

Thermal building model identification using time-scaled identification methods

Paul Malisani, François Chaplais, Nicolas Petit and Dominique Feldmann

Abstract—The aim of this paper is to propose a robust and accurate method for the parametric identification of the thermal behaviour of low consumption buildings. These buildings are known to have a two-time scale structure, which, if not handled properly, results in poor conditioning of the parametric identification.

We compare three identification methods, one uses the data on the whole frequency domain (ARX) when the other methods use the same data but separated on local frequency domain (time scaled methods).

All three methods identify a reduced second order model. Robustness is tested by corrupting the input and output before the identification, and comparing the simulation results for the various models and the original uncorrupted input. The numerical results clearly show that the time scaled methods are superior both in accuracy (noise free identification and simulation) and robustness (when identification is performed on corrupted data).

I. INTRODUCTION

As current norms on energy consumption become more restrictive, efficient control of the heating of low consumption buildings has emerged as a topic of interest. A way to address this problem of practical interest is to develop optimal control laws generating optimal trajectories under constraints. These constraints bear on the control and the state; they account for the in-door comfort and the power limitations (see [1] [2]). To compute such an efficient optimal control law, especially for long time periods, we need a low order model for the thermal behaviour of the building.

According to ([3], [4], [5]), low order linear models form a good set of models to describe the general thermal behaviour of buildings. But, as it has been stressed in [6], these models can give quite good results on prediction errors while providing poor estimates of the building's physical characteristics. This is a serious problem in the presented context of optimal control (especially under constraints) which requires good estimates of poles, zeros and static gains.

Usually, such bad performances can be the result of a bad conditioning of the identification's optimization problem. For the three identification methods presented here, these optimizations are formulated as quadratic problems, and the

condition is the conditioning of the excitation matrix (or matrices).¹It is related to the sensitivity of the solution of $Ax = b$ with respect to variations of A or b .

Ill conditioning of the excitation matrix(es) can be the result of insufficient frequency content in the input data; it can be also related to near collinearity of the state and future input subspaces [8]. However, it has been proved in [9] that, even for inputs which are rich enough in the frequency domain, the excitation matrix of two time scaled systems (such as low consumption buildings) is asymptotically degenerate as the ratio between the large and small time constants of the system tends to the infinity. Identifying these systems locally in the frequency domain removes these degeneracy problem.

It should be noted that, in the last two or three decades, time scales have been largely associated to wavelet transforms. Wavelets can be used in several ways in dynamical systems identification. The first usage is for data filtering. Indeed, we could use wavelet transforms to separate frequency bands in the data. However, if one sticks to the popular dyadic transforms, one is limited to time scales which are equal to powers of 2. More classical low-pass and high-pass filters are more flexible, and quite sufficient for our purpose. The other usage is to model the system directly in the wavelet domain. Characterization of finite dimensional systems in this domain have been studied in ([10], [11]). Reference [12] covers a similar topic. A limitation is that these processes are hardly (or even not at all) related to classical (rational) Linear Time Invariant (LTI) systems. The most visible reason for this is that the transforms from the time domain to the wavelet domain and back are not causal; therefore it seems unlikely that operations in the wavelet domain can be turned into causal operations in the time domain.

The purpose of this paper is to compare the performance of a classical ARX identification procedure to two variants of the two time scaled identification (see [9]), for the purpose of modelling a low consumption building with a second order model. The difference with [9] is that we are never in the model matching case. The performance is considered both in terms of simulation error with respect to a high order model, and in robustness with respect to data corruption.

This paper is organized as follows.

In Section II, we describe the plant we wish to identify, and define various data sets that will be used for that purpose. For comparison purposes, we introduce here data sets where each input generates a separate output; actually, we currently have

P. Malisani (corresponding author) is with EDF R&D, Département EnerBat, Centre des renardières, 77818 Moret-sur-Loing and is Ph.D. candidate in Mathematics and Control, C.A.S., Unité Mathématiques et Systèmes, MINES-ParisTech, 60 bdv Saint-Michel 75272 PARIS cedex 06, France paul.malisani@mines-paristech.fr

F. Chaplais and N. Petit are with the Centre Automatique et Systèmes, Unité Mathématiques et Systèmes, MINES-ParisTech, 60 bdv Saint-Michel 75272 PARIS cedex 06, France firstname.name@mines-paristech.fr

D. Feldmann is with EDF R&D, Département EnerBat, Centre des renardières, 77818 Moret-sur-Loing France.

¹We recall that, for the L^2 norm, the condition number [7] of a matrix A is the ratio between the largest and the smallest eigenvalues of $A^T A$.

access to the sum of these outputs, that is, the temperature inside the building. It is interesting to consider this possibility because it gives more information on the system, and we wish to evaluate the benefits of having access to that extra information.

In Section III, we describe the various model classes within which we will look for a model, and how we parameterize them with a finite set of numbers. This where we introduce two time scaled models. We detail how the parameters of a model class are related to the parameters of another one.

In section IV, we define the various optimization problems which, with the parameterizations of III and the data sets of section II, will define how the various parameters used in the model classes are obtained from the data sets. The definition of these optimization problem are important because the plant does not match any model of any class of section III. Indeed, the output data is generated by a LTI system of order 47 (possibly corrupted with noise), whereas we are looking for a model of order 2. Therefore the choice of the optimization problems greatly influences the determination of the system parameters.

In Section V we compare the results obtained in terms of static gains identification, statistical properties of simulation errors, conditioning of the optimization problems and poles and zeros locations. This is done using various data sets, models, and model parameterizations. These results are interpreted in the light of simulation accuracy and robustness with respect to data corruption.

In Section VI, we conclude on the results and show the substantial efficiency of the time scaled method in terms of simulation errors and robustness of the parameters identification to noises.

II. PLANT AND DATA

Our desired goal is to obtain a low-order thermal model of a one-area building describing the general behaviour of the internal temperature depending on several inputs. We need those models to optimally control the heating of a building under constraints. At the time this article is written, an actual *low consumption* building is not (yet) available to us for measurements. In its place, we shall use a high order (47^{th}) linear system as the “true” input-output mapping. This high-order model is a spatial discretization of the heat equation in the building.

The inputs and output and listed in table I. The control of this system is a part of the last input, together with human activities. We consider a person to be a constant input of $100W$ and we also know the heat provided by the devices inside the house. For identification purposes, we use inputs which are an average of chronicles over several decades. These data are experimentally measured weather histories sampled with a period of one hour over one year; due to their poor time-resolution it is likely that these signals are not well shaped to perform a good identification (see [8]).

The knowledge of the building’s geometric shape and its orientation, allows us to generate the input of the system.

TABLE I
INPUT-OUTPUT

Output	input
Internal temperature	External temperature
	Solar flux on the floor
	Solar flux on the walls
	Heating flux on the air node

These preliminary transformations are non-linear, and because a linear model is sought after, one cannot directly use the measured data but the transformed data to perform the identification. These non-linear transformations are described in [13]. The output is then computed by simulation using a LTI model of order 47 which accounts for the three-dimensions geometry of the building.

This data set the *noise free data*. By contrast, we will call *noisy data* the same data set to which we add noise *independently* on each input and output. The noises on each signal are Gaussian white noises of standard deviation equal to one thirtieth of the standard deviation of the signal. Because the signals are not stationary it represents a quite strong noise on the signals. For instance, this represents a standard deviation of $.3^{\circ}C$ on a temperature measurement, which is a realistic value for a temperature sensor. This signal/noise ratio is consistent with real application.

In addition, we shall use another data set, which we call *separated output* data set. It is obtained by separating (in simulation) the influence of each input within the internal temperature. This gives much more information on the plant behavior. We shall also allow ourselves to corrupt the data with independent noises; in this case, the data set will be called *noisy separated output*.

III. MODEL CLASSES AND PARAMETERIZATIONS

It is well known (see [14]) that the system detailed in section II can be efficiently represented by a second order linear model. This can be done in several manners, which we now discuss.

A. Classical ARX model

This is the classical LTI model with rational transfer function. The order here is two. We have restricted our study to strictly proper transfer. This model class, together with the chosen parameterization (see equation (1)), has been found to represent the best trade-off between robustness and simulation accuracy in numerical results.

The parameterization is given by (see [15])

$$y[k] + a_1 y[k-1] + a_2 y[k-2] = \sum_{i=1}^4 b_{i1} u_i[k-1] + b_{i2} u_i[k-2] \quad (1)$$

with $i = 1, \dots, 4$, and where the models parameters are $a_1, a_2, b_{11}, \dots, b_{14}, b_{21}, \dots, b_{24}$.

B. Two Time scale transfer

The difference with the previous model class (see equation (1)) is the introduction of a parameter $\epsilon \ll 1$ which represents the ratio between the "slow" and the "fast" time scales. Specifically, the transfer is expressed as

$$T_\epsilon(s) = T_f(s)T_s\left(\frac{s}{\epsilon}\right) \quad (2)$$

T_s and T_f are slow and fast transfer functions independent of ϵ . Thermal models are known to be two time scale and then can be represented by the equation (2) (see [16]).

For a given ϵ , the model class is the same as the ARX; however, it suggests a different parameterization and an adequate handling of each time scale. To do so, the following definition is needed:

Definition 1: We define the fast transfer $\tau_f(s)$ and the slow transfer $\tau_s(s)$ as follows

$$\lim_{\epsilon \rightarrow 0} T(i\omega) = T_s(i\infty)T_f(i\omega) \stackrel{def}{=} \tau_f(i\omega) \quad (3)$$

$$\lim_{\epsilon \rightarrow 0} T(i\epsilon\omega) = T_s(i\omega)T_f(i0) \stackrel{def}{=} \tau_s(i\omega) \quad (4)$$

Observe that, as ϵ goes to zero, the slow and the fast transfer keep a similar magnitude if and only if the slow transfer is biproper as defined in [17]. As suggested by Definition 1, T_ϵ behaves like τ_s in the low frequencies and like τ_f in the high frequencies. For a given ϵ , we can recover T_ϵ from τ_s and τ_f if the static gain of the fast transfer is equal to the high frequency gain of the slow transfer.

If some knowledge of a frequency that separates the two parts of T_ϵ in the frequency domain is available, we can design a low-pass pre-filter F_l and a high-pass pre-filter F_h from which the following model class and parameterization are defined:

Definition 2: The two time scale model class for the filters F_l and F_h are described in transfer form by

$$F_l y = \tau_s F_l u \quad (5)$$

$$F_h y = \tau_f F_h u \quad (6)$$

For a given T_ϵ the orders of τ_s and τ_f are given by definition 1. These two transfers are parameterized linearly as in the ARX class and are subject to the constraint that

$$|\tau_s(i\infty)| = |\tau_f(0)| \quad (7)$$

Several observations can be made

- a suitable change of time scale in the differential operator, as suggested by (4), makes (6) independent of ϵ .
- for a finite ϵ , a system with transfer T_ϵ does *not* satisfy (5,6). However, there is a one-to-one correspondence between the parameters of T_ϵ and the parameters of τ_s and τ_f when (7) holds. This is essentially similar to the correspondence of the linear parameterization of ARX models and their gain/poles/zeros description.
- if one uses a classic least square method to identify T_ϵ , the excitation matrix, i.e. the Hessian of the cost, is asymptotically degenerate as ϵ tends to zero [9]. Therefore this method is not robust for small ϵ .

- it has been proven in [9] that, if one considers the classical L^2 prediction error as cost for the models (5,6), then its minimum tends to zero when ϵ tends to zero if (y, u) satisfy $y = T_\epsilon u$. Further, the limit excitation matrix is nondegenerate.

In the experiments carried-out on the discussed thermal model, it has been observed that the poles given by the ARX identification provide a good indication of the value of the cutting frequency that should be used to design the low and high pass pre-filters (see eq. (5) and (6)). Figure 1 shows the amplitude Bode plot of the high order model for the heating control, and its value when multiplied by the low-pass and high-pass pre-filters F_l and F_h , respectively used in the following numerical experiments. The filters are Butterworth filters.

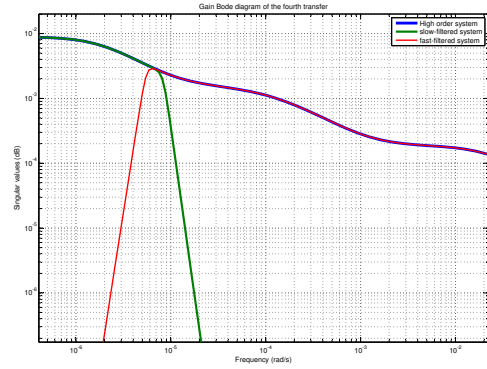


Fig. 1. Global and pre-filtered heating transfers. The right part of the plot, where the slope goes back to -1, is irrelevant to the identification because the data sample rate makes it disappear.

IV. THE PARAMETRIC IDENTIFICATION PROBLEMS

Here, we define optimization problems to perform the identification of the parameters for each model class. Some emphasis is put on the difference between the separated and non separated output data sets.

A. Global ARX model

Using (1), we minimize the L^2 norm of the prediction error, as defined by the difference between the two sides of (1). In practice, we use Matlab's ARX routine to determine optimal coefficients.

B. Two time scales identification with a global measurement of the inside temperature

1) *Parameterization:* The number of poles and zeros of each transfer function has to be set. Since we want a model of order two, we chose a model T_ϵ with two poles, with one pole in τ_s and one pole in τ_f . The third pole that is visible in figure 1 is irrelevant because its time constant is significantly faster than the sampling rate. As in Section III-B, the method requires a slow zero. A fast zero could be considered too. This one is visible in figure 1. It turns out that, for the

data set where the global inside temperature is measured, the best trade-off between robustness and simulation accuracy is achieved by including a fast zero in the fast transfers. Thus the parameterization for the slow and fast models are

$$\tau_s(t) = \frac{1}{s + \alpha} (k_1 s + z_1, \dots, k_4 s + z_4) \quad (8)$$

$$\tau_f(t) = \frac{1}{\beta s + 1} (\rho_1 s + p_1, \dots, \rho_4 s + p_4) \quad (9)$$

2) *The identification problem:* To perform the identification, we follow the two following steps

- Step 1 : Use of high-pass and low-pass pre-filtering data for which the approximations as given by Definition 1 are as accurate as possible.
- Step 2 : Perform separate identifications of τ_n and τ_s under the constraint that

$$|\tau_{si}(i\infty)| = |\tau_{fi}(i0)| \quad (i = 1 \dots 4) \quad (10)$$

Let $y_s(t)$ and $u_s(t)$ (resp. $y_f(t)$ and $u_f(t) = (u_{s1}(t) \dots u_{s4}(t))$) be the low-pass (resp. high-pass) filtered data, then the corresponding differential equations are given by

$$\frac{d}{dt} y_s(t) + \alpha y_s(t) = \sum_{i=1}^4 k_i \frac{d}{dt} u_{si}(t) + z_i u_{si}(t) \quad (11)$$

$$\beta \frac{d}{dt} y_f(t) + y_f(t) = \sum_{i=1}^4 \rho_i \frac{d}{dt} u_{fi}(t) + p_i u_{fi}(t) \quad (12)$$

Using finite differences we obtain, using usual discrete-time notations,

$$\frac{y_s^{k+1} - y_s^k}{\Delta_s} + \alpha y_s^k = \sum_{i=1}^4 k_i \frac{u_{si}^{k+1} - u_{si}^k}{\Delta_s} + z_i u_{si}^k \quad (13)$$

$$\beta \frac{y_f^{k+1} - y_f^k}{\Delta_f} + y_f^k = \sum_{i=1}^4 \rho_i \frac{u_{fi}^{k+1} - u_{fi}^k}{\Delta_f} + p_i u_{fi}^k \quad (14)$$

where Δ_s and Δ_f are rescaling parameters chosen to improve the conditioning of the problem by adapting the finite difference to the considered time scale (see [18]). Note that in (13) the sampling rate may be smaller than Δ_s since y_s has been pre-filtered by a low pass filter.

The problem is linear with respect to the parameters so it is convenient to use a least squares method to identify the two transfer matrices. Moreover, this parameterization of the transfer matrix allows to write the constraints linearly with respect to the parameters as shown in (15)

$$\nu^T = \nu_s^T - \nu_f^T = (k_1 - p_1, \dots, k_4 - p_4) = 0 \quad (15)$$

where the parameters k_i, p_i are appearing in the equations (11), (12), (13) and (14).

3) *Problem statement:* We can now formulate an optimization problem. Given a set of data, the problem is to find the parameters vectors $\theta_s = (k_1 \dots k_4 \ z_1 \dots z_4 \ \alpha)$ and $\theta_f = (p_1 \dots p_4 \ \rho_1 \dots \rho_4 \ \beta)$, corresponding to the

parameters from the equations (11), (12), (13) and (14), by solving the following problem

$$\min_{\theta_s, \theta_f} J_s(\theta_s) + J_f(\theta_f) \quad (16)$$

$$\nu = 0$$

where $J_s(\theta_s)$ (resp. $J_f(\theta_f)$) is the least squares cost of the slow (resp. fast) matrix transfer given by

$$J_s(\theta_s) = \frac{1}{M} \sum_1^M \hat{\epsilon}_s^2[k, \theta_s] \quad (17)$$

$$J_f(\theta_f) = \frac{1}{M} \sum_1^M \hat{\epsilon}_f^2[k, \theta_f] \quad (18)$$

where

$$\hat{\epsilon}_s^2[k, \theta_s] = \frac{y_s^{k+1} - y_s^k}{\Delta_s} - \varphi_s[k] \theta_s \quad (19)$$

$$\hat{\epsilon}_f^2[k, \theta_f] = y_f^k - \varphi_f[k] \theta_f \quad (20)$$

$$\varphi_s[k] = \left(\frac{u_{s1}^{k+1} - u_{s1}^k}{\Delta_s}, \dots, \frac{u_{s4}^{k+1} - u_{s4}^k}{\Delta_s} \dots \right. \\ \left. u_{s1}^k, \dots, u_{s4}^k, -y_s^k \right) \quad (21)$$

$$\varphi_f[k] = \left(u_{f1}^k, \dots, u_{f4}^k, \frac{u_{f1}^{k+1} - u_{f1}^k}{\Delta_f}, \dots, \frac{u_{f4}^{k+1} - u_{f4}^k}{\Delta_f} \dots \right. \\ \left. - \frac{y_f^{k+1} - y_f^k}{\Delta_f} \right) \quad (22)$$

In [9] it has been proved that, if the real transfer is indeed T_ϵ , the minimum of (16) is asymptotically reached (as ϵ tends to zero) by the parameters corresponding to the slow and fast transfers. Moreover, the Hessians of J_s and J_f are not degenerate when ϵ tends to zero.

4) *Problem solving:* While this is not a requirement, we chose to solve problem (16) with Uzawa's algorithm (see [19]). Its main feature is that, at the minimization stage, each subproblem is very similar to an identification problem on the relevant frequency range, (see [15]), in the sense that the Hessian of the inner optimization problem is a matrix that contains the signals covariance. Moreover, the gradient step of the maximization problem is adapted to each constraint.

C. Two time scales identification with a separation of the influences of each input

Using a data set which is different from the data set used in the previous section leads to a different tradeoff between accuracy and robustness. Indeed, we have observed that for separated outputs it was best to make some of the fast zeros "vanish" from the parameterization.

1) *Parameterization:* Even if using the two time scaled method to identify the system allows a clear improvement of the results in terms of simulation errors and parameters identification, as compared to the classical least squares method, an even better identification can be achieved. One explanation is that the system has four inputs and just one output. These inputs are really poorly balanced and some of

them do not excite the system in an appropriate frequency range. For instance, the solar fluxes are almost perfectly 24 hours-periodic signals. Therefore, it is difficult to clearly identify the influence of these inputs on the temperature inside the building. That is why in this part we now separate the influence of each input on the temperature. Instead of identifying a transfer matrix we identify four separate transfer functions. This method is referred to as the separated time scaled method. Because we look for a second order model we have to impose that the four slow (resp. fast) transfer share the same poles.

Observe that, even if the model class may appear similar to the one in Section IV-B (once the equality of the poles in the four transfers is duly accounted for), the cost that we will minimize in (32) is not the same as in (16), because we add four prediction error costs.

In other words, the sum of the excitation matrices of four signals is different from the excitation matrix of the sum of these four signals.

In this case, we have four transfer functions $T_i(s) = T_{si}(s/\epsilon)T_{fi}(s)$ ($i = 1 \dots 4$). Each transfer can be decomposed into a fast and a slow transfer as mentioned in Definition 1. We now separately identify the four slow (resp. fast) sub-systems in their own time scale under the following constraints :

- for each transfer function the high frequency gain of the slow system must be equal to the static gain of the fast system ($|\tau_{sj}(i\infty)| = |\tau_{fj}(i0)|$ $j = 1 \dots 4$)
- the fast (resp. slow) sub-systems share the same poles.

To perform the identification, we follow the two following steps

- Step 1 : use of high-pass and low-pass pre-filtering data for which the approximations as given by definition (1) are accurate.
- Step 2 : perform separate identifications of τ_n and τ_s under the constraint that $|\tau_{sj}(i\infty)| = |\tau_{fj}(i0)|$ ($j = 1 \dots 4$) and that the transfer functions τ_{sj} (resp. τ_{fj}) share the same poles.

Let $y_{si}(t)$ and $u_{si}(t)$ (resp. $y_{fi}(t)$ and $u_{fi}(t)$) be the low-pass (resp. high-pass) filtered simulations data of the i^{th} transfer function, then the corresponding differential equations are given by a slow subsystem

$$\frac{d}{dt}y_{si}(t) + \alpha_i y_{si}(t) = k_i \frac{d}{dt}u_{si}(t) + z_i u_{si}(t) \quad (i = 1 \dots 4) \quad (23)$$

and a fast subsystem

$$\beta_i \frac{d}{dt}y_{fi}(t) + y_{fi}(t) = p_i u_{fi}(t) \quad (i = 1 \dots 3) \quad (24)$$

$$\beta_4 \frac{d}{dt}y_{f4}(t) + y_{f4}(t) = \rho_4 \frac{d}{dt}u_{f4}(t) + p_4 u_{f4}(t) \quad (25)$$

This model class has been found to achieve the best trade-off between robustness and simulation accuracy. In particular, deleting the zeros in (24) achieves the best trade off between robustness and simulation accuracy.

Using finite differences we have, using the same notations employed in Section IV-B

$$\frac{y_{si}^{k+1} - y_{si}^k}{\Delta_{si}} + \alpha_i y_{si}^k = k_i \frac{u_{si}^{k+1} - u_{si}^k}{\Delta_{si}} + z_i u_{si}^k \quad (26)$$

$(i = 1 \dots 4)$

$$\beta_i \frac{y_{fi}^{k+1} - y_{fi}^k}{\Delta_{fi}} + y_{fi}^k = p_i u_{fi}^k \quad (i = 1 \dots 3) \quad (27)$$

$$\beta_4 \frac{y_{f4}^{k+1} - y_{f4}^k}{\Delta_{f4}} + y_{f4}^k = \rho_4 \frac{u_{f4}^{k+1} - u_{f4}^k}{\Delta_{f4}} + p_4 u_{f4}^k \quad (28)$$

Once again, the constraints can be expressed linearly with respect to the parameters. Actually, the constraints of the identification problem are :

$$\alpha_i - \alpha_{i+1} = 0, \quad i = 1 \dots 3 \quad (29)$$

$$\beta_i - \beta_{i+1} = 0, \quad i = 1 \dots 3 \quad (30)$$

$$k_i - p_i = 0, \quad i = 1 \dots 4 \quad (31)$$

Thus, the vector of constraints $\nu = \nu_s - \nu_f$ is given by the concatenation of the ten equalities given by (29), (30) and (31).

2) *Problem statement:* Given a set of data, the problem is to find the four parameters vectors $\theta_{si} = (k_i \ z_i \ \alpha_i)^T$, the three $\theta_{fi} = (p_i \ \beta_i)^T$ ($i = 1 \dots 3$) and $\theta_{f4} = (p_4 \ \rho_4 \ \beta_4)^T$ by solving the following problem

$$\min_{\substack{\theta_{si}, \theta_{fi} \\ \nu = 0}} \sum_{i=1}^4 J_{si}(\theta_{si}) + J_{fi}(\theta_{fi}) \quad (32)$$

where $J_{si}(\theta_{si})$ (resp. $J_{fi}(\theta_{fi})$) is the least squares cost of the i^{th} slow (resp. fast) transfer function.

$$J_{si}(\theta_{si}) = \frac{1}{M} \sum_{k=1}^M \hat{\epsilon}_{si}^2[k, \theta_{si}] \quad (33)$$

$$J_{fi}(\theta_{fi}) = \frac{1}{M} \sum_{k=1}^M \hat{\epsilon}_{fi}^2[k, \theta_{fi}] \quad (34)$$

where

$$\hat{\epsilon}_{si}^2[k, \theta_{si}] = \frac{y_{si}^{k+1} - y_{si}^k}{\Delta_{si}} - \varphi_{si}[k] \theta_{si} \quad (35)$$

$$\hat{\epsilon}_{fi}^2[k, \theta_{fi}] = y_{fi}^k - \varphi_{fi}[k] \theta_{fi} \quad (36)$$

$$\varphi_{si}[k] = \left(\frac{u_{si}^{k+1} - u_{si}^k}{\Delta_{si}} \quad u_{si}^k \quad -y_{si}^k \right) \quad (37)$$

$$\varphi_{fi}[k] = \left(u_{fi}^k \quad -\frac{y_{fi}^{k+1} - y_{fi}^k}{\Delta_{fi}} \right) \quad (i = 1 \dots 3) \quad (38)$$

$$\varphi_{f4}[k] = \left(u_{f4}^k \quad \frac{u_{f4}^{k+1} - u_{f4}^k}{\Delta_{f4}} \quad -\frac{y_{f4}^{k+1} - y_{f4}^k}{\Delta_{f4}} \right) \quad (39)$$

3) *Problem solving:* Here again, we use Uzawa's algorithm to solve this problem.

V. NUMERICAL RESULTS

A. Conditioning of the problems

To perform a robust parameter identification, the Hessian of the optimization problem has to be well conditioned (see [15]). Yet, a two-time scaled system usually induces bad

conditioning (see [9]). The time scaled identification has been designed to improve the conditioning of the optimization problem. For the classical least squares method there is one conditioning number, while there are two conditioning numbers for the time-scaled method (one for the slow transfer matrix and one for the fast one), and there are eight conditioning numbers for the separated time scaled method (one for each subsystem). The conditioning numbers are given in the table II. We use everywhere data without noise corruption resulting from the high order model (see section II).

TABLE II
CONDITIONING NUMBERS

	least squares identification	Time scaled identification	Separated time scaled
conditioning numbers	$r_{LS} = 2.6/10^{10}$	$r_s = 1.4/10^8$ $r_f = 1.4/10^9$	$r_{s1} = 0.0011$ $r_{s2} = 0.00083$ $r_{s3} = 0.00078$ $r_{s4} = 0.0016$ $r_{f1} = 0.043$ $r_{f2} = 0.013$ $r_{f3} = 0.037$ $r_{f4} = 0.020$

As one can see it on Table II, the separated time scaled method improves the conditioning of the problem. But, we can also see that using a non separated time scaled method does not improve the conditioning numbers as well as the previous method. Having separated outputs provides extra information on the system as we have virtually three extra sensors.

The bad conditioning of the least squares method (ARX) is highly problematic because the results of the identification are very poor, in simulation results and in parameter identification. Concerning the non separated time scaled method, in the following, we will see that, despite the bad conditioning of the system, this method yields better results in simulation and in parameter identification than the least squares method. On the other side, it will be seen that this method fails to estimate the location of the zeros of the system, particularly when the identification is performed using noisy data.

Finally, we can see that the separation of the transfers allows us to normalize the problem and then to improve the conditioning numbers. As a result, this method is really robust with respect to noises and consistent results² in parameter identification are obtained whether noisy or noise free data are used.

B. Simulation results

1) *Simulation protocol:* This protocol is decomposed in four steps:

- 1) Using noise free input data described in section II and using the high order model, we get the four noise free corresponding outputs.

²This comparison is made with noises which have the same statistical properties.

- 2) Then, we perform a first identification using the previous data.
- 3) Further, we add independent noises on the inputs and the outputs collected from the first step. Then, we perform three identifications using this noisy data and the three identification methods.
- 4) Finally a validation step is performed. We simulate all the models from 2 and 3 using noise free inputs to obtain the global temperature of the building. We compare these temperature to the global output of step 1. The Table III gives some statistical properties of the simulation error between the global temperature from the high order model and the global temperature of each of the six identified models.

2) *Results:* Figure 2 shows the errors of simulation between the high order reference model and the three identified systems, the latter being identified using noisy data. These simulations are performed using noise free inputs over 25 days. As can be seen in Figure 2 the ARX model identified using MATLAB's identification toolbox does not give good results in terms of simulation errors. Moreover, one can see in Figure 2 that the standard deviation of the simulation error seems to be better with the separation of the influences of the inputs than without.

TABLE III
COMPARISON OF STATISTICAL PROPERTIES OF THE SIMULATION ERROR WITH RESPECT TO THE NOISE FREE SIMULATION USING THE HIGH ORDER MODEL

	Stat. prop.	Least squares identification	Time scaled identification	Separated time scaled
Noise free data	Mean	-0.0989	-0.0024	-0.0088
	std dev	0.93	0.35	0.266
Noisy data	Mean	0.0461	0.0021	-0.0045
	std dev	0.641	0.49	0.34

a) *Using noise free data:* The table III shows some statistical properties of the simulation errors of the three identified systems. Considering the identification using noise free data, one can see that the best results are obtained by the time-scaled methods. Indeed, the statistical properties of the simulation error obtained with the time scaled methods are similar. One can also notice that the worst results are clearly obtained with the ARX model.

b) *Using noisy data:* Let us focus on the identification using noisy data. one can see that the best results are again achieved with the time-scaled methods. One can also observe that the deterioration of the standard deviation is less important when we separate the influences of the input than with a global measurement of the temperature. We can also notice that even if the ARX model is still the worst, the addition of noises has clearly improved its performances.

3) *Conclusion:* One can see that using a time scaled method allows a good improvement of the results in terms of simulation error. The comparison between the least squares method and the time scaled method shows that the results are better using the latter.

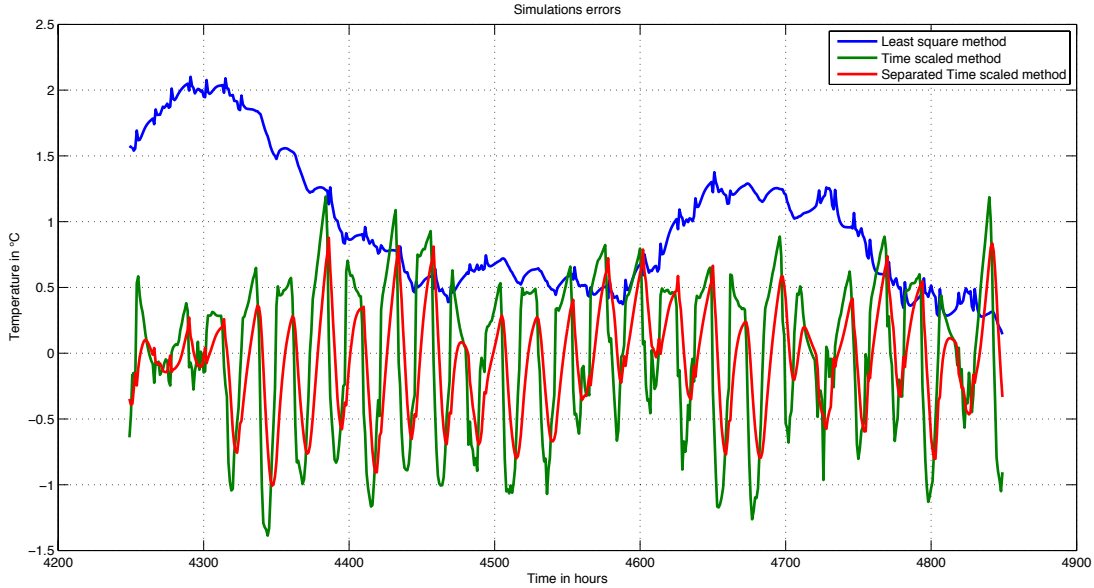


Fig. 2. Comparison of simulations errors using noise free data obtained with the three identified models which have been identified using noisy data. The reference output is computed with the high order model using noise free data. This figure shows the benefits in terms of simulation error provided by using the time scaled identification. This plot represent a 25 days simulation extracted from a whole year simulation.

C. Static gains, poles and zeros identification

Concerning poles and zeros, we give, in tables VI, VII and VIII the corresponding time constants. Those time constants are calculated using the discrete model provided by equations (1), (13), (14), (26), (27) and (28)

1) *Static gains identification:* Let us see the results of the three identification on the static gains

TABLE IV

COMPARISON OF THE IDENTIFIED STATIC GAIN USING NOISE FREE DATA

	High order model	least squares identification	Time scaled identification	Separated time scaled
Gain 1	1	0.903	1.0013	1.0004
Gain 2	0.0088	0.0058	0.0089	0.0088
Gain 3	$6.75 \cdot 10^{-5}$	$5.68 \cdot 10^{-4}$	$5.48 \cdot 10^{-5}$	$6.75 \cdot 10^{-5}$
Gain 4	0.009	0.0125	0.009	0.009

TABLE V

COMPARISON OF THE IDENTIFIED STATIC GAIN USING NOISY DATA

	High order model	least squares identification	Time scaled identification	Separated time scaled
Gain 1	1	0.959	1.0008	1.0007
Gain 2	0.0088	0.006	0.0088	0.0088
Gain 3	$6.75 \cdot 10^{-5}$	$5 \cdot 10^{-4}$	$6.73 \cdot 10^{-5}$	$6.75 \cdot 10^{-5}$
Gain 4	0.009	0.0107	0.009	0.009

As one can see on the tables IV and V, using a time-scaled identification method yields a substantial improvement compared to the classical least squares method. In fact,

the classical least squares method never correctly estimates the static gains whereas the time-scaled methods estimate the gains of the transfer matrix adequately.

Moreover, the separation of the influences of each input allows one to reach the same accuracy using noisy or noise free data, whereas the others methods give better results using noisy data.

2) *Time constants and zeros location:*

a) *Time constants identification.:* The table VI reports the identified time constants using noisy and noise free data. We can see that with and without noises the identification of the two time constants of the system are similar using the time-scaled methods, whereas the ARX model provides time constants quite different of the other models. Since the simulation results are better with the models identified by time scaled methods, one can suppose that the time constants are well identified by these methods.

TABLE VI

IDENTIFIED TIME CONSTANT IN HOURS

		least squares identification	Time scaled identification	Separated time scaled
Noise free data	Slow	117	143	147
	Fast	1.1	2.5	2.9
noisy data	Slow	108	144	147
	Fast	1.2	2.4	2

b) *Zeros time constants:* The tables VII and VIII give the zeros time constants of each transfer using respectively noise free and noisy data to perform the identification. As one can see on these tables, the only method yielding a weak dispersion of the identified parameters is the time scaled

identification with separation of the influences of the inputs. Moreover, using Moore's method to reduce the system does not keep the two time scaled structure of the system. Indeed, looking at the Bode diagram of the reduced system, one can notice that the two time scaled structure exhibited by both the high order model and the identified one is not preserved by the reduced one.

TABLE VII

IDENTIFIED ZEROS' LOCATION USING NOISE FREE DATA. THE † SYMBOL MEANS THAT THE ZERO HAS BEEN FOUND TO BE UNSTABLE.

		least squares identification	Time scaled identification	Separated time scaled
First Transfer	Slow zero	7.8	14.8	24
	Fast zero	None	4.6	None
Second Transfer	Slow zero	2.1	9.9	5.1
	Fast zero	None	0.85	None
Third Transfer	Slow zero	1.65	197†	11.1
	Fast zero	None	0.44	None
Fourth Transfer	Slow zero	9.99	21.3	24.9
	Fast zero	None	0.051†	0.76

TABLE VIII

IDENTIFIED ZEROS' LOCATION USING NOISY DATA. THE † SYMBOL MEANS THAT THE ZERO HAS BEEN FOUND TO BE UNSTABLE.

		least squares identification	Time scaled identification	Separated time scaled
First Transfer	Slow zero	5.3	15.5	23.2
	Fast zero	None	4.4	None
Second Transfer	Slow zero	2.8	9.9	5.2
	Fast zero	None	1.1	None
Third Transfer	Slow zero	1.9	116†	11.6
	Fast zero	None	0.023†	None
Fourth Transfer	Slow zero	11.2	22.6	24.3
	Fast zero	None	0.019†	0.44

3) *Conclusion:* The least squares method does not really identify static gains, time constants and zeros of the system, the time scaled method with global measurement allows us to identify the static gains and the time constants, but shows poor results in the identification of the zeros. Finally, the time-scaled method with separation of each inputs allows to identify all these parameters with robustness to noises.

VI. CONCLUSION

In this paper, it was shown that using a time-scaled identification method (see [9]) allows a substantial improvement of the model identification compared to a classical least squares method, in terms of prediction error and of parameters sensitivity to measurement noises. It was also emphasized that to clearly identify a time-scaled system it is needed to find a good compromise between simulation error and robustness to noise. The time-scaled method allows a great improvement in the search of this compromise compared to the least squares method.

Nevertheless, it has been observed that the identification of the zeros of the system is too sensitive to the noises using that method. To improve the conditioning of the system the inputs

and the output should be normalized. Using a separation of the influences of each input is a solution. Then, it has been shown that separating the influences of the inputs and using a time scaled method can provide a good compromise between identification error and robustness toward noises since the results obtained with or without noises are quite similar.

In summary, this work proposes an efficient method, based on a two time scale models to identify a low order linear model describing the thermal behaviour of the system. This efficiency is measured in terms of simulation errors and in terms of robustness of the parameters identification to noises. This is due to the normalization of the two time scale problems, both in magnitude of the signals and in their frequency range. The model obtained by this method can be used in simulation and it can also be used in constrained optimal control since the parameters are well identified.

REFERENCES

- [1] J. Bonnans and T. Guilbaud, "Using logarithmic penalties in the shooting algorithm for optimal control problems," *Optim. Control Appl. Meth.*, vol. 24, pp. 257–278, 2003.
- [2] R. Hartl, S. Sethi, and R. Vickson, "A survey of the maximum principles for optimal control problems with state constraints," *SIAM review*, vol. 37, no. 2, pp. 181–218, 1995.
- [3] M. Gouda, S. Danaher, and U. C.P., "Building thermal model reduction using nonlinear constrained optimization," *Building and Environment*, vol. 37, pp. 1017–1031, 2002.
- [4] G. Fraisse, C. Viardot, O. Lafabrie, and G. Achard, "Development of a simplified and accurate building model based on electrical analogy," *Energy and Buildings*, vol. 34, pp. 1017–1031, 2002.
- [5] M. Jiménez, H. Madsen, and K. Andersen, "Identification of the main thermal characteristics of building components using matlab," *Building and Environment*, vol. 43, pp. 170–180, 2008.
- [6] M. Jiménez and H. Madsen, "Models for describing the thermal characteristics of building components," *Building and Environment*, vol. 43, pp. 152–162, 2008.
- [7] J. Stoer and R. Burlish, *Introduction to Numerical Analysis*. Springer-Verlag, 1993.
- [8] A. Chiuso and G. Picci, "On the ill-conditioning of subspace identification with input," *Automatica*, vol. 40, pp. 575–589, 2004.
- [9] F. Chaplais and A. Alaoui El Azher, "Two time scaled parameter identification by coordination of local identifiers," *Automatica*, vol. 32, no. 9, pp. 1303–1309, 1996.
- [10] M. Basseville, A. Benveniste, and S. Willsky, "Multiscale autoregressive processes. part i: Schur-levinson parameterizations," *IEEE Trans. Aut. Control*, vol. 40, no. 8, pp. 1915–1934, 1992.
- [11] —, "Multiscale autoregressive processes. part ii: Lattice structures for whitening and modelling," *IEEE Trans. Aut. Control*, vol. 40, no. 8, pp. 1935–1954, 1992.
- [12] G. Stephanopoulos, O. Karsligil, and M. Dyer, "Multiscale theory for linear dynamic processes. part I. foundations," *Computers and Chemical Engineering*, vol. 32, pp. 857–884, 2008.
- [13] A. Sodja and B. Zupančič, "Modelling thermal processes in buildings using an object-oriented approach and modelica," *Simulation Modelling Practice and Theory*, vol. 17, pp. 1143–1159, 2009.
- [14] A. Le Mouel, "Contribution à l'étude des caractéristiques dynamiques réduites des systèmes thermiques complexes," Ph.D. dissertation, Ecole des Mines de Paris, 1995.
- [15] L. Ljung, *System Identification, Theory for the User*. Prentice-Hall, 1987.
- [16] D. Luse and H. Khalil, "Frequency domain results for systems with slow and fast dynamics," *IEEE Trans. Aut. Control*, vol. AC-30, no. 12, pp. 1171–1178, 1985.
- [17] T. Kailath, *Linear Systems*. Prentice Hall, 1980.
- [18] R. Middleton and G. Goodwin, "Improved finite word length characteristic in digital control using delta operators," *IEEE Trans. Aut. Control*, vol. IC-31, no. 1, pp. 1015–1021, 1986.
- [19] A. K. J., L. Hurwicz, and H. Uzawa, *Studies in Linear and Non Linear Programming*. Stanford University Press, CA, 1972.