Model Tuning Approach For Energy Management
Of Office and Apartment Settings

Lisa Scanu[1,2], Stephane Ploix[2], Pierre Bernaud[1], Etienne Wurtz [1]
[1]Univ. Grenoble Alpes, CEA, LITEN, INES, LGEB, F-38000 Grenoble, France
[2]Univ. Grenoble Alpes, CNRS, Grenoble INP, GSCOP, F-38000 Grenoble, France

Abstract

This paper presents a methodology for parameter estimation of a suitable model for energy management services in office and apartment settings. The objective of this work is to identify model structures and tune parameters to fit recorded data. Once these parameters have been identified , the model will support energy services (prediction, explanation). Grey-box models are proposed to estimate the temperature and CO_2 concentration based only on few sensors. Then, two different estimation parameter methods have been applied relying on a descent algorithm or a genetic one. Different structures have been compared. Finally, the resulting model has been applied to two case studies: an office and an apartment.

1 Introduction

As buildings become more energy-efficient, the effect of human activities on their global consumption increases. To limit this impact, a first attempt was made to increase the level of automation in buildings. However, from surveys led in residential and office buildings, Carassus (2011) and Sidler (2011) conclude that automation does not reduce the gap between predicted and effective consumption. This work reveals that static rules can not be sufficient. In order to reach the goal of low energy consumption in the operation phase of buildings, occupants must then be involved in the energy management process. The proposed work is part of a research project focusing on a system embedding energy services with a high level of interaction (called "e-consultant") to empower users with energy services. The role of these energy services is to allow users to define their own objective and consequently generate information and advice concerning the best controls to apply on active systems of building (HVAC systems, sun-shading devices, windows opening,). An analysis of the sociological literature was made in order to implement relevant services for the users and favouring real and durable changes in their behaviours.

The identified services require thermal and aeraulic behavioural models easy to set up. The work presented in this paper focuses on the development and implementation of such a model. It involves to deal with: the definition of model structures and the parameter estimation method with validation on case studies. The model must be simple enough to ease its set-up with estimation and prediction capabilities. The limits in terms of complexity and computational time are discussed later. This e-consultant should target both old and new dwellings and be configured by end-users thus it must not require much expert knowledge. Consequently, some sensors are needed to estimate the thermal and aeraulic model. Tests have been performed to limit the instrumentation requirements. This paper describes a methodology to determine relevant model structures on different case studies. These structures should be generic: it should adapt to both mono-zone and multi-zone case studies and it should be based on a minimal number of sensors and requiring little expert knowledge.

The paper is organized as follows. The office setting is described in Section 2. Then, the two different parameters estimation methods are presented : meta-optimization in section 3 and genetic algorithms in section 4. The latter section is devoted to a complex multi-zones case study for validation: an apartment.

2 Proposed approach

2.1 First study case: Mono-zone office

In order to validate model structure, a real case study has been chosen: an office occupied by 1 to 4 persons. In terms of instrumentation, the office is equipped with almost 200 sensors. The office itself is equipped with 26 sensors measuring temperatures, CO_2 and COV concentrations, illuminance, humidity and power consumption, door and windows contacts, motions and weather (*cf.* figure 1). Only indoor temperature, CO_2 sensors, window and door contacts and weather station for outdoor temperature and nebulosity are used to estimate the parameters of the models. Some estimators were defined for occupancy and heating during winter season. For occupancy, power-consumption measurements are used to estimate the number of in-use computers and thus the number of occupants. Regarding the heating power, a temperature sensor has been placed on the surface of the
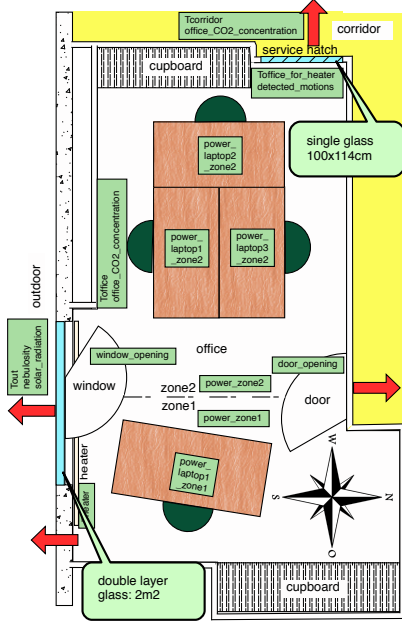
Figure 1: Map of the office with sensors

heater and the power heating is computed using equation 1

$$P_{heat} = K(T_{heater} - T_{office}) \qquad (1)$$

where K is estimated.

2.2 Literature review

Semi-physical models are considered because of their limited complexity. A literature review reveals that the use of electric analogy modelling is prevalent among grey box models (Hazyuk et al. (2012), Mathews et al. (1994), Bacher and Madsen (2011)) since they offer a good representation of the different components of the buildings and a high ease of use.

To obtain models requiring little computations, a first very simple structure has been implemented, information being then added so as to meet an adequate accuracy. Opening of doors and windows are taken into consideration since Fabi et al. (2012) claim that the opening of windows is one of the user-accessible adaptive actions most impacting energy consumption, comfort and air quality. They are the occupant possible actions during summer period. Castillo et al. (2014) also came to the conclusion that taking into account the opening of the door on the corridor led to significant improvements in their results. The simple structure is built by locating the whole inertia in a fictive wall represented by a single equivalent capacitor (see figure 2). The model with 4 capacitors is the most complex model tested in this study (see figure 3) taking into account the available data and the instrumentation constraint. It includes simultaneously the capacitor of the fictive wall, the ones of the inner and exterior walls and the one of the indoor air which represent the dynamics of the building. Between those two structures, 7 others were developed, deleting the

inertia of the reference model or adding more capacitors representing the inertia of the air and the inner or outer wall. The different structures are summarized in table 1 and the RC-networks are displayed in Annexe A.
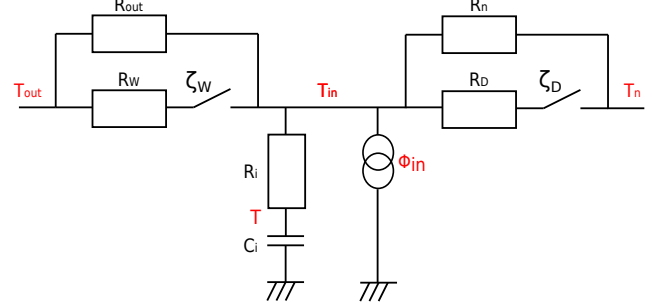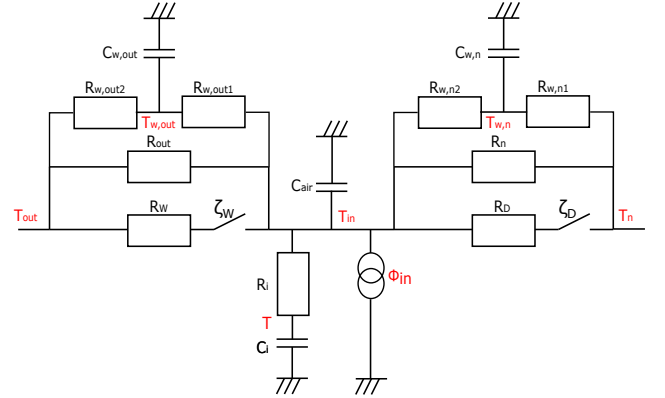

Figure 2: Reference model for the office

Figure 3: Model with 4 capacitors

The full model includes 4 state variables, each one representing the temperature in a part of the study case:

- T_{in}: The temperature of the office ambient air, which is used as the model output
- τ: The temperature of a fictive wall
- $T_{w,n}$: The surface temperature of the inner wall
- $T_{w,out}$: The surface temperature of the outer wall

The parameters of the model represent different thermal properties of the building. This includes thermal resistances:

- R_i: between the fictive wall average temperature and the wall surface
- R_W: between the exterior and the interior through the window
- R_{out}: between the exterior and the interior
- $R_{w,out1}$: between the average wall temperature and the inner surface of the exterior wall
- $R_{w,out2}$: between the average wall temperature and outer surface of the exterior wall
- R_n: between the adjacent room and the interior
- R_D: between the adjacent room and the interior through the door

- $R_{w,n1}$: between the average wall temperature and inner surface of the interior wall
- $R_{w,n2}$: between the average wall temperature and outer surface of the interior wall

The heat capacities of different parts of the building are represented by:

- C_i: for the fictive wall
- C_{air}: for the interior
- $C_{w,out}$: for the outside wall
- $C_{w,n}$: for the interior wall

Door and window openings are represented respectively by ζ_D and ζ_W standing for opening time ratios, ϕ_{in} represents the internal gains: solar, electric and occupancy gains.

Thermal models are coupled with an aeraulic model presented in figure 4 and described by equation 2 .
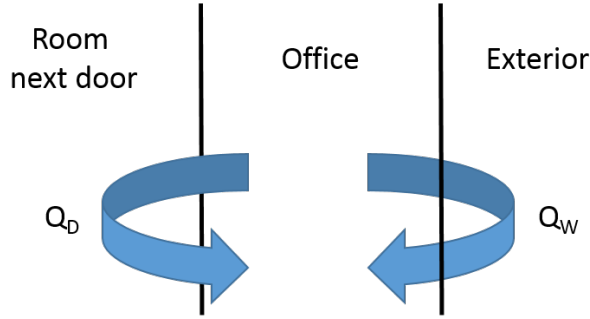

Figure 4: Aeraulic model

$$V\frac{d\Gamma_{in}}{dt} = -(Q_W+Q_D)\Gamma_{in}+Q_W\Gamma_{out}+Q_D\Gamma_n+S_{CO_2}^{body}n \tag{2}$$

where Γ_{in}, Γ_{out} and Γ_n represent the CO_2 concentrations of the office, the outdoor and the neighbourhood, $S_{CO_2}^{body}$ the average production of CO_2 per person and n the number of occupants in the room.

2.3 Models parameter estimation

Using physical models for energy management in buildings requires a non-linear optimisation step which consists in estimating the parameters of the models using recorded data during a training period. During a validation period, the estimated parameters are then used to check the error between the predicted and the measured output as an objective function. Different parameter estimation methods exist: some of them require expert knowledge to select a relevant initial dataset such as usual descent algorithms in a rough parameter space while others require a linear model such as linear regressions. Since the models used in this study are not linear and do not rely on expert knowledge some methods can be eliminated. Non-linear optimization algorithms considered in this work are meta-heuristic and descent algorithms. The descent algorithms are based on the computation of the gradient and sometimes Hessian of the objective

function. Thus, according to an initial point, they can be stuck in a local optimum but they are less costly in terms of computational time than meta-heuristic algorithms. That is why a "meta-optimization" procedure has been implemented in order to try to avoid that problem while keeping a low computational cost. Genetic algorithms have been used because they find the global optimum with a sufficient computational time but they are usually fairly more costly than descent algorithms.

2.4 Models validation and selection methods

Once the parameters estimated, several model selection and validation procedures can be used. They differ according to the objective(s) of the model. In this work, the chosen procedure is inspired from the work of Bacher and Madsen (2011) and the PhD thesis of Reynders (2015). It consists in validating the robustness and the accuracy of the identified model but also in selecting the most suitable order.

2.4.1 Tools

To do so, different tools are needed: the standardized version of the Root Mean Square Error (sRMSE) and a sensitivity analysis. The RMSE enables to quantify the forecast error between the indoor simulated temperature and the measured one. From here, the sRMSE is computed per equation 3 allowing the comparison of the model on different datasets.

$$sRMSE = \frac{\sqrt{\frac{1}{N}\sum_{k=1}^{N}(\hat{y}_k - y_k)^2}}{y_{max} - y_{min}} \tag{3}$$

where y_k represents the measured value and $\hat{y}$ the predicted one of the indoor temperature.

A sensitivity analysis is performed in order to ensure that each parameter is useful for forecasting the temperature and CO_2 concentrations in the case study. Different methods exist such as the Sobol indices described by Sobol (2001) or the one described by Saltelli et al. (1999) but both require a lot of simulations. Morris (1991) developed another method requiring fewer simulations and still a few hypotheses, which makes it faster than the Sobol and FAST (Fourier Amplitude Sensitivity Testing) enabling to take into consideration the interactions between the parameters. As the goal of the sensitivity analysis in this research is to ensure that the structure of the model is consistent and that each parameter has an impact on the output, it is not completely suitable. For a first estimation, the method of Morris is implemented but the method of Sobol should be further investigated.

2.4.2 Validation method

The validation method consists in different steps:

- Validity of the model during estimation phase
- Capacity of the model to predict future behaviour

Table 1: Summary of studied models

Models	States	Observations	N
Reference	τT_{in}	$T_n T_{out} \zeta_D \zeta_W$	12
Model0C	-	$T_n T_{out} \zeta_D \zeta_W$	10
Model2Cair	τT_{in}	$T_n T_{out} \zeta_D \zeta_W$	13
Model2CwallT$_{out}$	$\tau T_{w,out}$	$T_n T_{out} \zeta_D \zeta_W$	15
Model2CwallT$_n$	$\tau T_{w,n}$	$T_n T_{out} \zeta_D \zeta_W$	15
Model3CT$_{out}$	$\tau T_{in} T_{w,out}$	$T_n T_{out} \zeta_D \zeta_W$	16
Model3CT$_n$	$\tau T_{in} T_{w,n}$	$T_n T_{out} \zeta_D \zeta_W$	16
Model3Cwall	$\tau T_{w,out} T_{w,n}$	$T_n T_{out} \zeta_D \zeta_W$	18
Model4C	$\tau T_{in} T_{w,out} T_{w,n}$	$T_n T_{out} \zeta_D \zeta_W$	19

- Stability of the model along seasonal variations

For the first two steps, the sRMSE is computed and a limit to 0.1 is set up to validate a model. Then, for the stability, the higher value on the validation phase is chosen between both summer and winter scenario. A range of $\pm 20\%$ of this value is defined and if the other value does not fit in that range, the model is considered as not valid.

2.4.3 Selection method

If the models verify every condition of the validation, then they are compared to each other in order to select the most suitable one. For that purpose, the sum of the sRMSE obtained for both summer and winter will be computed and then compared. The selected model will be the one with the lower value. To complete the selection process a sensitivity analysis will be performed to ensure that all parameters are significant.

3 Meta optimisation

3.1 Principle

The meta-optimization approach chosen has been presented by Le Mounier (2016) in her PhD thesis. The goal is to guarantee the physical validity of parameters and to guide the optimization in order to improve the results. It requires an *a priori* knowledge of the values of the parameters in order to initialize them. Parameters are initialized set at values with of the right order of magnitude. Then, lower and upper bounds are define very broad in order both to give more freedom to the algorithm and to limit the impact of the choice of initial values. Thereafter, the optimization is guided towards zones where the objective function is not monotonous in the parametric directions. The procedure consists in the implementation of a series of successive optimizations and in the selection of set of parameters minimizing the errors without letting them reach their bounds.

3.2 Limits

Thus this approach was applied to the different models but several limits appeared. First, the meta optimisation was applied to the model with one capacity and the results were good with a sRMSE between

0.06 and 0.08 according to the seasons and a computational time around of approx. 3 minutes. Then, the complexity of the models was increased step by step and the models with two capacitors were implemented. For example, for the model with an air capacitance, the sRMSE reached 0.7 - 0.8 with a computational time peaking 28 hours. Further investigation revealed that the algorithm was not able to converge to an optimum. It also showed that the procedure is not ergodic. Indeed, 406 simulations were run with 406 different values to initialize the parameters of the model. The results can be seen in figure 5: it can be noticed that depending on the initial point, the absolute average error can vary strongly reaching values up to 20°C, obviously widely outside the acceptable limit for an energy management model. According to the results and limits presented in that section, it seems obvious that this method can not be used here. Then, another method of parameter estimation is investigated allowing to compare models and to select a suitable thermal structure.
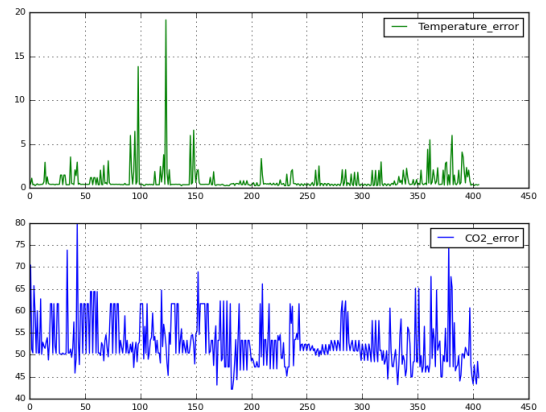

Figure 5: Absolute average error for 406 simulations

4 Genetic algorithm

To solve the problems of convergence and sensitivity to the initial point, it was considered to use genetic algorithms to free the models from the problems of convergence although they are usually more

costly in terms of computational time. Furthermore, the choice was made of a multi-objective algorithm to deal separately with CO_2 and temperature objectives. Several algorithms have been developed, such as the VEGA algorithm (Vectorial Evaluation Genetic Algorithm) Schaffer (1985), the NPGA (Niched Pareto Genetic Algorithm) which resorted mainly to a selection based on Pareto domination (*cf.* Horn et al. (1994)). Later, it was superseded by the algorithm NPGA2 which uses the degree of domination of an individual as a deciding factor for the tournament selection Erickson et al. (2001). Also available is the algorithm NSGA-II (Non Dominated Sorting Genetic Algorithm), an elitist genetic algorithm introduced by Deb et al. (2000) in the 2000s and based on a classification of individuals in several levels. This last algorithm speeds up the sorting process in comparison of the algorithm NSGA while preserving the diversity of populations (Shaikh et al. (2014)); it has been chosen in this study.

4.1 Principle

The basic principle of this algorithm lies on four main steps: creation of the initial population, selection of individuals, mutations and crossover operations and computation of the selection criteria. The sorting of the population is done according to the criteria of the Pareto domination favouring individuals with a high crowding distance so as to maintain a good diversity in the population.

4.2 Training and validation dataset

The summer scenario consists in an estimation of the model on the month of May 2015 and in its validation on the month of June 2015. It is a month of heavy heat in Grenoble since the outdoor temperatures oscillates between 20 and 32°C (*cf.* 6a). Measured data involves a great number of openings of the windows and of the door which have a clear impact on the office temperature and CO_2 concentrations. The winter scenario consists in estimating the model on the month of October 2015 and in validating it on the month of November 2015. This period is characterised by an indoor temperature relatively high for the season: around 22°C for the month of October and by a great fall on temperature in November going until 16°C. The outdoor temperature oscillates between 0 and 15°C (*cf.* 6b), which implies a heavy use of the heaters. The similar trends between the indoor air temperature and the outdoor temperature expresses the light inertia of the study case.

4.3 The reference model

Following the same approach as previously, the model with one capacity was tested first and then, the complexity of the models has been increased step by step. Figure 7 shows the temperature responses for the model with one capacity for both summer and winter scenarios. The red curve illustrates the first phase of
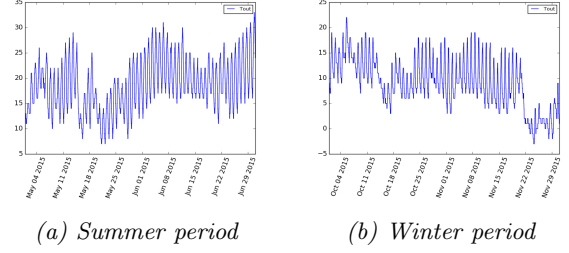

(a) Summer period	(b) Winter period

Figure 6: Outdoor temperature

validation of the model which consists in the simulation of the model on a dataset not used for estimating the parameters. The separation between the training and the cross validation phases is materialized by a colour change on the plots. It allows to perform a qualitative evaluation of the results of the model. Then, for a quantitative analysis of the estimation results, the standardized RMSE is computed (*cf.* table 2). The numerical values confirm the relevance and the accuracy of the model.
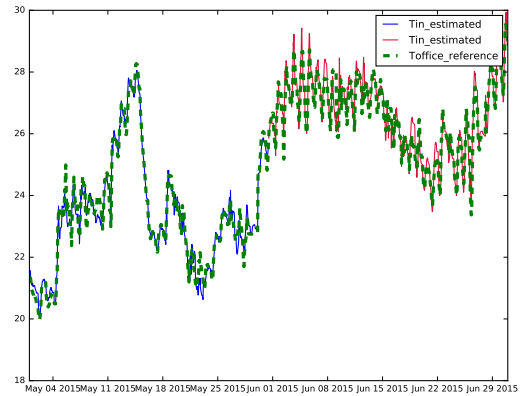
*(a) Summer Temperature*
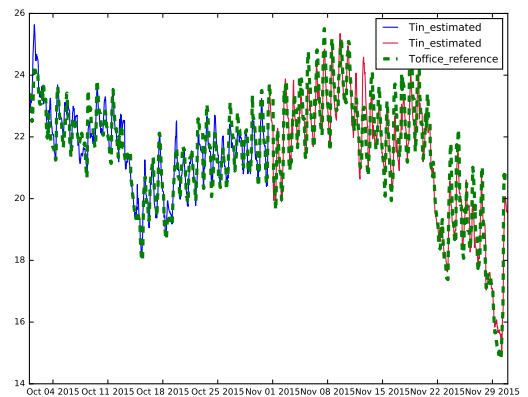
(b) Winter Temperature

*Figure 7: Results of the reference model*

Table 2: sRMSE values

	Estimation	Validation
Summer	0.13949	0.11774
Winter	0.14327	0.167634

4.4 Validation

According to the results of the sRMSE shown in figure 8, the "Model2Cair" and "Model3CT$_{out}$" can be eliminated because of the sRMSE value criteria and the ergodicity along the seasons. The models "Reference" and "Model0C" do not comply with the limit of sRMSE value.
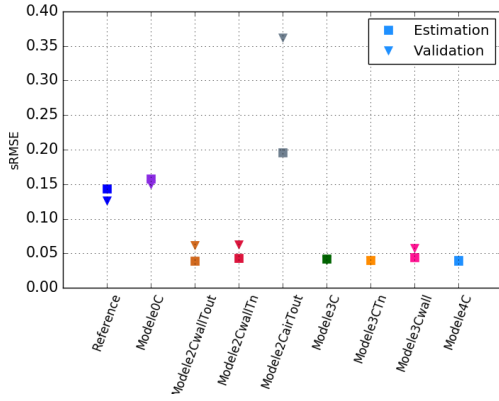
(a) Summer
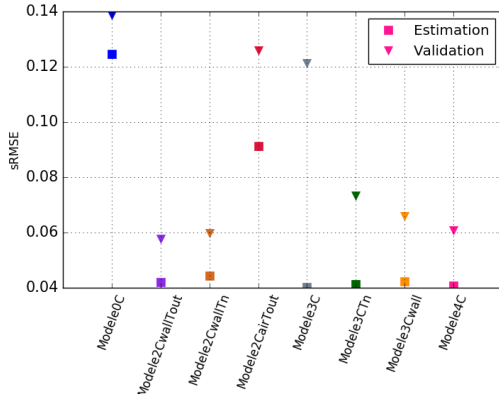
(b) Winter

Figure 8: Results of the different models in estimation and validation phase

4.5 Selection

The next step consists in selecting the most suitable model between "Model3CT$_n$", "Model2CwallT$_{out}$", "Model3Cwalls", "Model2CwallT$_n$" and "Model4C" which will be done by looking at the sum of the sRMSE values obtained for the different seasons. The result can be seen in the figure 9. It appeared directly

that the "Model4C" is the best one according to the criterion described above.
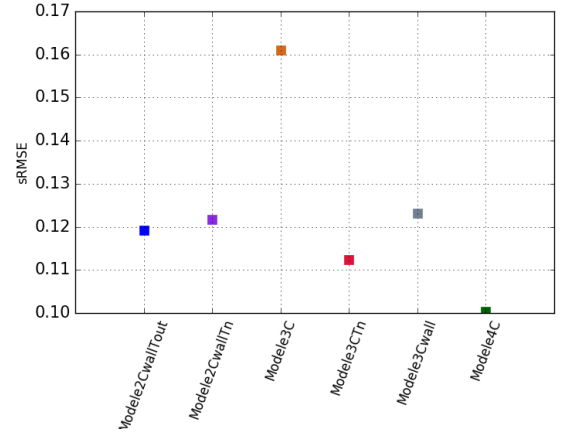
Figure 9: Sum of the sRMSE values for both seasons

4.6 Sensitivity analysis

The Morris analysis shown in figure 10 reveals that three main structural parameters of the model C$_{w,out}$, R$_{w,n2}$ and R$_{w,out2}$ seem not to have a negligible impact on the output. Yet, deleting one of the branch containing the parameters C$_{w,out}$ and R$_{w,out2}$ would mean considering one the structure with three capacities which performances (according to the criteria of accuracy) are less interesting. Hence, that result needs to be investigated further, especially studying the estimated values of the parameters. However, contrary to the other methods, the Morris method is qualitative which means that it is possible to classify the different parameters but not to ensure their relevance. As the goal of the sensitivity analysis in this research is to ensure that the structure of the model is consistent and that each parameter has an impact on the output, it is not suitable. Then, despite the computational cost, Sobol method will be further investigated in order to confirm these results.
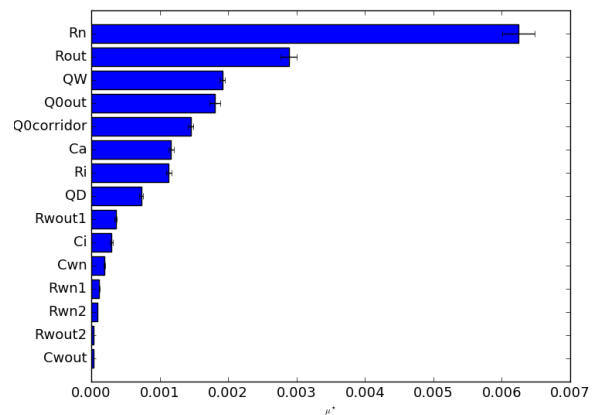

Figure 10: Morris analysis

4.7 Removal of sensors

In parallel, attempts have been made to quantify the impact of the loss of data on the accuracy of the results. For that purpose, the information of the door and window contacts used in the thermal model is no longer used in the estimation process. The results can be seen in figure 11. It can be noted that removing the contact data does not seem to have a significant impact in the sRMSE values and even tends to improve the results. The selected model can be applied to an apartment simply equipped with NetAtmo indoor and outdoor climate stations[1] which are sensors already marketed and adressing ordinary users. Results are described in the next section.
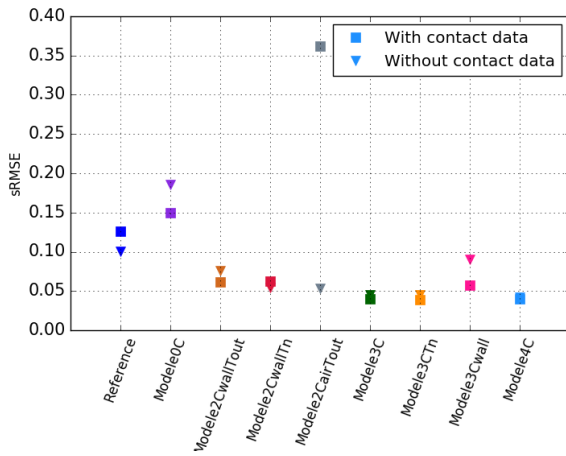
Figure 11: Comparison between with and without contact data

5 Application to a more complex case study

The first case study defined a single-zone model for a single room. The next step is to extend this model to a multi-zone case study. Besides, the multi-zone case study is a apartment, which brings a lot of differences in terms of uses and occupation rules.

5.1 Architecture and instrumentation

The architectural plan is shown in figure 12. The apartment is equipped with 5 NetAtmo stations, a main device in the Living Room which providing measurements of noise level, humidity, temperature, pressure and CO_2 concentration and three auxiliary devices, one in each room, delivering only measurements on humidity, temperature and CO_2 concentration. The bathroom device will not be used in a first time.

5.2 Modelling

In order to apply the model to a whole apartment, a combination of the mono-zone model has been implemented as illustrated in figure 13. Thereafter, the
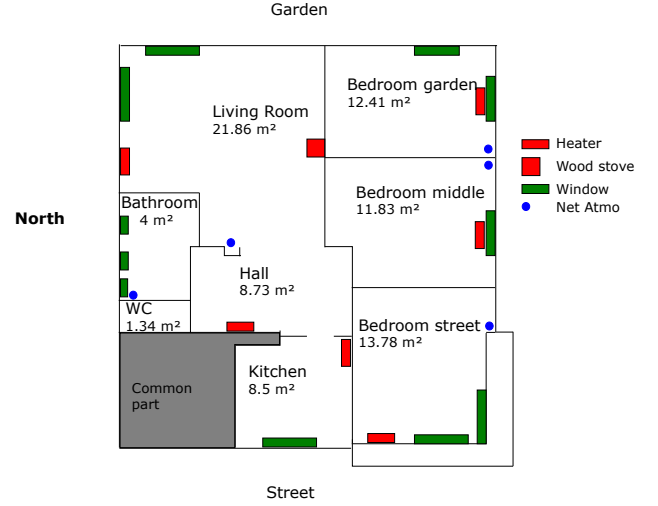
---

[1] http://www.netatmo.com

Figure 12: Architectural plan of the apartment

new type of living area requires to adapt all estimators of occupancy and heating since the uses are quite different. In a first time, it is done with really simple estimators of occupancy based on level of CO_2 in each room. Indeed, according to Jiang et al. (2016), CO2 concentration is the parameter that most closely correlates with the number of occupants.
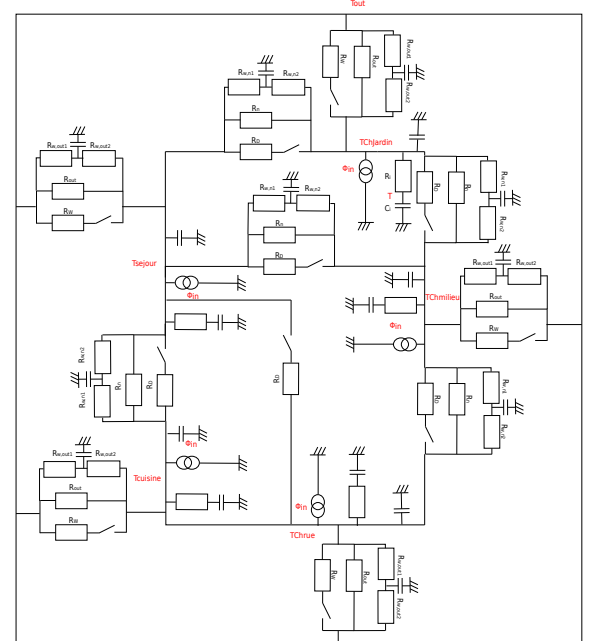

Figure 13: RC model of the apartment

5.3 Results

The "Model4C' was implemented for the whole apartment. From here, the temperature profiles shown in figure 15 are quite accurate: the trends are globally well represented but still need to be improved. As can be seen in figure 14, the CO_2 profiles are less accurate. The model seems to have issues representing the CO_2 peaks and while the global dynamic is well represented, it is not so good for the fast dynamics.
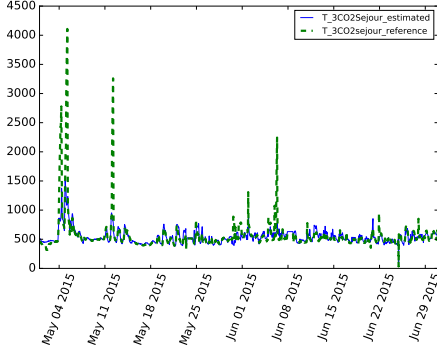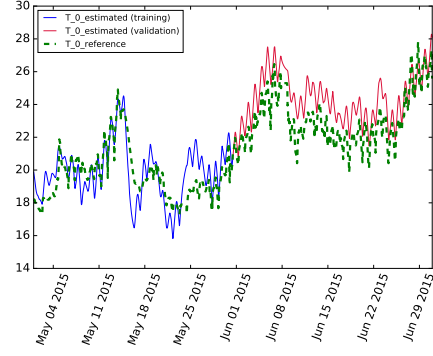
Figure 14: "Streets" bedroom CO_2

However, occupants are more sensitive to temperature accuracy than to CO_2 concentration. Problems of air quality occur when a threshold of 2000 ppm is reached, thus peak modelling should be improved accordingly.

6 Conclusion

This paper presented an innovative and easy way to model different types of buildings for energy management with very few data. This method has been applied to two case studies representative of the main types of living areas. The proposed models are grey-box models based on both data and prior physical knowledge. The configuration of those models can be done by anyone since no expert knowledge is required but only: area of rooms, orientation, architectural plan. In order to facilitate a further implementation, the data used have been reduced to a minimum: indoor temperature and CO_2 concentration in each room plus nebulosity and outdoor temperature forecasts. Those elements make the implementation of the energy management services accessible to a large public. Various helpful services allowing users to better understand their buildings can be imagined. The results presented in the multi zone case study are a promising first step. Further research will focus on ways to estimate more accurately the occupancy as well as the heating power in residential building. Further work will be led to understand the results of the Morris analysis by looking into the estimated parameters values and the cross-correlation of the parameters. The results of the models without contacts data need also to be investigated in order to either deleting that information of explaining the results. It must be understood also that deleting these data involves diminishing the intended services regarding the openings.
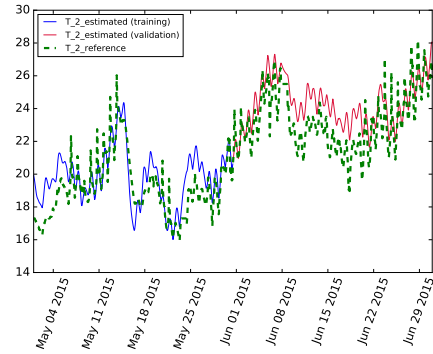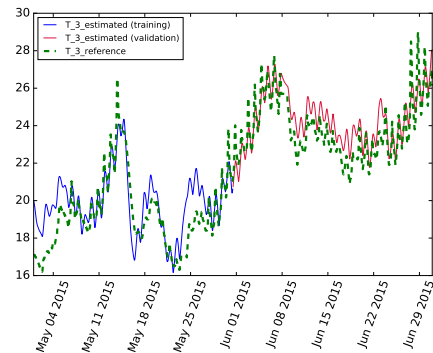
7 Acknowledgement

(a) Living Room
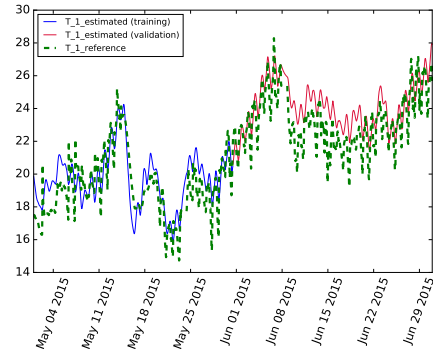
(b) "Garden" bedroom

(c) "Middle" bedroom

(d) "Streets" bedroom

Figure 15: Temperature profiles for the winter scenario

References

Bacher, P. and H. Madsen (2011). Identifying suitale models for the heat dynamics of buildings. *Energy and Buildings 43*, 1511–1522.

Carassus, J. (2011). Les immeubles de bureaux "verts" tiennent-ils leurs promesses? *Performances réelles, valeur immobilières et certification*.

Castillo, L., R. Enríquez, M. Jiménez, and M. Heras (2014). Dynamic integrated method based on regression and averages, applied to estimate the thermal parameters of a room in an occupied office building in madrid. *Energy and Buildings 81*, 337–362.

Deb, K., S. Agrawal, A. Pratap, and T. Meyarivan (2000). A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: Nsga-ii. In *International Conference on Parallel Problem Solving From Nature*, pp. 849–858. Springer.

Erickson, M., A. Mayer, and J. Horn (2001). The niched pareto genetic algorithm 2 applied to the design of groundwater remediation systems. In *International Conference on Evolutionary Multi-Criterion Optimization*, pp. 681–695. Springer.

Fabi, V., R. V. Andersen, S. Corgnati, and B. W. Olesen (2012). Occupants' window opening behaviour: A literature review of factors influencing occupant behaviour and models. *Building and Environment 58*, 188–198.

Hazyuk, I., C. Ghiaus, , and D. Penhouet (2012). Optimal temperature control of intermittently heated buildings using model predictive control. *Building and Environment Part I?Building Modeling*(51), 379?87.

Horn, J., N. Nafpliotis, and D. E. Goldberg (1994). A niched pareto genetic algorithm for multiobjective optimization. In *Evolutionary Computation, 1994. IEEE World Congress on Computational Intelligence., Proceedings of the First IEEE Conference on*, pp. 82–87. Ieee.

Jiang, C., M. K. Masood, Y. C. Soh, and H. Li (2016). Indoor occupancy estimation from carbon dioxide concentration. *Energy and Buildings 131*, 132–141.

Le Mounier, A. (2016). *Méta-optimisation pour la calibration automatique de modèles énergétiques bâtiment pour le pilotage anticipative*. Ph. D. thesis, Communauté Université Grenoble Alpes.

Mathews, E., P. Richards, and C. Lombard (1994). A first-order thermal model for building design. *Energy and Buildings* (21), 133–145.

Morris, M. D. (1991). Factorial sampling plans for preliminary computational experiments. *Technometrics 33*(2), 161–174.

Reynders, G. (2015). *Quantifying the Impact of Building Design on the Potential of Structural Storage for Active Demand Response in Residential Buildings*. Ph. D. thesis.

Saltelli, A., S. Tarantola, and K.-S. Chan (1999). A quantitative model-independent method for global sensitivity analysis of model output. *Technometrics 41*(1), 39–56.

Schaffer, J. D. (1985). Multiple objective optimization with vector evaluated genetic algorithms. In *Proceedings of the 1st international Conference on Genetic Algorithms*, pp. 93–100. L. Erlbaum Associates Inc.

Shaikh, P. H., N. B. M. Nor, P. Nallagownden, I. Elamvazuthi, and T. Ibrahim (2014). A review on optimized control systems for building energy and comfort management of smart sustainable buildings. *Renewable and Sustainable Energy Reviews 34*, 409–429.

Sidler, O. (2011). De la conception à la mesure, comment expliquer les écarts? *Evaluer les performances des bâtiments basse consommation, CSTB/CETE de l'OUEST*.

Sobol, I. M. (2001). Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates. *Mathematics and computers in simulation 55*(1), 271–280.
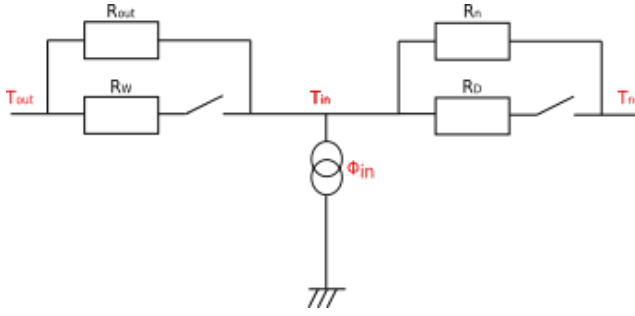
A   Annexe: Structures of models

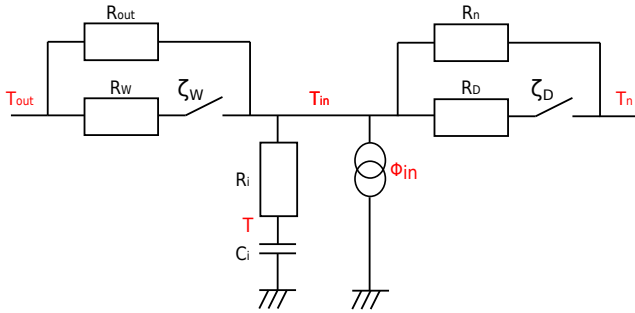
Figure 16: Model without capacitor: Model0C
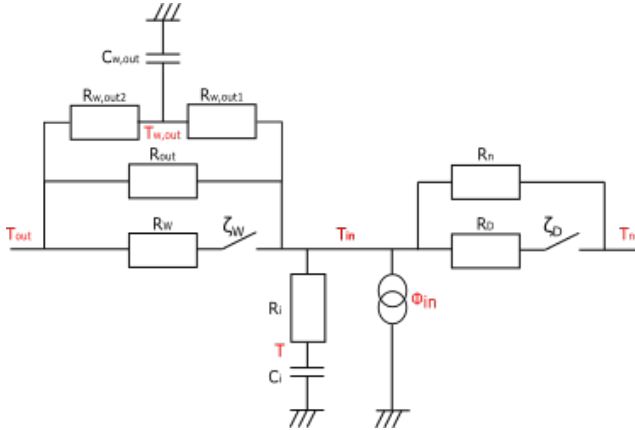

Figure 17: Reference model

Figure 18: Model with 2 capacitors: Model2CwallT$_{out}$
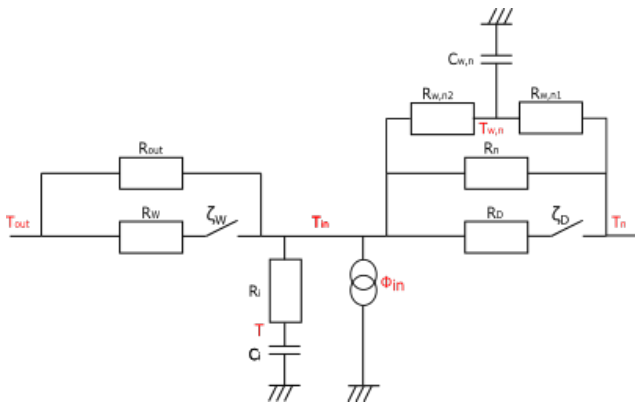

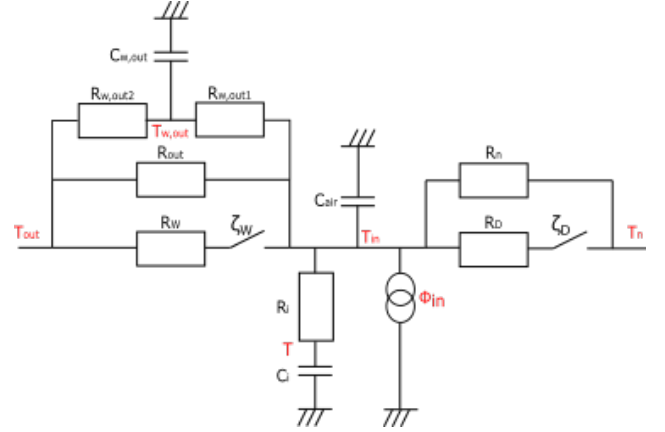*Figure 19: Model with 2 capacitors: Model2CwallT$_n$*

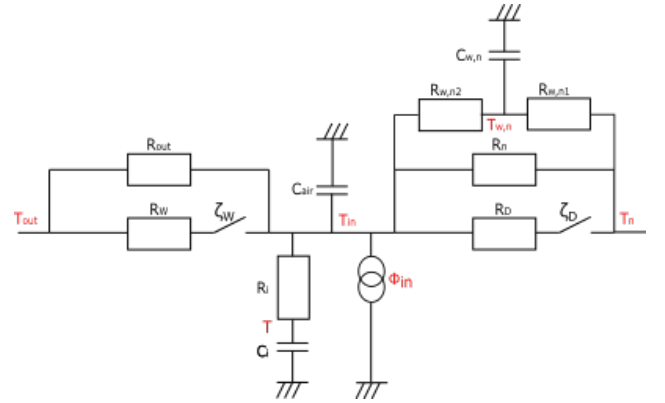*Figure 20: Model with 3 capacitors:Model3T$_{out}$*

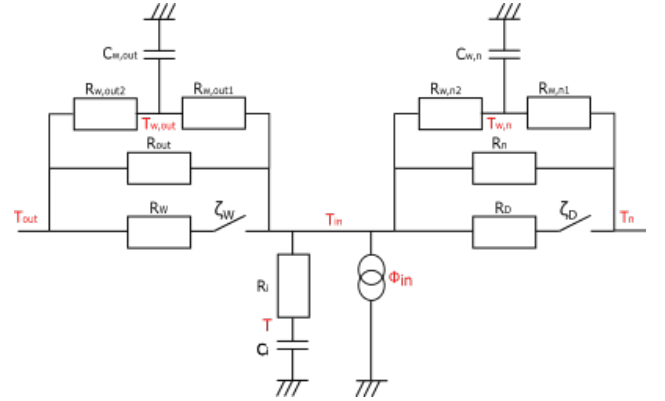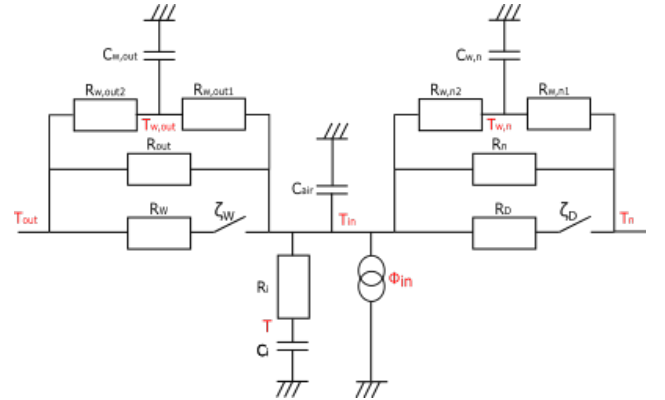*Figure 21: Model with 3 capacitors: Model3CT$_n$*

Figure 22: Model with 3 capacitors: Model3Cwalls

Figure 23: Model with 4 capacitors: Model4C