USING DESIGN OF EXPERIMENTS METHODS TO DEVELOP LOW ENERGY BUILDING MODEL UNDER MODELICA

Gilles PLESSIS, Sila FILFLI, Cristian MURESAN and Hassan BOUIA
EDF R&D EnerBat (Energy in buildings and territories department), Moret, France.
Contact information : gilles.plessis@edf.fr

ABSTRACT
Simplified energy simulation programs are increasingly used in the residential building sector to support design process. These tools are used to drive both new and existing buildings towards better energy performances. However, they have limited application ranges and are not adapted for low energy buildings or passive houses.

These simulation programs are mainly based on two technical appraisals. The first is derived from model order reduction methods applied to detailed models which describe the physical knowledge of the phenomena studied. The second is based on parametric studies. In this paper, combined methods of model order reduction and design of experiments are used on low energy building models.

INTRODUCTION
Purpose
Low energy buildings and passive houses are different from common buildings, the physical phenomena involved are strongly less intensive and therefore less perceptible. The whole concept of passivHaus means that the “free” heat gains must match the energy demand unlike in common building where it is mainly covered by the heating system. Nowadays, most of the design simulation tools, those using coarse or simplified model, are based on correlations which identify dynamic parameters of the model. These correlations are increasingly complex to find while building energy performances are improved. As mentioned in (Dufrestel et al., 2008), in order to take account of solar and internal heat gains, some changes have to be made. Usually, the dynamic building simulation for design process used one hour time step. This can be sufficient to account for outdoor temperature changes or when calculating only the yearly energy consumption. However, a finer time step must be considered for other solicitations (such as solar irradiation, air flow induced by the ventilation system, heating and cooling capacities, …) in order to predict more accurately energy demand.

This study aims to obtain simplified parametric models which are adapted to design process. The models should be significantly less time consuming than the original building models and be able to run on common computers. Besides, they should be able to account for much lower time steps than those commonly used in commercial software. Afterward, the reduced models generated (in state-space form) could be coupled with HVAC systems and control devices in order to simulate the global energy behavior. This new model generation will be appropriate to design the building and its HVAC systems as a whole system.

Methodology
The study presented in this paper is carried out in two main steps. The first one consists of a model order reduction which includes a linearization and a reduction of the detailed model. The second stage aims at finding multi-polynomial expressions of the reduced model coefficients thanks to design of experiments methods.

The reduction method selected is the balancing realization method described by (Moore, 1981). The Moore method is a direct method allowing to set the order of the reduced model related directly to the simulation time step. A preliminary parametric analysis on a one-zone model of a residential building shows that a 3th order model is sufficiently accurate to predict dynamics with respect to 15 minutes time step. The most sensitive transfer functions are related to the air supply temperature and internal heat gains.

The parametric analysis is used to find the multi-polynomial expressions of the reduced model coefficients. Design of experiments methods allow to organise the simulations needed to obtain parametric models and to reduce significantly their number thanks to fractional and D-optimal designs. The precision of D-optimal models are compared to full factorial design models. This method helps to detect easily the effects and interactions of the most influent parameters.

ORIGINAL BUILDING MODEL
The building model is based on a typology of French buildings and is called Mozart (Riederer and Partenay, 2010) (Noël, 2008). This is a one-zone model with about 100 m² floor representing an individual house. The detailed building model is
implemented thanks to a Modelica® library called OSMOSYS (Open Standard for MOdeling SYStem) developed by EDF. This library is based on acausal approach which eases the way a building envelope is created from components (window, wall, door, thermal bridge, air volume...) and likewise for HVAC systems. The model was validated experimentally at EDF. The building model links the different inputs to the observed output, the indoor air temperature, in a causal way. The outputs and the input of the building model can be seen on the Figure 1.

![Mozart model](image)

**Figure 1 Inputs and Output of the building model**

Depending mainly on the physical phenomena considered and the spatial discretization used, the order of the detailed system can be high. Accordingly, the simulation can be time-consuming especially for a 1-year simulation and not suitable for commercial softwares used during design process.

The physical phenomena considered in this model are linear:
- 1D conduction through building envelope
- Convection between walls and air volume or outdoor air
- Linearized radiation accounting for long wave radiation
- Net heat flows accounting for internal and solar heat fluxes (short wave radiation).

After a linearization, the model based on a set of differential algebraic equations is transformed into a causal state-space linear time-invariant system (Gao et al., 2008):

\[
\begin{align*}
\dot{T} &= AT + BU \\
T_s &= CT + DU
\end{align*}
\]

(1)

Where
- \( T \) is the state vector representing the temperature nodes
- \( U \) is the input vector representing the different solicitations (inputs) such as solar and internal heat gains
- \( T_s \) the output vector, which is in this case the indoor air temperature
- \( A \) is the state matrix
- \( B \) is the command matrix
- \( C \) is the observation matrix
- \( D \) is the matrix of direct transmission

The equation linking the net solar heat fluxes to the indoor air temperature is identical with respect to orientation. “Identical input” means that the zeros and poles of their transfer function are the same, however the static gain is not. Therefore the 15 inputs can be reduced to 6:
- Transmitted heat flux through windows
- Absorbed heat flux in windows
- Absorbed heat flux in walls
- Supply air temperature
- Outdoor air temperature
- Internal heat gains

Moreover, the inputs apply without thermal inertia to the indoor air node are identical. Consequently, the absorbed and transmitted solar flux by the windows, the supply air temperature and the internal heat gains are identical inputs. Finally, only 3 inputs can be considered:
- Absorbed solar flux by the walls
- Internal heat gains
- Outdoor air temperature

As previously mentioned, the order of the detailed model can be huge and thus inducing important CPU resources. As to overcome this issue, a model order reduction is performed.

**MODEL ORDER REDUCTION**

As previously mentioned, the Moore method is used with static gain conservation to perform model order reduction. This method was compared to other model order reduction methods, such as modal reduction (Marshall and Michailescos methods), aggregation (Roux, 1984) and was found more effective in previous papers (Déqué, 1997) (Palomo et Al., 1997). This method aims at finding a balanced basis depending on the controllability and observability concept (Moore, 1981). Each state can be described by its degree of controllability and observability which can be deduced from the gramians, noted respectively \( W_c \) and \( W_o \), solution of the Lyapunov equation. The whole method is based on state truncation which has either low controllability or observability degree. However, the state variables may not have the same degree of controllability and observability, the degree of controllability can be important and the observability can be very low or contrariwise. Therefore a balanced realization is computed on the system described in equation (1) in order to define new states with balanced degree of controllability and observability. The system after state coordinate transformation is:
The new observability and controllability gramians, $\bar{W}_o$ and $\bar{W}_c$, are equal. Depending on the reduction order, $N_f$, a states truncation is performed on the system described equation (2) so as to find the reduced order model:

$$
\begin{align*}
\dot{\tilde{X}}_{\text{red}} &= A_{\text{red}} \cdot X_{\text{red}} + B_{\text{red}} \cdot U \\
T_s &= C_{\text{red}} \cdot X_{\text{red}} + D_{\text{red}} \cdot U
\end{align*}
$$

(3)

Then the reduced model in the state-space form (equation (3)) is transformed into rational transfer functions in the frequency domain:

$$
TF_j(s) = K_j \prod_{i=1}^{i=N_f} \frac{s - z_{ij}}{s - p_{ij}} \quad j \in \{1, \ldots, N_{\text{in}}\}
$$

(4)

Where:
- $s$ the complex variable
- $TF_j$ the transfer function related to the input $j$
- $K_j$ the static gain
- $z_{ij}$ the zeros of the transfer function, which are either real or pair of complex conjugate, $z_{ij} \in \mathbb{C}$
- $p_{ij}$ the poles of the transfer function which are real, $p_{ij} \in \mathbb{R}$
- $N_{\text{in}}$ being the number of inputs.

Finally the transfer functions presented equation (4) are transformed into a simpler form:

$$
TF_j(s) = K_j \cdot \sum_{i=1}^{i=N_f} \frac{k_{ij}}{s - p_{ij}} \quad j \in \{1, \ldots, N_{\text{in}}\}
$$

(5)

Where $k_{ij} \in \mathbb{R}$.

The parametric transfer functions used to model the building take the form of equation (5).

The following step in the methodology is based on design of experiments techniques. The $K_{ij}$, $k_{ij}$ and $p_{ij}$ parameters of each transfer function are correlated to the different technological criteria (insulation level, heat transfer coefficients, air change rate, ...).

**DESIGN OF EXPERIMENTS**

Design of experiments methods aim to study the impact of different parameters, called the experimental factors, on a response variable. These methods enable one to understand the response variation as a function of experimental factors.

These factors, $x_i$, are considered as potential causes of the response variations, $y$.

$$
\hat{y} = f(x_i) \quad i \in \{1, \ldots, N_{\text{exp}}\}
$$

(6)

Where $N_{\text{exp}}$ is the number of factors considered and $\hat{y}$, the regression model of the response $y$.

The quality of the correlation is due to 3 main features:
- The regression expression
- The factors considered
- The factors ranges and therefore their levels

**Levels and factors selection**

Design of experiments requires a good knowledge of the phenomenon studied in order to consider the most significant factors. The domain and levels must be selected carefully. If the number of levels is limited over a large domain, the correlation can be very complicated to find or will not be very accurate. The levels chosen for this study encompass from very low energy building (low level) to old building with poor energy performances (high level). The design of experiments and correlation expression obtained are only valid inside this domain.

**Table 1**

<table>
<thead>
<tr>
<th>FACTORS</th>
<th>UNIT</th>
<th>LOW LEVEL</th>
<th>MEDIUM LEVEL</th>
<th>HIGH LEVEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>$U_{\text{win}}$</td>
<td>[W.m$^{-2}$.K$^{-1}$]</td>
<td>1.00</td>
<td>2.75</td>
<td>4.50</td>
</tr>
<tr>
<td>$v$</td>
<td>[h$^{-1}$]</td>
<td>0.40</td>
<td>1.20</td>
<td>2.00</td>
</tr>
<tr>
<td>$U_{\text{wall}}$</td>
<td>[W.m$^{-2}$.K$^{-1}$]</td>
<td>0.150</td>
<td>0.875</td>
<td>1.600</td>
</tr>
<tr>
<td>$U_{\text{roof}}$</td>
<td>[W.m$^{-2}$.K$^{-1}$]</td>
<td>0.10</td>
<td>0.70</td>
<td>1.30</td>
</tr>
</tbody>
</table>

Only the main levels are presented in the Table 1, however 2 intermediate levels are added equidistantly between the low and medium level from one side and between the medium and the high level on the other side.

For sake of convenience, only the main factors are presented here:
- $U_{\text{win}}$ the thermal conductance of the windows
- $v$ the air change rate of the building
- $U_{\text{wall}}$ the thermal conductance of the walls
- $U_{\text{roof}}$ the thermal conductance of the roof

These factors are chosen not only because they are significant but also because they can be changed during the design process. It should be noted that the thermal inertia of the building envelope was not
chosen as an experimental factor in this paper for sake of convenience albeit it is a significant one. However, as this project aims to generate accurate building models, all the significant factors should be considered.

The full factorial design presented in Table 1 requires $5^5 = 625$ experiments. Other design of experiments, such as fractional and D-optimal exist and are appropriate when the number of factors and levels is considerable (Filfili, 2006) (Goupy, 1999). They allow to decrease significantly the number of experiments with only a slight loss of accuracy. Thus a D-optimal design composed of only 15 experiments (3 main levels considered) is used for comparison.

The response variables studied are the poles and gain of the transfer functions presented in equation (5), $K_j$, $k_j$, and $p_{ij}$. Considering that the number of solicitations is 3, and the reduced order is 3, the number of responses to study is 21. In order to get a whole model representing the effects of 15 solicitations on the indoor air temperature, the other solicitations are studied to get only the $K_j$ values.

However in this paper, only the 3 solicitations previously mentioned are investigated.

The correlation used in this study is a polynomial type which is the most elaborate when considering 3 levels. The regression model for the full factorial design is defined as:

$$
\hat{y} = \sum_{i,j,k,l,m} a_i \cdot U_{x1}^1 \cdot U_{x2}^1 \cdot U_{x3}^1 \cdot U_{x4}^1 \cdot U_{x5}^1
$$

(7)

The number of $a$ coefficients appearing in equation (7) is 32. Considering the fact that 625 experiments are carried out by the full factorial design, the least square method is used to find the coefficients of the polynomial model, $a_i$.

In the D-optimal design, only the first order interactions are considered, $x_i \cdot x_j$.

$$
\hat{y} = \sum_{i,j,k,l,m} a_i \cdot U_{x1}^1 \cdot U_{x2}^1 \cdot U_{x3}^1 \cdot U_{x4}^1 \cdot U_{x5}^1
$$

(8)

As the number of coefficients match the number of experiments, 15, their deduction is straightforward. Finally the methodology described in this section provides 7 polynomial expressions ($K_j$, $3k_j$ and $3p_{ij}$) per solicitation (3 different solicitations).

Statistical indicators such as the relative mean error and its standard deviation are used to qualify the regression. A relative mean error near zero signifies that the regression is well calibrated. A low standard deviation means that the regression is precise. The relative mean error is defined for each experiment as:

$$
e = 100 \cdot (\hat{y} - y) / y
$$

(9)

Figure 2 presents the comparison between the response $p_{11}$ of both polynomial regressions. The vertical axis and horizontal axis represent respectively the true values from the model order reduction and the interpolated values from design of experiments methods.

The full factorial design is slightly better than D-optimal design, however the difference is not significant. The mean error is around 0 % for both designs and the standard deviation is 1 % for the D-optimal and nearly 0 % for the full factorial design.

Figure 3 presents the response $K_5$ of both polynomial regressions.

The two compared designs are well calibrated with a relative mean error value of 0 %. Moreover, in this case, the D-optimal design is again slightly less precise (standard mean deviation of 3 % against 2 %). These results are representative of other comparisons carried out.
DYNAMIC SIMULATIONS

As to ensure that the whole methodology is suitable to model low energy buildings some dynamic simulations are performed. The 3 solicitations selected previously are investigated:

- Absorbed solar flux by the walls
- Internal heat gains
- Outdoor air temperature

As mentioned in the introduction, this study aims at finding a parametric reduced model available to model low energy buildings and therefore appropriate for simulations with low time step. Each input excitation is a sinus wave with a bias. In this case, it is chosen to set the excitation time period depending on the solicitation. For solar heat flux absorbed by the walls and internal heat gains, a time period of 15 minutes is chosen. For the solicitation representing the outdoor air temperature a time period of 24 hours is chosen. These time periods are identified as sufficient especially for solar heat flux and internal heat gains which are the most dynamic solicitations.

The following figures present some comparisons between the results obtained with the detailed model and results obtained thanks to a D-optimal polynomial expression. Only one building is presented in this section. This building has the following specifications:

<table>
<thead>
<tr>
<th>FACTORS</th>
<th>UNIT</th>
<th>LEVEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>$U_{\text{win}}$</td>
<td>W.m$^{-2}$.K$^{-1}$</td>
<td>1.00</td>
</tr>
<tr>
<td>$v$</td>
<td>h$^{-1}$</td>
<td>1.20</td>
</tr>
<tr>
<td>$U_{\text{wall}}$</td>
<td>W.m$^{-2}$.K$^{-1}$</td>
<td>0.150</td>
</tr>
<tr>
<td>$U_{\text{roof}}$</td>
<td>W.m$^{-2}$.K$^{-1}$</td>
<td>0.10</td>
</tr>
</tbody>
</table>

These specifications represent those which can be used to design a very low energy building (Effinergie, 2008). The building investigated has particularly very low thermal conductances due to high insulation level.

Solar flux on the walls

The solicitation considered for the net solar heat flux on the wall is a sinusoidal heat flux of 300 W/m² amplitude and a bias of 400 W/m². Equation 10 gives the function of this solicitation:

$$I_{\text{SHF}}(t) = 300 \cdot \sin\left(\frac{2 \cdot \pi \cdot t}{\tau}\right) + 400$$

$$\tau = 900 \, \text{s}$$

The simulation is performed over a week. All the other solicitations are set to zeros, especially the outdoor air temperature and the supply air temperature, this is why the indoor air temperature elevation is rather small, about 0.1°C.

Internal heat gains

The solicitation considered for the internal heat gains is a sinusoidal heat flux of 100 W amplitude with a bias of 600 W for the whole building.

Equation 11 gives the function of this solicitation:

$$I_{\text{INT}}(t) = 100 \cdot \sin\left(\frac{2 \cdot \pi \cdot t}{\tau}\right) + 600$$

$$\tau = 900 \, \text{s}$$

All the other solicitations are also set to zero. The indoor air temperature elevation is obviously much important than the one resulting of the solar heat flux solicitation.

Outdoor air temperature

The solicitation considered for the outdoor air temperature is a sinusoidal temperature treament of 10°C amplitude with a bias of 5°C. Equation 12 gives the function of this solicitation:
\[ I_{OAT}(t) = 10 \cdot \sin\left(\frac{2 \cdot \pi}{\tau} \cdot t\right) + 5 \] \hspace{1cm} (12)

\[ \tau = 86400 \text{ s} \]

The simulation is performed over a week. All the other solicitations are set to zero. In this case, a time period of 24 hours seems more appropriate than 15 minutes.

DISCUSSION AND RESULT ANALYSIS
This part presents both discussion and result analysis for the 2 previous sections, design of experiments and dynamic simulations.

Design of experiments
In both Figure 2 and Figure 3, the D-optimal design is slightly less precise than the full factorial design (1 % against 0 % and 3 % against 2 %). However the 2 designs seems well calibrated, they both give a mean error value of 0 %. The number of experiments required for the D-optimal design is tremendously decreased compared to the full factorial design, respectively 15 against 625. This could be of great interest when increasing the number of factors and levels considered.

These preliminary results can not be sufficient to generalize that the parametric models will fit the reduced models. Indeed, these statistical analysis give aggregated indicators, and a more thorough analysis is required to validate the dynamic behaviour of such models.

Dynamic simulations
Dynamic simulations aim to validate the methodology and reveal issues which may arise. It is chosen to select a very low energy building, actually this model is close to the domain boundaries for the factors studied.

From Figure 4 and Figure 5, one can say that D-optimal design provides an accurate model, at least for the solar heat flux and the internal heat gains. The difference between the detailed model and D-optimal model is always under 0.5 °C and the dynamics evolution of temperature is well captured. These results can be better by improving the polynomial expression defining the static gain of the internal heat gains solicitation.

However, on Figure 6 presenting the outdoor air temperature solicitation, the D-optimal design is strongly incorrect. On the contrary, on Figure 7, the polynomial expressions obtained with the full factorial design enable to get an accurate model of the detailed building. The difference between these two, detailed and full factorial model, is always under 0.3°C during a one week simulation. The exact reduced model is also presented on Figure 7. It is very close from the detailed model both in term of dynamic and absolute temperature values.

Perspectives
Some additional investigations were performed in order to understand the issue illustrated by Figure 6. It seems that for some experiments, the real part of one or two zeros of the reduced transfer functions presented equation (4) may encounter a discontinuity.
and change their sign. The zeros can modify the dynamic behavior of the system. In this case, the full factorial design thanks to the least square method, smooths the coefficients variations and is able to fit the detailed model accurately. When a slow zero has positive real part, an undershoot (such as the behavior expose on the Figure 6 when the response goes in the opposite direction) is obtained. No physical meaning of the undershoot was found. Additional investigations revealed that a weakness of the reduction method causes the undershoot. The Lyapunov based balanced truncation provides accurate results only if it is applied to stable open-loop systems. More evaluated reduction methods based also on balanced truncation will be explored in order to overcome this issue. Jonckheere and Silverman (Jonckheere, 1983) introduced the LQG balanced truncation method for unstable closed-loop systems. The main advantage of this method is to first close up the system with a positive feedback by means of an optimal LQG compensator. After algebraic manipulations the contribution of the each state component to the closed-loop system is obtained and the redundant state components can be deleted.

The reduction of the detailed model requires the preservation of certain properties of the model such as passivity and stability. Passivity is an important feature of linear systems. In network theory, a passive component consumes but does not produce energy (thermodynamic passivity). The Lyapunov based balanced truncation does not guarantee, in general, the passivity preservation of the reduced-order model. Appropriate reduction methods (Opdenacker, 1988) based on the generalization of algebraic Riccati equations allow to preserve passivity.

CONCLUSION

The methodology proposed in this paper allows to generate parametric reduced models appropriate for low time step simulation. These models are generated over a large domain of building specifications, from very low to high energy performances.

The model being reduced, the computer run time is highly decreased. In order to evaluate the computer run time gain, 2 simulation runs over a year were performed with a common computer (AMD Athlon 64 X2 Dual-Core processor 5000B, 2.6 Ghz, 2 GB RAM) and the DASSL solver of Dymola. The Mozart building model previously described was used with a weather file of Moret in France. The computer run time for the detailed building was 30 seconds and for the reduced model, 2 seconds.

The results presented in this paper are promising for both parts, design of experiments and dynamic simulations. The correlation expressions obtained from experimental designs are accurate, with a mean error of 0 % and a standard deviation around 2 %. Concerning the dynamic simulations, the temperature differences are below 0.5°C over a one week simulation except for the model obtained thanks to D-optimal design and solicited with outdoor air temperature.

In order to overcome the observed discrepancies, future work will inspect more accurate model order reduction techniques for linear time invariant system. This work will be coupled with the error estimation of the combined approach (model order reduction and design of experiments).

REFERENCES


