	-	-	- (Ti ->Th)
TiTm	4.20	±0.15	0.160	±0.020	-9.2	0 (Ti ->Tm)
TiTmTh	-	-	-	-	-	- (TiTm ->TiTmTh)
TiTeTh	13.70	±27.90	3.400	±1.000	-48.9	1 (TiTe ->TiTeTh)
TiTmTe	-	-	-	-	-	- (TiTe ->TiTeTh)
TiTmTeTh	-	-	-	-	-	- (TiTeTh ->TiTmTeTh)

From this test, the p-value shown in the last column of Table 6 is obtained. These p-values show the relation between the different models. The lower this p-value is, the better the quality of the model will be for the corresponding data. From these p-values can be concluded that the most suitable model is TiTe. There are also two other models which show 0 as p-value. However, they are showing an extra parameter in the model (TiTeRia) or a lower Log-likelihood result (TiTm), with no extra improvement in the p-value.

Moreover, the residuals are satisfactory for the model TiTe and the obtained HLC result is equal to the value obtained with LORD, 4.1 W/K. The same happens with the solar aperture value, where 0.15m^2 is obtained. Moreover, the residuals for the best model (TiTe) are shown in Figure B.1 (Annex B).

4.2 Validation and results for dataset 17:

The same procedure is followed when analyzing dataset 17. In this case, the selected period is between the 19th and the 22nd of December 2013, three from the available 9 days, since none of the models was able to provide good results for the whole dataset. This proves that a decision making phase based on statistical information and common sense is required. Although the physical system has not changed, the data may contain too much or too less information for the chosen model.

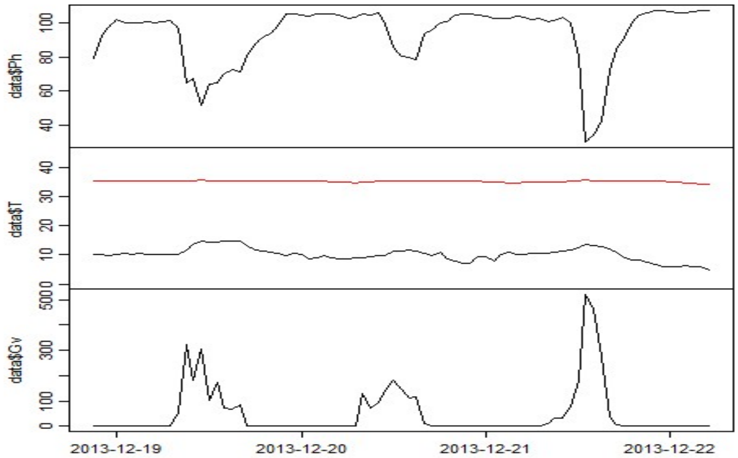


Figure 19. Input data for dataset 17

All the models are tested again for the reduced new dataset. From this test the Log-likelihood values are obtained, which will be indispensable for the Likelihood Ratio Test development. Unlike in dataset 16, there are no doubts selecting the most suitable period for the data. In this case, the lower p-value is also obtained with the model TiTe, resulting in the obtained value for the HLC, 4.2W/K, very similar to the obtained in dataset 16, 4.1W/K. However, if the aperture results are checked the results are exactly the same for both datasets, 0.15m^2. The results obtained are shown in Table 7:

Table 7. Results for the period in dataset 17

	HLC$_{corrected}$ [W/K]	**Error [W/K] (HTC 95% Confidence Interval of the estimation method)**	**Aperture (A_2) [m^2]**	**Error [m^2]**	**Log-likelihood**	**P-value**
Ti	4.2	±0.1	0.13	±0.03	82.7	
TiTe	**4.2**	**±0.09**	**0.15**	**±0.03**	**101.9**	**9.7* 10^{-8} (Ti ->TiTe)**
TiTeRia	4.2	±0.07	0.14	±0.03	93.1	2.1* 10^{-5} (Ti ->TiTeRia)
TiTh	-	-	-	-	-	- (Ti ->Th)
TiTm	4.3	±0.12	0.16	±0.04	93.4	2.8* 10^{-4} (Ti ->Tm)
TiTmTh	11.9	±16.9	3.9	±8.4	-76.3	1 (TiTm ->TiTmTh)
TiTeTh	-0.0008	±0.03	0.002	±0.08	76.5	1 (TiTe ->TiTeTh)
TiTeThRia	-	-	-	-	-	- (TiTe ->TiTeThRia)
TiTmTe	0.035	±0.19	4	±44	-100.8	1 (TiTe ->TiTeTh)
TiTmTeTh	-	-	-	-	-	- (TiTeTh ->TiTmTeTh)

5. FINAL CONCLUSIONS

During a number of research projects over the last 20 years, the need for advanced training for thermal performance assessment of the building envelope, based on two disciplines, e.g. building physical as well as statistical knowledge, has been identified. The importance of bench mark data for training to study relevant phenomena in real size buildings, is useful to develop skills in decision making regarding selection of variables and test periods suitable for analysis and modelling. Therefore, two different methods have been presented and applied to several data series from outdoor experiments. For each method different mathematical models have been used and results have been compared. The grey-box models applied in CTSM-R can be considered as robust. This statement has been demonstrated in the previously explained development where due to the analysis of the test box data, good results have been obtained. The same data has been applied using several methods in order to obtain the HLC and the solar aperture, gA of the test box. It has been demonstrated that LORD has been able to provide similar results. Moreover, the same happens between the two values obtained in the two different datasets recorded under completely different test conditions. For the HLC, the obtained value in LORD was 4.1W/K for dataset 16 and 4.2W/K for dataset 17, while the two results obtained using CTSM-R were 4.1W/K and 4.2W/K, leading to the conclusion that the obtained results are the same for both methods.

REFERENCES

1. Visit www.ctsm.info for all information and reference documents.
2. Gutschker O., Logical R Determination (LORD) – Software tool manual, available from www.dynastee.info (2003)
3. MATLAB – High performance numeric computation and visualization software. Reference guide. The Math Works; Natic; Mass;. http://www.mathworks.com/products/sysid (1992)
4. Söderström T, Stoica P. System Identification. Prentice Hall International: Hemel Hempstead, UK; (1989)
5. Bloem JJ (Ed.), System identification applied to building performance data. Commission of the European Communities. EUR 15885. (1994)
6. Bloem JJ. (Ed.). System Identification Competition. EUR 16359 EN (1996)
7. R is a free software environment for statistical computing and graphics http://www.r-project.org and RStudio http://www.rstudio.com
8. Guidelines for data analysis from dynamic experimental campaigns_Ch11.pdf; available from www.dynastee.info
9. Kristensen NR, Madsen H, Jørgensen SB. Parameter estimation in stochastic grey-box models. *Automatica*, **40**, 225-237; available from www.imm.dtu.dk/ctsm (2004)
10. M.J. Jimenez, et al.. "Report of Subtask 3 – Part 1. Thermal performance characterization based on full scale testing - description of the common exercises and physical guidelines". IEA EBC Annex 58; Final Reports. ISBN: 9789460189876, Published by KU Leuven, Belgium(2016).
11. K.R. Godfrey. Correlation Methods. *Automatica*, Vol **16**, 527-534. (1980)
12. D. Johnston, D. Miles-Shenton, D. Farmer, J. Wingfield. Whole House Heat Loss Test Method (Co-heating) http://www.leedsbeckett.ac.uk/as/cebe/ projects/cebe_coheating_test_method_june2013.pdf
13. Rhee-Duverne, S. and Baker, P. A retrofit of a Victorian terrace house: a whole-house thermal performance assessment, Historic England Research Report 103/2015. (2015)
14. Siviour, J. Experiment Thermal Calibration of Houses, Comparative Instrumentation of Low Energy Houses, Colloquium, University of Liege, Belgium. (1981)
15. Everett, R. Horton, A. and Doggart, J. Linford Low Energy Houses. Energy Research Group, Open University, Milton Keynes, UK. (1985)
16. P. Bacher, H. Madsen. Identifying suitable models for the heat dynamics of buildings, *Energy and Buildings*. **43** 1511-1522; (2011)

For further information and useful documents visit the web-site www.dynastee.info

Two summary articles can be downloaded from the *Overview* of the *Data analysis* menu-item. One is introducing physical aspects Analysis_physical_aspects.pdf while the second one introduces statistical aspects statistical_modelling.pdf of data analysis. The document Software techniques applied to thermal performance characteristics gives some further information about methods and tools and mentions as well benchmark data for testing these methods.

ANNEX A

All models with increased complexity (a to i) applied within the CTSM-R method

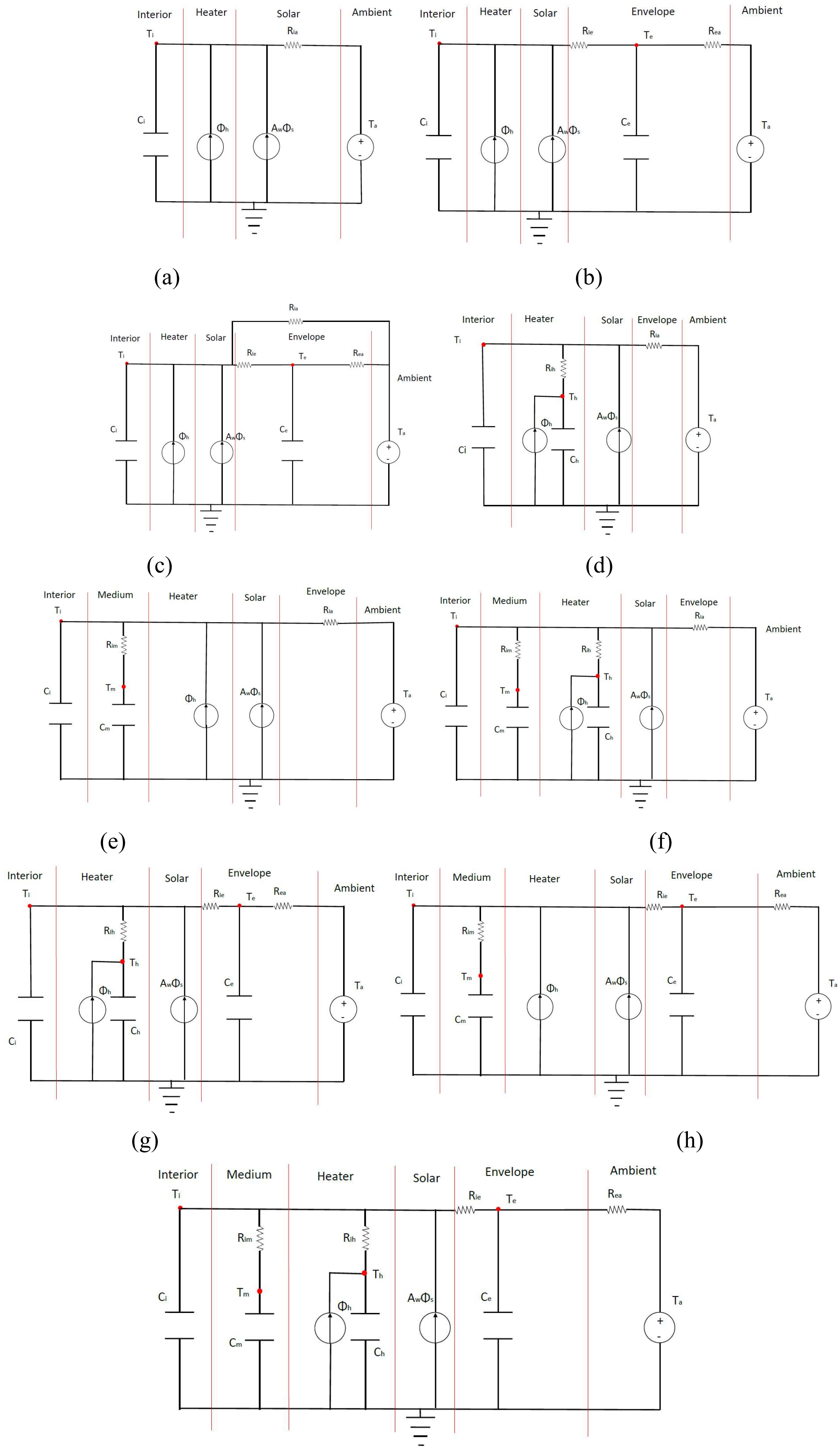
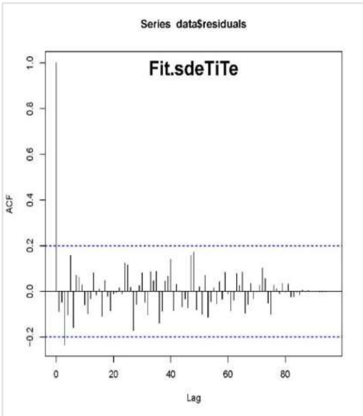


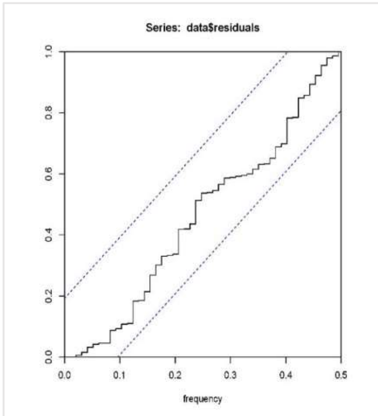
Figure A.1- All the tested models are (a) Ti, (b) TiTe, (c) TiTeRia, (d) TiTh, (e) TiTm, (f) TiTmTh, (g) TiTeTh, (h) TiTmTe and (i) TiTmTeTh.

The last model (i) contains seven parts that are combined in order to estimate the rest of the models, as shown in the rest of the models (a, b, c, d, e, f, g, h). The parts are the interior, the medium, the heater, the solar radiation, the envelope and the ambient. As seen, this model includes four state variables that represent the temperature in each part of the building; the interior temperature (Ti), the medium temperature (Tm), the heater temperature (Th) and the building envelope temperature (Te).

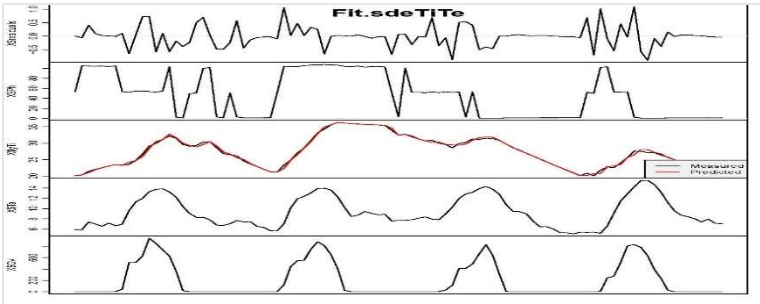
ANNEX B

The residuals of the best models are plotted in this section. Both of the datasets 16 and 17 were fitting suitably with model TiTe.



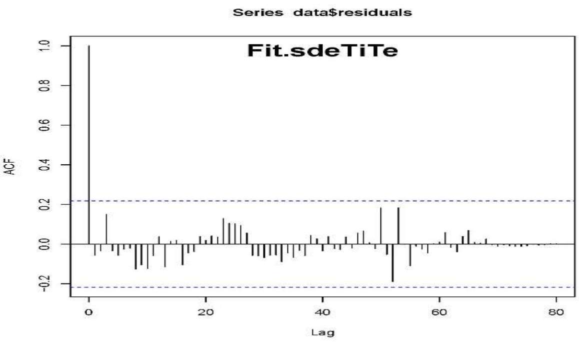


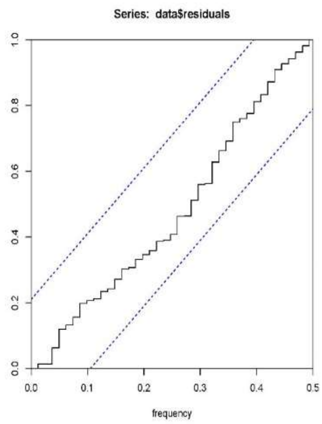
(a) (b)



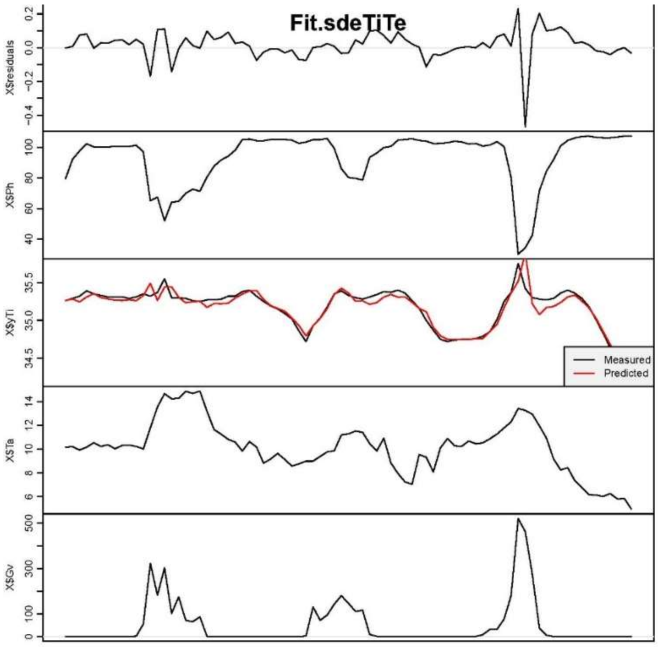
(c)

Figure B.1- Residuals of dataset 16: (a) Autocorrelation function, (b) periodogram and general residuals.





(a) (b)



(c)

Figure B.2- Residuals of dataset 17: (a) Autocorrelation function, (b) periodogram and general residuals.

The residuals obtained for both periods are quite good as shown in Figure B.1 and B.2. In case of dataset 16, the autocorrelation function (see Figure A.1-(a)) is not completely white noise but it is close to it. However, the autocorrelation function of the dataset 17 (see Figure A.2-(a)) can be assumed to be white noise. Moreover, the periodogram is describing suitably the dynamics of the models in both periods (see Figure B.1-(b) and Figure B.2-(b)).