Linking design and operation performance analysis through model calibration: Parametric assessment on a Passive House building

Published Version:
Lamberto Tronchin, M.M. (2018). Linking design and operation performance analysis through model calibration: Parametric assessment on a Passive House building. ENERGY, 165(A), 26-40 [10.1016/j.energy.2018.09.037].

Availability:
This version is available at: https://hdl.handle.net/11585/643314 since: 2018-09-20

Published:
DOI: http://doi.org/10.1016/j.energy.2018.09.037

Linking design and operation performance analysis through model calibration: Parametric assessment on a Passive House building

Lamberto Tronchin[*a], Massimiliano Manfren[b], Patrick AB James[b]
[a]Department of Architecture, University of Bologna, Via Cavalcavia 61 47521, Cesena, Italy
[b]Faculty of Engineering and the Environment, University of Southampton, Highfield, Southampton SO17 1BJ, United Kingdom

Abstract
Efficient buildings are an essential component of sustainability and energy transitions, which represent today a techno-economic and socio-economic problem. New paradigms are emerging both for new and existing buildings (e.g. NZEBs) and passive design strategies are becoming increasingly common. However, the adoption of these strategies in mild climates has to be carefully evaluated to prevent overheating in intermediate seasons and increasing cooling loads in summer, considering also climate change scenarios. Additionally, optimistic assumptions about building technology performance are often considered and the variability of occupant comfort preferences and behaviour is generally neglected in the design phase. The research presented aims at verifying the suitability of a simple, robust and scalable calibration approach (based on multivariate linear regression) to link design and operational performance analysis transparently, using a Passive House case study building. First, the original baseline design configuration is compared with a larger spectrum of data generated by means of parametric simulation, following a Design of Experiment (DOE) approach. After that, regression models are trained first on simulation data and then progressively calibrated on measured data during a three year monitoring period. The two fundamental objectives are evaluating the robustness of design phase performance analysis through parametric simulation (i.e. detecting potentially critical assumptions) and maintaining a continuity with operation phase performance analysis (i.e. exploiting the feed-back from measured data).

Keywords: Parametric modelling; behavioural modelling; building performance simulation; Passive House; performance monitoring; multivariate regression.

Highlights:

- Buildings are a relevant element in sustainability transition policies.
- Rigorous schemes for energy efficiency are important tools for designers.
- Robustness of performance estimates has to be considered in design phase.
- Design and operational performance analysis have to be linked transparently.
- Automated model calibration is necessary to ensure long-term performance monitoring.

Table of contents

1 Introduction

Efficient buildings are an essential component of sustainability and energy transition policies today and represent a techno-economic and socio-economic problem. The decarbonisation of building stock is one of the most important goals of policies, considering the impact of buildings at the global scale [1] and, in particular, in highly developed countries [2]. Building stock decarbonisation process embodies the necessity of increasing energy efficiency in end-uses, reducing demand and providing a relevant quota of energy supply by renewable sources. Energy efficiency paradigms are emerging both for new and existing buildings (i.e. Nearly Zero Energy Buildings, or NZEBs) [3] and passive design strategies, exploiting solar and internal gains to balance heat losses due to transmission and ventilation (in heating mode), are becoming increasingly common. These strategies can be particularly effective where heating constitutes, in most of the cases, the predominant part of energy consumption. However, the adoption of these strategies in mild climates has to be carefully evaluated to prevent overheating [4, 5] in intermediate seasons and increasing cooling loads in summer, considering also climate change problem [6], as buildings are long-term assets.

More in general, despite the great research effort put in design tools and technical standards in the last decades, both "re-bound" and "pre-bound" effect have been found empirically and, therefore, the gap between simulated and measured performance has been widely investigated in recent years [7, 8]. The "re-bound" effect [9] in efficient buildings is determined by inappropriate operation strategies, while the "pre-bound" effect [10] in inefficient buildings is determined by a more conscious consideration of the costs of energy services by occupants. Consequently, we have to acknowledge the fact that design phase assumptions and calculation methodologies can highly impact the reliability of our estimates of building performance, considering the essential problem of matching simulated and measured performance [11, 12] through calibration techniques. Additionally, in most of the cases the variability of the impact of occupants' comfort preferences and behaviour on performance is generally neglected in the design phase [13-15]. Finally, we can identify also an increasing commitment towards resource efficiency [16] in the built environment and the need for a holistic view on the topic of building sustainability [17], considering the whole life cycle impact of technologies for the building sector in a more realistic and reliable way [18-20]. All these elements constitute the motivation for the research presented.

As anticipated, model calibration is essential to link design and operational performance analysis under uncertainty [8] and the research is based on two fundamental tools: parametric simulation to produce a large spectrum of possible building energy performance outcomes (considering realistically the impact of the user behaviour and variable operating conditions from the very beginning), and model calibration employing a simple, robust and scalable technique (i.e. multivariate linear regression). A Passive House building is employed as case study to illustrate our approach. First, the original baseline design configuration is compared with a larger spectrum of data generated by means of parametric simulation, following a Design of Experiment (DOE) approach. After that, regression models are trained first on simulation data and then progressively calibrated during a three year monitoring period. In synthesis, the two fundamental research objectives are increasing the robustness of performance estimates in design phase, through parametric simulation, and maintaining, at the same time, a continuity with operational phase performance analysis, through model calibration. In this way, it is possible to detect first critical assumptions already in the design phase and then to derive critical insights as a feed-back from measured data, during operation phase. The techniques used are chosen because of their simplicity, robustness and scalability. The latter is particularly important as shown in recent research on knowledge discovery in large scale building stock datasets [21, 22] and on Model Predictive Control for the integration of renewables in the built environment [23]. For these reasons, the chosen approach is potentially suitable for both individual buildings, which can have a minimal cost automated performance monitoring (to keep performance under control at a reasonable effort, in long-term monitoring), but also for large scale studies [24-26] aimed at energy planning and policy, using inexpensive data acquisition and processing procedures.

Nomenclature

Variables and parameters

A	average value
a,b,c,d,e,f	regression coefficients
Cv(RMSE)	coefficient of variation of RMSE
D	deviation, difference between measured and simulated data
I	solar radiation
M	measured/simulated data
MAPE	mean absolute percentage error
NMBE	normalized mean bias error
q	specific energy transfer rate (energy signature)
P	predicted data
R^2	determination coefficient
RD	relative deviation
RMSE	root mean square error
S	simulated
SS	sum of the squares
y	numeric value
θ	temperature
ε	error term

Subscripts and superscripts

–	average
^	predicted value
b	baseline
c	cooling
h	heating
i	index
n	number of points
res	residual

2 Research methodology

The importance of parametric and probabilistic analysis of building performance is becoming evident [27-30], both in new construction and retrofit interventions [31, 32]. Cost-optimal [33] levels of investment have to be considered for the effective deployment of energy efficiency practices and, consequently, for the credibility and success of policies in this direction. However, occupants’ comfort preferences and behaviour [14, 15, 34] can lead to a relevant gap between simulated and measured performance [7], undermining the effectiveness of policies that have to confront with real behaviour [8, 9, 35].

In order to overcome this fundamental issue, a methodological continuity should be established between performance analysis practices across life cycle phases (i.e. model based analysis), using parametric simulation in design phase (generally only a limited amount of parameter configurations is considered for design phase simulations) and progressively calibrating building models to measured data (lo learn from feed-back). A great effort has been put in recent years on optimization [36] and simulation-based optimization [37] of building energy performance. Further, Design of Experiments (DOE) and parametric design have received also an increasing attention [27-30], together with Monte Carlo simulation to test the robustness of performance modelling [15, 28, 38].

Meta-models [39] (i.e. surrogate models, reduced-order models) are considered among the most promising techniques to overcome the limitations determined by the dimension of the optimization problems or parametric simulations. The choice of a specific technique can dependent on several factors [40]. Indeed, meta-models can be successfully used for different purposes, e.g. in design optimization, [37] calibration [39] and control [41]. In fact, they are very flexible and they can be employed to link design and operation phase performance analysis [42], considering, however, the trade-offs between complexity, predictive ability and transparency (i.e. black-box Vs grey-box models) [40]. In this research we propose piecewise linear multivariate regression models for calibration. This choice is motivated in detail in Section 2.1, considering both design and operational phase issues.

2.1 Motivations for regression modelling approach

Building performance can be studied by means of Key Performance Indicators (KPIs) [43-45], generally aimed at aggregating a larger set of data in a single representative quantity. Clearly, KPIs can be used to characterize both design and operational performance. This section presents the motivations for using a regression-based approach in this sense.

As anticipated, meta-models are flexible techniques which can be used for multiple purposed during building life cycle phases. With respect design phase issues, we can find in recent literature several examples of multi-variate regression models to support design optimization [46-50], considering also topics such as robustness of energy performance contracting and cost-optimal analysis [38, 51]. Further, with respect to operation phase issues, models are acceptable for calibration if they are able to satisfy the thresholds of measurement and verification (M&V) protocols [52-54], which constitute the minimal requirements. The motivations for the choice of a regression modelling approach in this research are connected to previous research conducted in the field and future prospects, considering relevant topics such as:

1. conceptual simplicity and ease of implementation compared to other meta-model based techniques for calibration [39];
2. automated or partially automated model selection capabilities [55, 56];
3. possibility to account for the impact of different operational strategies and conditions [13-15], considering different levels of thermal inertia [57];
4. scalability and applicability with respect to different types of end-uses [58] and multiple temporal [59, 60] and spatial scales [24, 26];
5. visualization of the impact of users' behaviour [14];
6. model robustness testing, under different behavioural conditions, using Monte Carlo simulation [15];
7. use of Bayesian analysis [61, 62] as an extension of conventional regression;

Finally, the use of simplified but robust and scalable models could potentially open up new perspectives for the application of large scale optimization of distributed energy resource in the built environment [23, 63-68], considering the problem of updating model parameters through periodic recalibration in evolving conditions [6, 69]. In order to render these applications more transparent and automated, further research should be oriented towards the definition of multi-scale and multi-level performance metrics [58, 70] and corresponding visualization techniques.

2.2 Methodology for case study analysis

The research presented is based on a case study analysis. In Section 3.1, the data from the original building design are used as baseline (initial design simulation) and then compared to parametric simulation runs obtained using Design of Experiment (DOE) approach. Therefore, parameters in DOE simulations have been varied with respect to the baseline configuration. Initial design involved the use of PHPP semi-stationary calculation methodology [71], specifically developed for Passive House buildings. In this research simulations are conducted using a validated grey-box dynamic model, suitable to perform multiple runs in a reduced time frame [72, 73], mantaining, at the same time, an acceptable level of reliability. Further, this choice corresponds to the necessity of enabling a future development of the research oriented to the non-intrusive identification of relevant physical parameters of the building [74]. In this research the grey-box lumped model parameters have been initially calibrated to the original baseline configuration in PHPP, to ensure comparability of results, and then varied following two-level full factorial design experiment plans [75], to compute every possible combination of factors and levels. Generally, a full factorial DOE cannot be used because of the computational effort: due to the exponential growth of experiments' number, this is only feasible for a limited number of factors and levels (as in this case study). The alternative choice would be running different fractional designs, where a selection of factor combinations is identified to reduce the number of experiments while maintaining an appropriate exploration of the design space and supporting a faster design workflow. However, by reducing the number of experiments we could possibly neglect some configurations which could be important for the analysis. In principle, we could have looked for a fractional design for this case study, but it would have been specific for the case study itself [29]. In order to derive more general rules for DOE, it would be necessary to apply the regression based approach presented in this paper to groups or typologies of reference buildings [22, 33], but this goes beyond the scope of this research. However, this can constitute the basis for future research, considering previous multi-scale simulation experience [58, 76].
In this case study, multiple DOE runs are used to account for the performance variability determined by envelope components and by occupant's comfort preferences and behaviour. Ideally, the parametric approach aims at understanding the impact of factors and to detect potentially critical assumptions already at the preliminary design level and to ensure the robustness of energy performance evaluation [28, 38]. In real building operation these variations can determine a very relevant gap between simulated and measured performance and, consequently, can compromise the cost-effectiveness of investments in energy efficiency, undermining the credibility of energy efficiency practices [33]. In other words, the objective of DOE simulation is that of addressing critically (i.e. with less optimistic assumptions) the effects of performance variability.
After that, in Section 3.2 the result for baseline design configuration is described more in detail, highlighting visually the relevant components characterizing building energy balance. Further, Section 3.3 describes the necessary steps and tools (in the workflow)

to link design and operational phase performance analysis through model calibration, and to test the applicability of regression models for performance prediction, using energy signatures [77].

Parametric simulation data are used to train multiple piecewise linear multivariate regression models. Finally, models are used for progressive calibration on measured data over a three year time monitoring, described in Section 3.4. In model training and testing phases visualization techniques are used in combination with numeric ones to enable an intuitive interpretation of results and to ease human interaction in an automated (or partially automated) calibration process.

3 Case study analysis

The case study chosen is a Passive House standard residential building constructed at south border of the Province of Forlì-Cesena, near Rimini, in the Emilia Romagna Region in Northern Italy. The case study building is characterized by highly insulated envelope components, a mechanical ventilation system with heat recovery (all-air system), a ground-source reversible heat pump system (GSHP) serving the mechanical ventilation system for heating and cooling demands and the domestic hot water demand. Further, a photovoltaic system for on-site electricity production and a solar thermal system for domestic hot water production integration are present. In the parametric simulations heat recovery has been considered in winter mode operation, taking into account also the relevant impact of auxiliaries [78].

3.1 Parametric simulation using Design of Experiment approach

As anticipated, the baseline configuration chosen for simulation is the one used originally for building design. The envelope parameters used in the grey-box model (lumped parameters) have been calibrated to reproduce the same heating demand of the original model in PHPP. Grey-box models are highly flexible, scalable and represent a good compromise between detail and accuracy when modelling building energy dynamics [79, 80]. These models have been used for yearly simulations, including all the energy demands from the building:

1. heating;
2. cooling;
3. domestic hot water (DHW);
4. lighting;
5. appliances.

Internal gains assumed in simulation and reported in Table 1 are averaged on a daily base and are very modest, considering the fact that the building, despite being very large, is actually used only by 4/5 people. It has to be underlined the fact that baseline configuration and DOE run 1 use constant operating schedules, as reported in Table 1, to maintain a comparability with the original PHPP model, but more realistic schedules are considered in the parametric simulation runs 2 (behaviour 1) and 3 (behaviour 2). In two-level DOE vary between two values, indicated with -1 and +1. The number of simulations depends on the amount of parameters chosen and on the combinatorial logic chosen. In this research we consider a full factorial DOE, for the reasons outlined in Section 2.1. The overall simulation data are summarized in Table 1.

Table 1: Baseline and Two-level Design Of Experiment simulation data

Group	Type	Unit	Baseline	Design of experiment Levels -1	Design of experiment Levels +1
Climate	UNI 10349:2016	-			
Geometry	Gross volume	m³	1557		
	Net volume	m³	1231		
	Heat loss surface area	m²	847		
	Net floor area	m²	444		
	Surface/volume ratio	1/m	0,54		
Envelope	U value external walls	$W/(m^2K)$	0,18	0,23	0,27
	U value roof	$W/(m^2K)$	0,17	0,21	0,26
	U value transparent components	$W/(m^2K)$	0,83	1,04	1,25
Activities	Internal gains (lighting, appliances and occupancy, daily average)	W/m^2	1	1	1.5
Control and operation	Heating set-point temperature	°C	20	20	22
	Cooling set-point temperature	°C	26	26	28
	Air-change rate (infiltration and mechanical ventilation with heat recovery in heating mode)	vol/h	0,2	0,2	0,4
	Shading factor (solar control summer mode)	-	0.5	0.5	0.7
	Domestic hot water demand	l/person/day	50	50	70
	Schedules – DOE constant operation	-	0.00-23.00	0.00-23.00	0.00-23.00
	Schedules – DOE behaviour 1	-	7.00-22.00	7.00-22.00	7.00-22.00
	Schedules – DOE behaviour 2	-	7.00-9.00, 17.00-22.00	7.00-9.00, 17.00-22.00	7.00-9.00, 17.00-22.00

In terms of temperature set-points, it has been considered an increase of two degrees in heating mode and an increase of two degrees also in cooling mode, to account respectively for an increased heating demand and for a reduced cooling demand. In terms of ventilation rate, infiltration and mechanical ventilation with heat recovery in heating mode have been considered.
Technical systems consist of a GSHP system, providing heating, cooling and domestic hot water (DHW), a rooftop photovoltaic plant (BIPV) and a solar thermal system with storage to integrate DHW production. Relevant sizing data of technical systems are reported in Table 2.

Table 2: Technical system sizing data

Group	Technology	Type	Unit	Value
Heating/Cooling system	GSHP (Ground-source heat pump)	Brine/Water Heat Pump	kW	8.4
		Borehole heat exchanger (2 double U boreholes)	m	100
On-site energy production	Building Integrated Photo-Voltaic (BIPV)	Polycrystalline silicon	kW_p	9.2
	Solar thermal	Glazed flat plate collector	m^2	4.32
		Domestic hot water storage	m^3	0.74

In order to simulate realistic operation conditions, coherent operating schedules have been created for heating, cooling, air-change rate (ventilation/infiltration) and internal gains (lighting, appliances, people). Schedules have been created using the methodology described in detail in previous research [14, 15] and the corresponding normative references [81]. As anticipated, the DOE simulation runs conducted are 3, one for each set of operation schedules, simulating different behavioural patterns of people living in the building:

1. operation is continuous as in baseline design configuration (constant operation profile);
2. operation is concentrated between 7.00 and 22.00 (variable operation profile, behaviour 1);
3. operation is concentrated between 7.00 and 9.00 and between 17.00 and 22.00 (variable operation profile, behaviour 2).

The indicators chosen for simulation output analysis are the following ones:

1. thermal demand for heating and cooling;
2. electricity demand for end-use (heating, cooling, DHW, appliances and lighting);
3. self-consumption of on-site RES electricity production;
4. renewable energy ratio (RER) [82];
5. load matching and grid interaction index [83, 84];
6. non-renewable primary energy demand;
7. CO_2 emission.

Most of the performance indicators have been calculated according to the methodology proposed in the standard ISO 52000-1 [85], which will be adopted in the future energy efficiency legislation at the EU level (overarching framework for the Energy Performance of Buildings, or EPB). Further, it has to been underlined the fact the KPIs chosen are substantially scalable, up to neighbourhood/district [65] scale, city scale [86] and regional/national scale [87].

As introduced before, the whole building energy demand has been taken into account, weighting delivered and imported electricity asymmetrically. The primary energy and emission factors assumed for calculation are the ones contained in Italian legislation regarding energy efficiency in buildings. However, while the delivered energy weight assumed is 1, the exported energy weight assumed here is 0.4, differently from the current building performance rating scheme adopted at the national level, which gives a 0 weight for exported energy.

The results obtained from DOE simulations have been used to report KPIs on a yearly base, considering respectively lower bound (LB) and upper bound (UB) of values obtained. The data are reported in Table 3, showing values for:

1. baseline design configuration;
2. lower and upper bound of overall data (DOE run 1, 2, 3);
3. constant operation data (DOE run 1);
4. behaviour 1 data (DOE run 2);
5. behaviour 2 data (DOE run 3).

Table 3: Baseline and Two-Level Design Of Experiment simulation data comparison – lower bound and upper bound of KPI yearly values

Balance level	KPI	Unit	Baseline	Design of Experiments							
				Overall		Constant		Behaviour 1		Behaviour 2	
				LB	UB	LB	UB	LB	UB	LB	UB
Zonal	Heating demand	kWh/m^2	19.3	17.2	39.6	19.2	39.6	18.0	36.2	17.2	33.8
	Cooling demand	kWh/m^2	10.8	0.8	12.6	0.8	12.3	1.2	12.6	1.1	11.2
Meter	Self-consumption	%	26.9	16.7	42.6	24.2	30.7	26.4	42.6	16.7	22.2
	Renewable Energy Ratio	%	91.7	75.8	97.3	81.3	94.6	79.4	97.3	75.8	93.2
	Load matching index	%	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	Primary Energy	kWh/m^2	5.0	1.5	24.3	3.2	18.5	1.5	20.1	3.9	24.3
	CO_2 Emission	kg/m^2	1.1	0.3	5.4	0.7	4.1	0.3	4.4	0.9	5.4

3.2 Analysis of baseline design configuration

Parametric simulation runs described in the previous Section are performed to create a possible spectrum of performance data, under uncertainty. On the other hand, baseline configuration represents simply the initial design hypothesis. In this Section baseline configuration is analysed to verify graphically, first of all, the suitability of a regression-based approach. For this reason, we report monthly data of indicators, plotted against average monthly external air temperature [39, 55, 56, 58], to identify correlations. For energy quantities in particular, we transform monthly data to derive the average power calculated over a monthly operation period; this method is called energy signature [77]. The objective of energy signatures is deriving weather normalized visualizations, suitable for monitoring and calibration in different climate conditions. Monthly monitoring of energy performance is not data intensive and can be done both manually and automatically, by means of data acquisition systems from meters. Further, it can easily scale from single buildings to building stock [58] and cities [24].

Monthly electricity demand composition and related energy signatures are reported in Figure 1 for the baseline configuration, showing the proportion of the different components of electricity demand in the building. The shape of data in energy signatures indicates the possibility of fitting total electric energy demand with a piecewise-linear regression model, while heating and cooling demand can be fitted with two separate linear regression models, as reported in literature [58, 88], allowing a physical interpretation of regression coefficients. The electricity meter balance with respect to demand and on-site production is reported in Figure 2, while delivered and exported energy data are reported in Figure 3, together with the related signatures. In this case also the data patterns can be approximated by linear and piece-wise linear models.

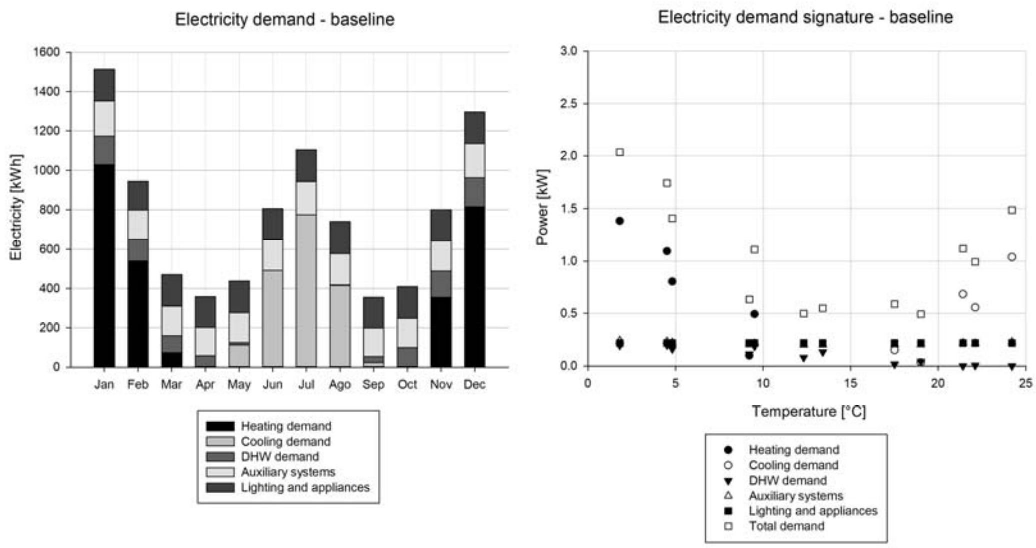

Figure 1: Electricity demand composition – monthly data and signatures

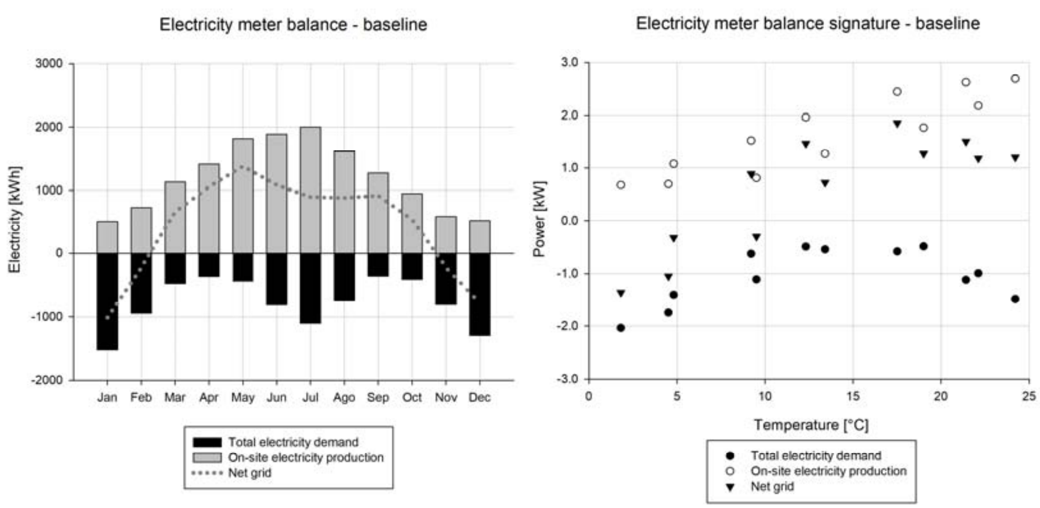

Figure 2: Electricity meter balance - on-site production and demand - monthly data and signatures

The values represented in Figure 2 highlight the fact that the photovoltaic system is able to satisfy the total electricity demand of the building on a yearly base. Further, the values reported in Figure 3 show the interaction of the building with the grid, by means of the patterns of delivered and exported energy. The analysis of these patterns shows indirectly when (on a daily base) the activity at the building level is concentrated, because we can discriminate the quantity of energy self-consumed depending on the climatic variables (temperature and solar radiation). In this way, it is possible to test if the schedules assumed for dynamic simulation are approximately correct even with low resolution data (monthly in this case). Therefore, further research development in this direction is possible by introducing more information about user behaviour (e.g. integrating long-term monthly measurements with periodic short-term measurements at hourly/sub-hourly intervals [59, 60, 89]).

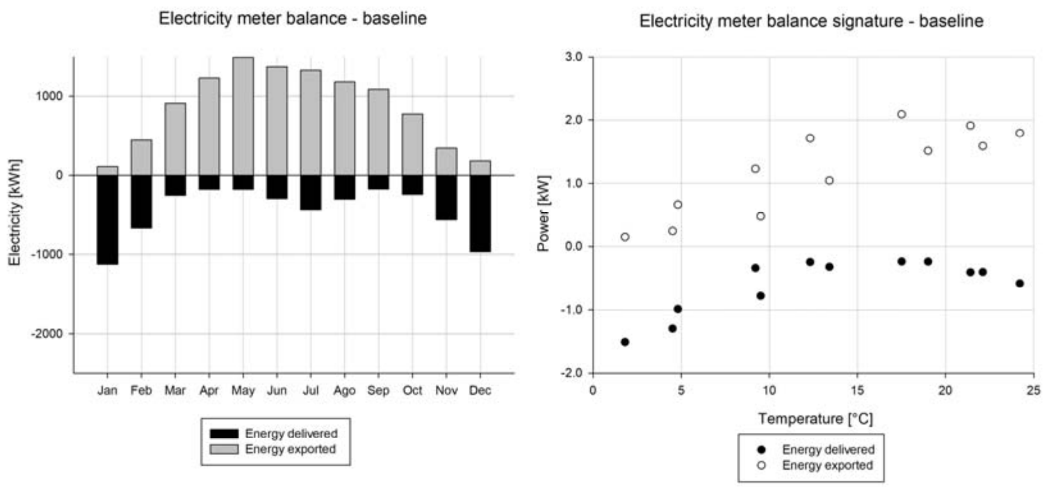

Figure 3: Electricity meter balance - delivered and exported energy - monthly data and signatures

Another way of accounting for the variability of the building interaction with the grid are load matching and grid interaction indexes, which are reported in Figure 4. Load matching index assumes the maximum value of 100% by definition [83, 84].

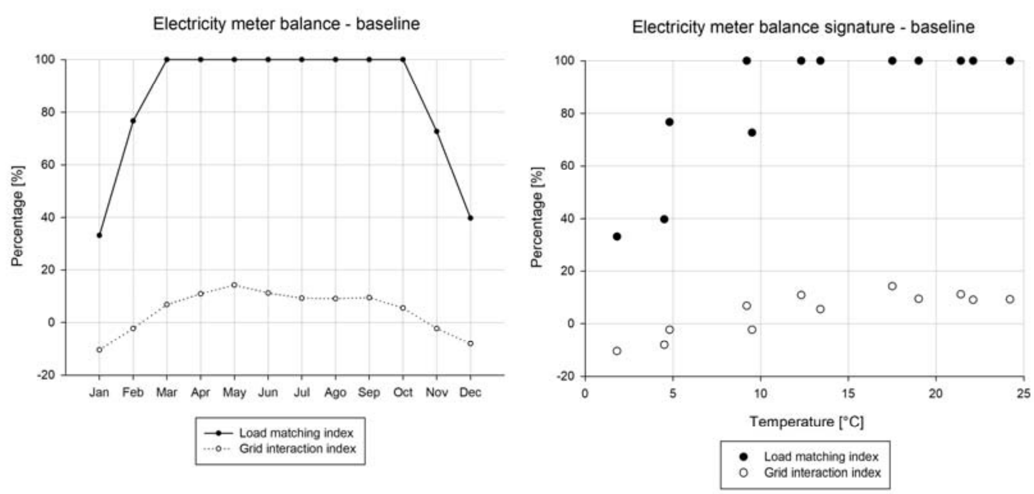

Figure 4: Electricity meter balance – load matching and grid interaction indexes – monthly data and signatures

3.3 Linking design and operational performance analysis

The aim of this research was establishing a link between DOE simulation data and operational data, in order to calibrate progressively simple predictive models, maintaining at the same time a comparability with initial parametric estimates. Regression models are essential for two fundamental reasons:

1. providing a simple but effective approach for performance monitoring, for the reasons outlined in Section 2;
2. performing weather normalization of simulation results, generated with a standard climate data file, reported in Table 1.

The choice was adopting a piecewise linear multivariate regression approach, considering the general motivations reported in Section 2.1. Actually multiple types of meta-models can be considered for calibration purpose as described in Section 2, but we decided to use the simplest possible approach to ease model calibration and, consequently, performance monitoring, creating a procedure that could possibly scale with respect to temporal [59] and spatial resolution of data [24, 39], using multi-level analysis [70]. Further, among all the data presented in Sections 3.1 and 3.2, we decided to focus on the total aggregated electricity demand, plotted in Figure 2 for baseline design configuration, even though the model can be further decomposed with respect to zonal energy balance components [58], represented in Figure 1.
The piecewise linear multivariate regression models proposed are reported in Table 4. The overall predictive model is the combination of three linear submodels, respectively for heating, cooling and baseline demand. Two types of models are considered:

1. type 1, accounting only for external air temperature dependence;
2. type 2, accounting for both external air temperature and solar radiation dependence.

Table 4: Regression models for heating, cooling and baseline demand analysis

Demand	Model type 1	Model type 2
Heating	$q_{h,1} = a_0 + a_1\theta_e + \varepsilon$	$q_{h,2} = b_0 + b_1\theta_e + b_2 I_{sol} + \varepsilon$
Cooling	$q_{c,1} = c_0 + c_1\theta_e + \varepsilon$	$q_{c,2} = d_0 + d_1\theta_e + d_2 I_{sol} + \varepsilon$
Baseline	$q_{b,1} = e_0 + e_1\theta_e + \varepsilon$	$q_{b,2} = f_0 + f_1\theta_e + f_2 I_{sol} + \varepsilon$

External temperature is the most important regressor for weather normalization [90]. However, we decided to include also solar radiation as a regressor, considering the fact that we are analysing a Passive House standard building, in which the impact of solar gains is relevant and a solar thermal system for the integration of DHW production is present as well. Nonetheless, similar approaches can be used for solar photo-voltaic [91] and solar thermal plants [92, 93].
In order to evaluate and compare properly simulation data in design phase and measured data in operation phase, we used a set of statistical indicators. We decided to train first the two different types of multivariate piecewise linear regression models on simulated data, in order to test them in the first year of operation with respect to measured data. Then, from the second year onward, models are directly trained on measured data. This part of the research is described in detail in Section 3.4.
Going back to statistical indicators, the goodness of fit of a regression model can be expressed by the determination coefficient R^2 that can assume values ranging from 0 to 1 (or 0 to 100%, if expressed in percentual terms), where 1 means that the data fitting is perfect. The formula for R^2 is the following one:

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} = 1 - \frac{\sum_i \left(y_i - \hat{y}_i\right)^2}{\sum_i \left(y_i - \overline{y}_i\right)^2} \tag{1}$$

R^2 is an important indicator of the goodness of fit, but it is not the only one to be considered. We decided to consider also *MAPE* (Mean Absolute Percentage Error), to account for the average absolute value of the difference among measured and predicted data, normalized with respect to measured data themselves. *MAPE* is calculated as follows.

$$MAPE = \frac{1}{n}\sum_i \left|\frac{M_i - P_i}{M_i}\right| \cdot 100 \tag{2}$$

Further, in the state-of-the-art of model calibration procedures [11, 12, 52-54] other two metrics are employed, *NMBE* and *Cv(RMSE)*. *NMBE* (Normalized Mean Bias Error) is the total sum of the differences between measured (or simulated, before operation) and predicted energy consumption at the calculation time intervals (e.g. monthly, hourly) of the considered period. The difference is then divided by the sum of the measured (or simulated) energy consumption.

$$NMBE = \frac{\sum_i (M_i - P_i)}{\sum_i M_i} \cdot 100 \tag{3}$$

A positive value of *NMBE* implies a model overestimation of energy consumption, viceversa a negative value implies an underestimation.
The *RMSE* (Root Mean Squared Error) is a measure of the sample deviation of the differences between measured values and values predicted by the model. *Cv(RMSE)* is the Coefficient of Variation of *RMSE* and is calculated as the *RMSE* normalized to the mean of the measured values. *Cv(RMSE)* represents a normalized measure of the variability among measured (or simulated, before operation) and predicted data. It specifies the overall uncertainty in the prediction of the building energy consumption, reflecting the errors size and the amount of scatter. Lower *Cv(RMSE)* values indicate a better calibrated model.

$$Cv\left(RMSE\right) = \frac{RMSE}{A} \cdot 100 \tag{4}$$

$$RMSE = \sqrt{\frac{\sum_i \left(M_i - P_i\right)^2}{n}} \tag{5}$$

$$A = \frac{\sum_i M_i}{n} \tag{6}$$

The threshold limits considered at the state-of-the-art are reported in Table 5, considering the most relevant protocols for measurement and verification (M&V) existing today.

Table 5: Threshold limits of metrics for model calibration with monthly data

Metric		ASHRAE Guidelines 14	IPMVP	FEMP
MBE	%	± 5	± 20	± 5
Cv(RMSE)	%	15	-	15

Simulated parametric data (DOE) are used as reference to link design (when no measured data are available) and operational performance analysis. As specified before, we concentrated on the analysis on simulated total aggregated electric energy demand, training regression models respectively on:

1. lower bound (LB) and upper bound (UB) data for the overall DOE runs dataset (runs 1, 2, 3);
2. three subsets of data, corresponding to constant operation (DOE run 1), behaviour 1 (DOE run 2) and behaviour 2 (DOE run 3).

The results obtained are reported in Table 6, showing the goodness of fit of piecewise linear regression models to simulated data in all the conditions.

Table 6: Training of regression models on DOE simulation data

Regression model	Dataset		Training - simulation data DOE			
			R^2	*MAPE*	*NMBE*	*Cv(RMSE)*
			%	%	%	%
Type 1	Overall	LB	93.65	9.34	0.06	13.58
		UB	96.64	7.33	0.02	9.01
	Constant	LB	93.97	9.66	0.07	14.22
		UB	96.16	8.81	0.01	10.63
	Behaviour 1	LB	93.44	9.38	0.07	13.19
		UB	96.56	7.31	0.01	9.07
	Behaviour 2	LB	93.43	9.33	0.06	12.79
		UB	96.52	7.23	0.01	8.96
Type 2	Overall	LB	99.90	1.42	-0.02	1.65
		UB	99.77	1.93	-0.01	2.36
	Constant	LB	99.86	3.77	4.82	8.33
		UB	99.65	8.55	-3.57	5.89
	Behaviour 1	LB	99.91	1.02	0.03	1.47
		UB	99.77	1.88	0.00	2.32
	Behaviour 2	LB	99.92	1.11	0.03	1.42
		UB	99.69	2.17	0.00	2.67

We can also represent easily the results of model training process graphically. In this research we decided to plot the distribution of simulated monthly data on a yearly base, together with the corresponding energy signatures (lower and upper bound of simulation data envelopment), compared with model type 1 and model type 2 regression results. The results are represented in Figure 5 for the overall dataset, and in Figures 6, 7 and 8, respectively for constant operation, behaviour 1 and behaviour 2. The use of interval data for parametric simulation is substantially comparable to an epistemic uncertainty assumption [94].

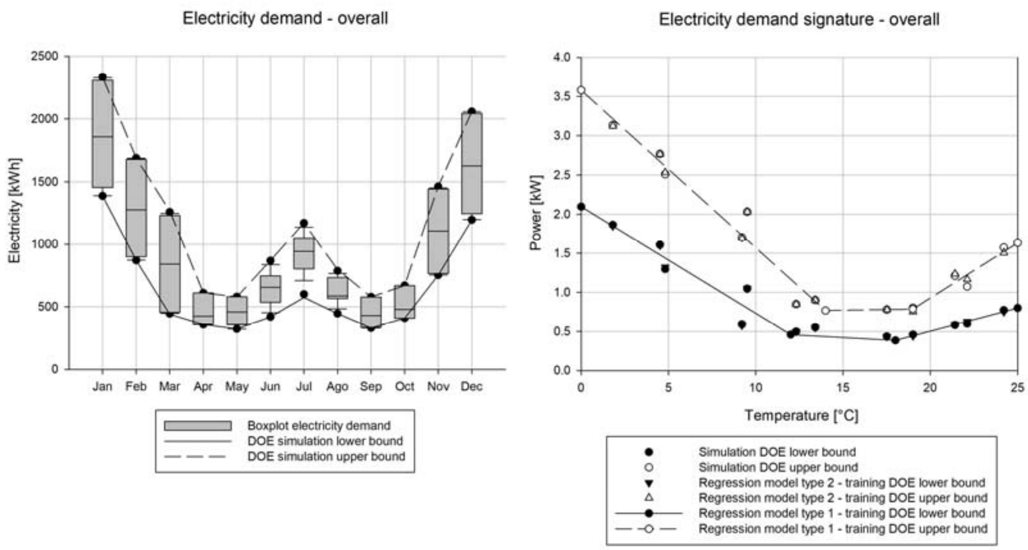

Figure 5: Total simulated monthly electricity demand distribution (boxplot) and comparison between simulated and piecewise linear multivariate regression (energy signatures) – overall data

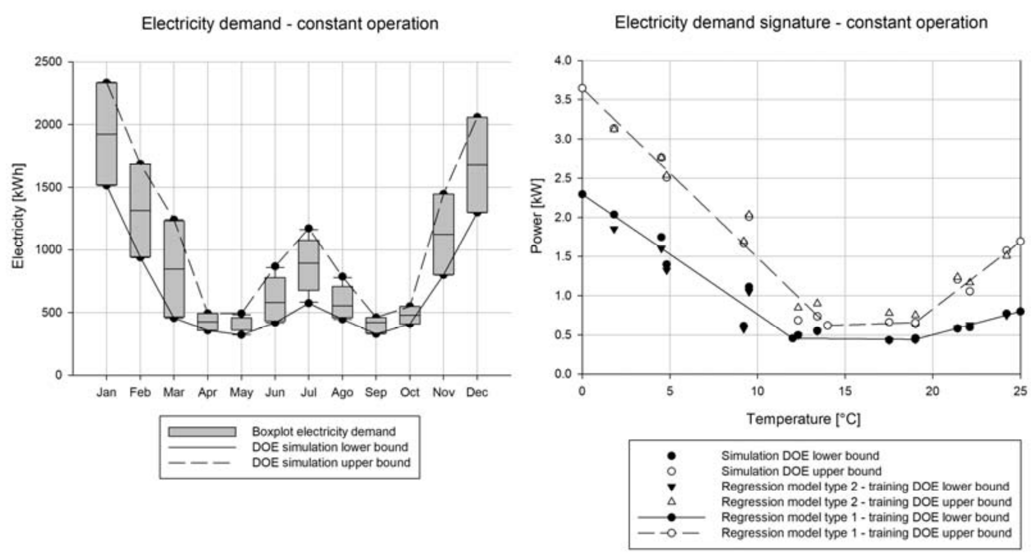

Figure 6: Total simulated monthly electricity demand distribution (boxplot) and comparison between simulated and piecewise linear multivariate regression (energy signatures) – constant operation data

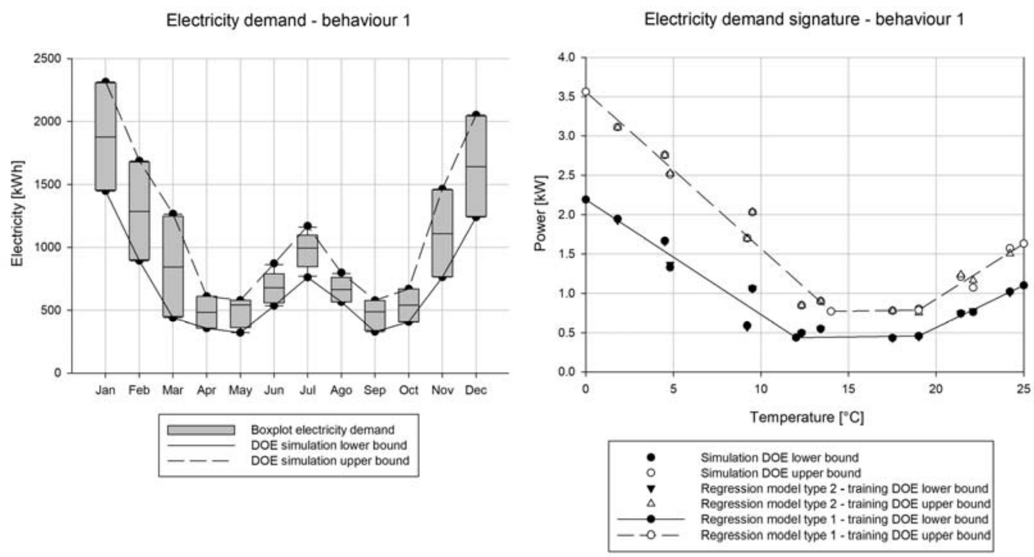

Figure 7: Total simulated monthly electricity demand distribution (boxplot) and comparison between simulated and piecewise linear multivariate regression (energy signatures) – behavior 1 data

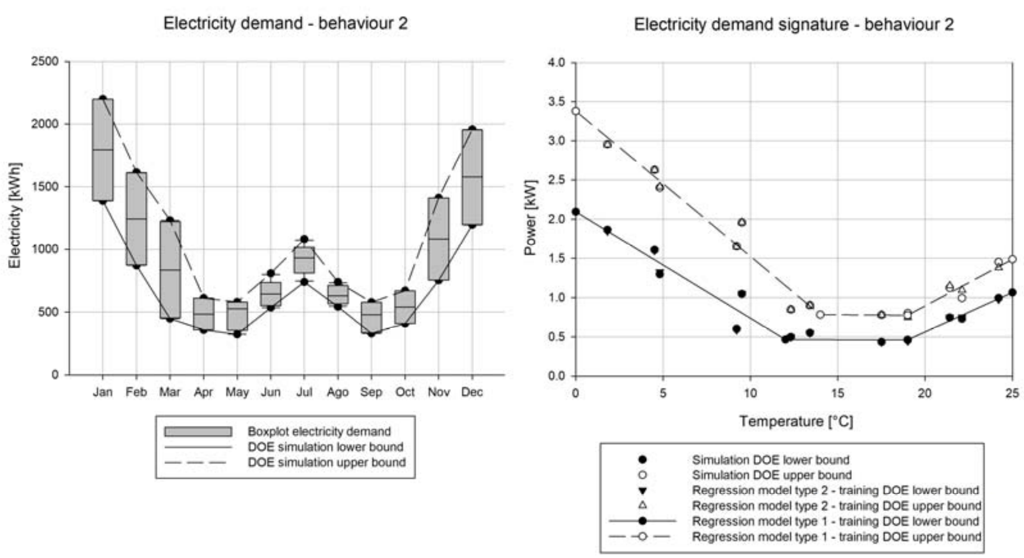

Figure 8: Total simulated monthly electricity demand distribution (boxplot) and comparison between simulated and piecewise linear multivariate regression (energy signatures) – behavior 2 data

3.4 Monitoring and incremental model calibration

We decided to use both model type 1 and type 2 for monitoring and incremental model calibration process. The analysis is concentrated on total aggregated electricity demand, as specified before. Models are initially trained respectively on the lower and upper bounds of overall DOE runs data, when measured data are not available (design phase). In this way, we consider the largest possible spectrum of data variability, given by the underlying assumptions for the generation of DOE cases, reported in Section 3.1. After the first year of operation, models are trained on measured data.

The results of model training and testing for the three years of monitoring period are plotted in Tables 7 and 8, respectively for model type 1 and type 2. The phases and subphases of the process are reported in Tables, considering:

1. design phase, model training on DOE simulation data;
2. operation phase, initial operation, uncalibrated model;
3. operation phase, partial calibration, models don't reach calibration thresholds reported in Table 5;
4. operation phase, calibration, model reaches calibration thresholds reported in Table 5.

In general, the results highlighted the necessity of considering multiple statistical indicators in the calibration process. In fact, R^2 is highly dependent on the scatter of data and therefore cannot be considered as the only parameter for predictive model validation, because this could lead to misleading conclusions. In fact, model R^2 can be high even if the model is uncalibrated, uncovering a systematic error. Therefore, the predictive model is acceptable only if its calibration indicators *NMBE* and *Cv(RMSE)* are within the limits reported in Table 5, according to calibration protocols in M&V.

Table 7: Incremental calibration during three years of operation – Model type 1

Phase	Sub-phase	Training dataset	Testing dataset	R^2	MAPE	NMBE	Cv(RMSE)
				%	%	%	%
Design	Model training	Simulated data DOE - Overall LB		93.65	9.34	0.06	13.58
Design	Model training	Simulated data DOE - Overall UB		96.64	7.33	0.02	9.01
Operation	Initial operation		Measured data – Year 1	76.88	35.51	-50.23	37.60
Operation	Initial operation		Measured data - Year 1	73.08	33.35	20.59	44.39
Operation	Partial calibration	Measured data – Year 1		81.33	12.03	0.02	14.60
			Measured data – Year 2	91.97	13.08	-13.82	16.12
Operation	Partial calibration	Measured data – Year 1 and 2		82.64	11.44	0.04	13.44
			Measured data – Year 3	69.74	18.40	-6.95	19.75

Table 8: Incremental calibration during three years of operation – Model type 2

Phase	Sub-phase	Training dataset	Testing dataset	R^2	MAPE	NMBE	Cv(RMSE)
				%	%	%	%
Design	Model training	Simulated data DOE - Overall LB		99.90	1.42	-0.02	1.65
Design	Model training	Simulated data DOE - Overall UB		99.78	1.93	-0.01	2.36
Operation	Initial operation		Measured data – Year 1	69.91	38.80	-36.46	41.86
Operation	Initial operation		Measured data - Year 1	75.99	28.16	21.33	40.64
Operation	Partial calibration	Measured data – Year 1		85.93	8.05	0.04	12.76
			Measured data – Year 2	88.45	13.72	-13.75	17.07
Operation	Calibration	Measured data – Year 1 and 2		86.07	9.97	0.05	12.02
			Measured data – Year 3	87.54	11.97	-2.21	12.50

As we can see from the data in Tables 7 and 8, model type 1 remains partially calibrated even in the third year of operation, while model type 2 reaches calibration. With low temporal resolution data (i.e. monthly data) we need at least two years of measured data to be able to calibrate a model. It is worth noting that two years of data are also generally considered as a minimal requirement in energy audits. The research highlights the fact the we can monitor easily and inexpensively long-term performance with a spatial scalability up to the utility level [21, 24, 25]. Additionally, models can scale in time up to daily and hourly data resolution [59, 60] to reach calibration within a more limited time-frame of operation, when more data are available. In any case, we consider periodic recalibration fundamental to monitor long-term performance evolution, as indicated also in other studies [89]. Beside statistical indicators used in the calibration process, it is important to provide simple visual analytical tools to render the process of calibration and long-term performance monitoring more intuitive and transparent. In this research we decided to use three visualization tools:

1. time series of measured and predicted energy consumption data (electricity demand in this case), Figure 9;
2. time series of model deviations among measurements and predictions, Figure 10;
3. time series of cumulative sum of deviations (CUSUM) chart, Figure 11.

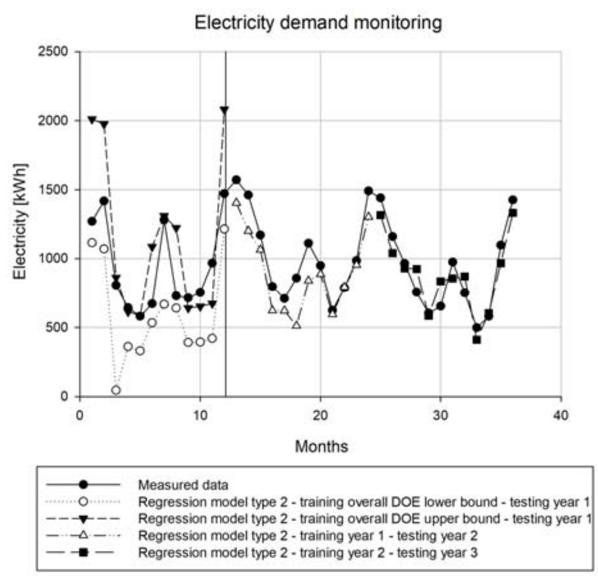

Figure 9: Electricity demand monitoring – time series of monthly data measured and predicted by different models, three years monitoring period

The time series in Figure 9 highlight the progressive calibration process, reached in the third year of operation as explained before, with the substantial alignment among measured and predicted data. The underlying model (a monthly model for the prediction of aggregated electricity consumption) is a “static” model (energy signature), as there is no explicit dependence on time but only on weather conditions and operating hours considered [77]. Subsequently, the deviations among measurements and predictions are calculated according to the following formula.

$$D_i = M_i - P_i \tag{7}$$

A positive deviation implies that the model is underestimating energy consumption at that point in time (i.e. the measured consumption is higher than predicted), while a negative deviation implies an overestimation of energy consumption (i.e. the measured consumption is lower than predicted). In this case study we can see how deviations in Figure 10 are progressively decreasing and how calibrated model deviations tend to oscillate around zero.

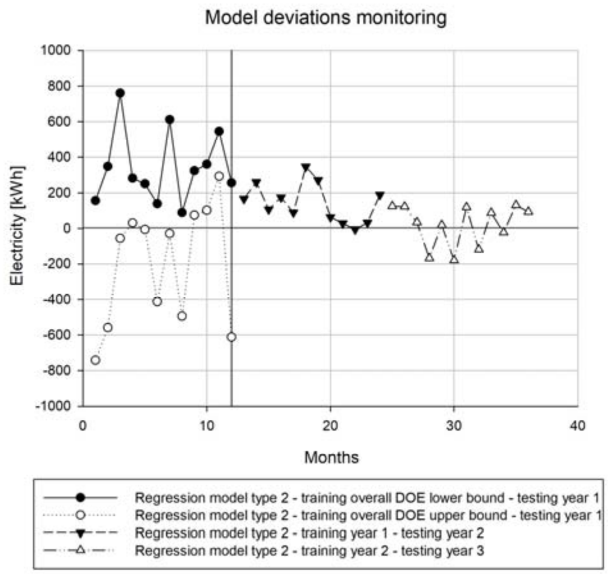

Figure 10: Electricity demand monitoring – deviations among measured and predicted data, three years monitoring period

Further, the cumulative sum of deviations is reported to ease the detection of model drift with respect to measured data. By using the incremental sum of deviation we can identify the cumulative difference between measured and predicted data at a point in time. A positive sum of deviations indicates that the actual energy demand is higher than predicted (i.e. model is underestimating consumption), while a negative sum of deviations indicates that actual energy demand is lower than predicted (i.e. model is overestimating consumption). In this research, the cumulative sum of deviations in the third year of operation for model type 2 is practically equal to zero, with a minimal difference between measurement and prediction (around 2%), confirming the reliability of the calibrated model.

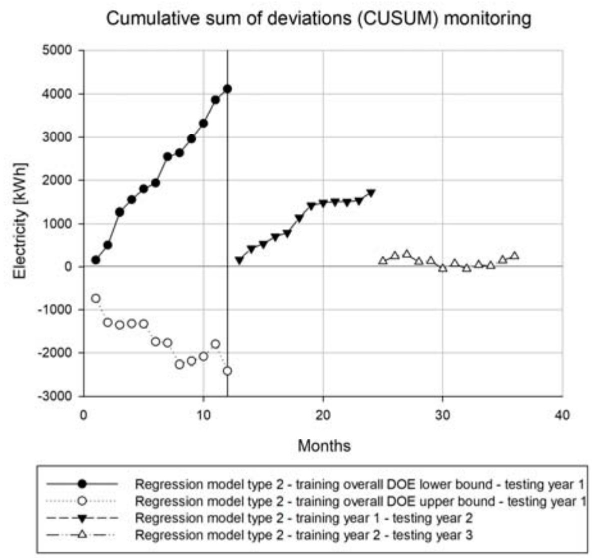

Figure 11: Electricity demand monitoring – cumulative sum of deviations among measured and predicted data, three years monitoring period

Finally, Figure 12 summarizes the whole procedure representing, on the left side, a priori parametric DOE estimates, reported previously in Figure 5, comparing them with measured data. On the right side of Figure 12, calibrated models (a posteriori) are reported.

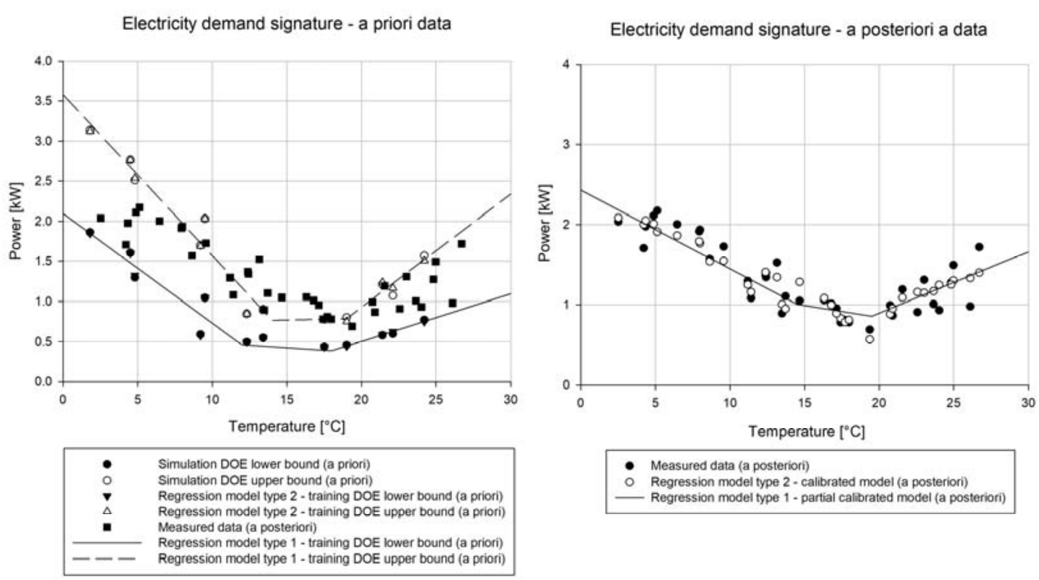

Figure 12: Electricity demand monitoring – overall analysis of a priori and a posteriori data, three years monitoring period

It is worth noting that, even if model type 1 remains partially calibrated, it is still useful to get a simple visual representation of the relevant differences with respect to heating, cooling and baseline demand, by comparing positions and slopes of regression lines. A further analysis of the components of the energy balance can help detecting root causes of anomalies in energy demand (i.e. considering a grey-box interpretation of regression coefficients) [88], and will be part of future research on this case study.

4 Conclusion

Design optimization in buildings has often been oriented towards specific paradigms without considering properly variability and uncertainty in design assumptions and without questioning relevant factors that could undermine the fundamental goals of paradigms themselves. Passive House standard is a rigorous voluntary scheme for high efficiency buildings, but the use of this standard in the Mediterranean area, characterized by a mild climate, can be debatable, considering climate change scenarios, and relevant uncertainties in performance simulation. For this reason, we selected a Passive House building in Italy as case study. The ability to monitor long-term performance inexpensively and to use easily accessible data is important for multiple stakeholders in the building sector. In fact, the analysis of building performance data using simple, robust and scalable techniques can provide relevant analytical insights improve design and operational practices, as well as to orient policies. In other words, our decisions can be based on feedbacks from the actual performance of building stock, rather than on (simulation-based) estimates that can be very far from reality in many cases, leading to a consistent performance gap. In this research we illustrated how parametric simulation (to test robustness of design configurations) can be combined with regression-based calibration approaches (state of the art of performance monitoring), establishing a continuity between design and operational phase analysis. In this way, we can assume a more critical perspective on building performance, necessary to ensure the credibility of energy efficiency practices, especially with respect to innovative business models where the analysis of cost-optimal levels of investment is a pre-requisite. In fact, risk analysis for efficiency investments is a particularly relevant problem today, embodying the necessity of evaluating performance variability in depth. Additionally, variability in performance outcomes determined by occupants' preferences and behaviour have been often neglected in design but they are essential for the success of innovative practices and policies in buildings. While in the case study presented we concentrated on the analysis of aggregated electricity demand, there are other relevant quantities, such as delivered and exported energy or the percentage of self-consumption of RES production, which can change radically when realistic operation profiles are used instead of standardized assumptions. Even an analysis of low temporal resolution data (e.g. monthly automatically metered data) conducted in an appropriate way (i.e. when sufficient metadata are available) can help unconvering the impact of user behaviour. This impact can determine a large variation of performance both in economic terms, depending on the specific business model adopted, and in environmental terms, because of temporal variation of interaction with energy infrastructures (i.e. delivered and exported energy patterns). Finally, the approach can be developed further when thermal metering data are available, and this will be part of future research.

As a conclusion, instead of simply evaluating the formal correctness of modelling approaches, it is necessary to introduce progressively parametric design in practice and in policy, considering, on the one hand, more realistic operation profiles for buildings and, on the other hand, more detailed and realistic data for grid interaction (energy conversion factors, tariffs, CO_2 emission, etc.). In this way, design practices in the built environment could evolve coherently with energy infrastructures, exploiting sinergies in terms of technology and business models. However, in order to progressively overcome limitations, it is necessary to work coherently on modelling and on the availability of relevant design and operational data, integrating efficiently long-term (low resolution) with short-term (high resolution) monitoring.

Acknowledgments

The authors would like to acknowledge Eng. Giuseppe Salvà Gagliolo for the initial data collection and analysis on the case study building.

References

[1] Berardi U. A cross-country comparison of the building energy consumptions and their trends. Resources, Conservation and Recycling. 2017;123:230-41.

[2] BPIE. Europe’s buildings under the microscope. Buildings Performance Institute Europe (BPIE); 2011.

[3] D'Agostino D, Zangheri P, Cuniberti B, Paci D, Bertoldi P. Synthesis Report on the National Plans for Nearly Zero Energy Buildings (NZEBs). JRC EU Commission; 2016.

[4] Porritt SM, Cropper PC, Shao L, Goodier CI. Ranking of interventions to reduce dwelling overheating during heat waves. Energy and Buildings. 2012;55:16-27.

[5] Tabatabaei Sameni SM, Gaterell M, Montazami A, Ahmed A. Overheating investigation in UK social housing flats built to the Passivhaus standard. Building and Environment. 2015;92:222-35.

[6] Jentsch MF, Bahaj AS, James PAB. Climate change future proofing of buildings—Generation and assessment of building simulation weather files. Energy and Buildings. 2008;40(12):2148-68.

[7] de Wilde P. The gap between predicted and measured energy performance of buildings: A framework for investigation. Automation in Construction. 2014;41:40-9.

[8] Imam S, Coley DA, Walker I. The building performance gap: Are modellers literate? Building Services Engineering Research and Technology. 2017;38(3):351-75.

[9] Herring H, Roy R. Technological innovation, energy efficient design and the rebound effect. Technovation. 2007;27(4):194-203.

[10] Rosenow J, Galvin R. Evaluating the evaluations: Evidence from energy efficiency programmes in Germany and the UK. Energy and Buildings. 2013;62(Supplement C):450-8.

[11] Coakley D, Raftery P, Keane M. A review of methods to match building energy simulation models to measured data. Renewable and Sustainable Energy Reviews. 2014;37:123-41.

[12] Fabrizio E, Monetti V. Methodologies and Advancements in the Calibration of Building Energy Models. Energies. 2015;8(4):2548.

[13] Tagliabue LC, Manfren M, De Angelis E. Energy efficiency assessment based on realistic occupancy patterns obtained through stochastic simulation. Modelling Behaviour: Springer; 2015. p. 469-78.

[14] Tagliabue LC, Manfren M, Ciribini ALC, De Angelis E. Probabilistic behavioural modeling in building performance simulation—The Brescia eLUX lab. Energy and Buildings. 2016;128:119-31.

[15] Cecconi FR, Manfren M, Tagliabue LC, Ciribini ALC, De Angelis E. Probabilistic behavioral modeling in building performance simulation: A Monte Carlo approach. Energy and Buildings. 2017;148:128-41.

[16] Dodd N, Donatello S, Garbarino E, Gama-Caldas M. Identifying macro-objectives for the life cycle environmental performance and resource efficiency of EU buildings. JRC EU Commission; 2015.
[17] CESBA. Common European Sustainable Building Assessment (http://wiki.cesba.eu/wiki/Main_Page), accessed 26/05/2017. 2017.
[18] EUSSD. European Commission, Environment, Sustainable Buildings (http://ec.europa.eu/environment/eussd/buildings.htm), accessed 20/04/2017. 2017.
[19] Kyriakidis A, Michael A, Illampas R, Charmpis DC, Ioannou I. Thermal performance and embodied energy of standard and retrofitted wall systems encountered in Southern Europe. Energy. 2018;161:1016-27.
[20] Pomponi F, Moncaster A. Scrutinising embodied carbon in buildings: The next performance gap made manifest. Renewable and Sustainable Energy Reviews. 2018;81:2431-42.
[21] Acquaviva A, Apiletti D, Attanasio A, Baralis E, Bottaccioli L, Castagnetti FB, et al. Energy signature analysis: Knowledge at your fingertips. Conference Energy signature analysis: Knowledge at your fingertips. IEEE, p. 543-50.
[22] Pistore L, Pernigotto G, Cappelletti F, Romagnoni P, Gasparella A. From energy signature to cluster analysis: an integrated approach. IV High Performance Buildings Conference at Purdue: Purdue University; 2016.
[23] Stadler P, Girardin L, Ashouri A, Maréchal F. Contribution of Model Predictive Control in the Integration of Renewable Energy Sources within the Built Environment. Frontiers in Energy Research. 2018;6(22).
[24] Abdolhosseini Qomi MJ, Noshadravan A, Sobstyl JM, Toole J, Ferreira J, Pellenq RJ-M, et al. Data analytics for simplifying thermal efficiency planning in cities. Journal of The Royal Society Interface. 2016;13(117).
[25] Meng Q, Mourshed M. Degree-day based non-domestic building energy analytics and modelling should use building and type specific base temperatures. Energy and Buildings. 2017;155(Supplement C):260-8.
[26] Kohler M, Blond N, Clappier A. A city scale degree-day method to assess building space heating energy demands in Strasbourg Eurometropolis (France). Applied Energy. 2016;184(Supplement C):40-54.
[27] Jaffal I, Inard C, Ghiaus C. Fast method to predict building heating demand based on the design of experiments. Energy and Buildings. 2009;41(6):669-77.
[28] Kotireddy R, Hoes P-J, Hensen JLM. A methodology for performance robustness assessment of low-energy buildings using scenario analysis. Applied Energy. 2018;212:428-42.
[29] Schlueter A, Geyer P. Linking BIM and Design of Experiments to balance architectural and technical design factors for energy performance. Automation in Construction. 2018;86:33-43.
[30] Shiel P, Tarantino S, Fischer M. Parametric analysis of design stage building energy performance simulation models. Energy and Buildings. 2018;172:78-93.
[31] EEFIG. Energy Efficiency – the first fuel for the EU Economy, How to drive new finance for energy efficiency investments Energy Efficiency Financial Institutions Group; 2015.
[32] Saheb Y, Bodis K, Szabo S, Ossenbrink H, Panev S. Energy Renovation: The Trump Card for the New Start for Europe. JRC EU Commission; 2015.
[33] Aste N, Adhikari RS, Manfren M. Cost optimal analysis of heat pump technology adoption in residential reference buildings. Renewable Energy. 2013;60:615-24.

[34] Menezes AC, Cripps A, Bouchlaghem D, Buswell R. Predicted vs. actual energy performance of non-domestic buildings: Using post-occupancy evaluation data to reduce the performance gap. Applied Energy. 2012;97:355-64.

[35] Sunikka-Blank M, Galvin R. Introducing the prebound effect: the gap between performance and actual energy consumption. Building Research & Information. 2012;40(3):260-73.

[36] Evins R. A review of computational optimisation methods applied to sustainable building design. Renewable and Sustainable Energy Reviews. 2013;22:230-45.

[37] Nguyen A-T, Reiter S, Rigo P. A review on simulation-based optimization methods applied to building performance analysis. Applied Energy. 2014;113:1043-58.

[38] Ligier S, Robillart M, Schalbart P, Peuportier B. Energy performance contracting methodology based upon simulation and measurement. Conference Energy performance contracting methodology based upon simulation and measurement.

[39] Manfren M, Aste N, Moshksar R. Calibration and uncertainty analysis for computer models – A meta-model based approach for integrated building energy simulation. Applied Energy. 2013;103:627-41.

[40] Koulamas C, Kalogeras AP, Pacheco-Torres R, Casillas J, Ferrarini L. Suitability analysis of modeling and assessment approaches in energy efficiency in buildings. Energy and Buildings. 2018;158:1662-82.

[41] Aste N, Manfren M, Marenzi G. Building Automation and Control Systems and performance optimization: A framework for analysis. Renewable and Sustainable Energy Reviews. 2017;75:313-30.

[42] Østergård T, Jensen RL, Maagaard SE. A comparison of six metamodeling techniques applied to building performance simulations. Applied Energy. 2018;211(Supplement C):89-103.

[43] Talele S, Traylor C, Arpan L, Curley C, Chen C-F, Day J, et al. Energy modeling and data structure framework for Sustainable Human-Building Ecosystems (SHBE) — a review. Frontiers in Energy. 2018.

[44] Yoshino H, Hong T, Nord N. IEA EBC annex 53: Total energy use in buildings—Analysis and evaluation methods. Energy and Buildings. 2017;152:124-36.

[45] Kylili A, Fokaides PA, Lopez Jimenez PA. Key Performance Indicators (KPIs) approach in buildings renovation for the sustainability of the built environment: A review. Renewable and Sustainable Energy Reviews. 2016;56:906-15.

[46] Al Gharably M, DeCarolis JF, Ranjithan SR. An enhanced linear regression-based building energy model (LRBEM+) for early design. Journal of Building Performance Simulation. 2016;9(2):115-33.

[47] Asadi S, Amiri SS, Mottahedi M. On the development of multi-linear regression analysis to assess energy consumption in the early stages of building design. Energy and Buildings. 2014;85:246-55.

[48] Ipbüker C, Valge M, Kalbe K, Mauring T, Tkaczyk AH. Case Study of Multiple Regression as Evaluation Tool for the Study of Relationships between Energy Demand, Air Tightness, and Associated Factors. Journal of Energy Engineering. 2016;143(1):04016027.

[49] Hygh JS, DeCarolis JF, Hill DB, Ranjithan SR. Multivariate regression as an energy assessment tool in early building design. Building and Environment. 2012;57:165-75.

[50] Catalina T, Virgone J, Blanco E. Development and validation of regression models to predict monthly heating demand for residential buildings. Energy and buildings. 2008;40(10):1825-32.

[51] Kavousian A, Rajagopal R. Data-driven benchmarking of building energy efficiency utilizing statistical frontier models. Journal of Computing in Civil Engineering. 2013;28(1):79-88.
[52] ASHRAE Guideline 14-2014: Measurement of Energy, Demand, and Water Savings; American Society of Heating, Refrigerating and Air-Conditioning Engineers: Atlanta, GA, USA, 2014.
[53] IPMVP New Construction Subcommittee. International Performance Measurement & Verification Protocol: Concepts and Option for Determining Energy Savings in New Construction, Volume III; Efficiency Valuation Organization (EVO): Washington, DC, USA, 2003.
[54] FEMP. Federal Energy Management Program, M&V Guidelines: Measurement and Verification for Federal Energy Projects Version 3.0, U.S. Department of Energy Federal Energy Management Program. Washington, DC, USA, 2008.
[55] Masuda H, Claridge DE. Statistical modeling of the building energy balance variable for screening of metered energy use in large commercial buildings. Energy and Buildings. 2014;77:292-303.
[56] Paulus MT, Claridge DE, Culp C. Algorithm for automating the selection of a temperature dependent change point model. Energy and Buildings. 2015;87:95-104.
[57] Aste N, Leonforte F, Manfren M, Mazzon M. Thermal inertia and energy efficiency – Parametric simulation assessment on a calibrated case study. Applied Energy. 2015;145:111-23.
[58] Tronchin L, Manfren M, Tagliabue LC. Optimization of building energy performance by means of multi-scale analysis – Lessons learned from case studies. Sustainable Cities and Society. 2016;27:296-306.
[59] Jalori S, T Agami Reddy PhD P. A unified inverse modeling framework for whole-building energy interval data: daily and hourly baseline modeling and short-term load forecasting. ASHRAE Transactions. 2015;121:156.
[60] Jalori S, T Agami Reddy PhD P. A new clustering method to identify outliers and diurnal schedules from building energy interval data. ASHRAE Transactions. 2015;121:33.
[61] Li Q, Augenbroe G, Brown J. Assessment of linear emulators in lightweight Bayesian calibration of dynamic building energy models for parameter estimation and performance prediction. Energy and Buildings. 2016;124:194-202.
[62] Booth A, Choudhary R, Spiegelhalter D. A hierarchical Bayesian framework for calibrating micro-level models with macro-level data. Journal of Building Performance Simulation. 2013;6(4):293-318.
[63] Adhikari RS, Aste N, Manfren M. Multi-commodity network flow models for dynamic energy management – Smart Grid applications. Energy Procedia. 2012;14:1374-9.
[64] Manfren M. Multi-commodity network flow models for dynamic energy management – Mathematical formulation. Energy Procedia. 2012;14:1380-5.
[65] Adhikari RS, Aste N, Manfren M. Optimization concepts in district energy design and management – A case study. Energy Procedia. 2012;14:1386-91.
[66] Kraning M, Chu E, Lavaei J, Boyd S. Dynamic Network Energy Management via Proximal Message Passing. Found Trends Optim. 2014;1(2):73-126.
[67] Orehounig K, Evins R, Dorer V. Integration of decentralized energy systems in neighbourhoods using the energy hub approach. Applied Energy. 2015;154:277-89.
[68] Mancarella P. MES (multi-energy systems): An overview of concepts and evaluation models. Energy. 2014;65:1-17.

[69] Jentsch MF, James PAB, Bourikas L, Bahaj AS. Transforming existing weather data for worldwide locations to enable energy and building performance simulation under future climates. Renewable Energy. 2013;55(Supplement C):514-24.

[70] Yang Z, Becerik-Gerber B. A model calibration framework for simultaneous multi-level building energy simulation. Applied Energy. 2015;149:415-31.

[71] PHPP. The energy balance and Passive House planning tool.

[72] Lehmann B, Gyalistras D, Gwerder M, Wirth K, Carl S. Intermediate complexity model for Model Predictive Control of Integrated Room Automation. Energy and Buildings. 2013;58:250-62.

[73] Buonomano A, Montanaro U, Palombo A, Santini S. Dynamic building energy performance analysis: A new adaptive control strategy for stringent thermohygrometric indoor air requirements. Applied Energy. 2016;163:361-86.

[74] Strachan P, Svehla K, Heusler I, Kersken M. Whole model empirical validation on a full-scale building. Journal of Building Performance Simulation. 2016;9(4):331-50.

[75] Antony J. Design of Experiments for Engineers and Scientists: Elsevier Science, 2014.

[76] Tronchin L, Manfren M. Multi-scale Analysis and Optimization of Building Energy Performance – Lessons Learned from Case Studies. Procedia Engineering. 2015;118:563-72.

[77] ISO 16346:2013, Energy performance of buildings — Assessment of overall energy performance. 2013.

[78] Aste N, Pero CD, Fattore M, Mazzon M. Investigating on electric consumptions for residential buildings ventilation in different Italian climates. Conference Investigating on electric consumptions for residential buildings ventilation in different Italian climates. p. 305-10.

[79] Baetens R, De Coninck R, Van Roy J, Verbruggen B, Driesen J, Helsen L, et al. Assessing electrical bottlenecks at feeder level for residential net zero-energy buildings by integrated system simulation. Applied Energy. 2012;96:74-83.

[80] De Coninck R, Magnusson F, Åkesson J, Helsen L. Toolbox for development and validation of grey-box building models for forecasting and control. Journal of Building Performance Simulation. 2016;9(3):288-303.

[81] ISO/DIS 18523-2:2018 Energy performance of buildings - Schedule and condition of building, zone and room usage for energy calculation--Part 2: Residential buildings.

[82] Kurnitski J. Technical definition for nearly zero energy buildings. REHVA Journal, Technical (May). 2013:22-8.

[83] Frontini F, Manfren M, Tagliabue LC. A Case Study of Solar Technologies Adoption: Criteria for BIPV Integration in Sensitive Built Environment. Energy Procedia. 2012;30:1006-15.

[84] Voss K, Sartori I, Napolitano A, Geier S, Gonçalves H, Hall M, et al. Load matching and grid interaction of net zero energy buildings. Conference Load matching and grid interaction of net zero energy buildings.

[85] ISO/DIS 52000-1:2017, Energy performance of buildings — Overarching EPB assessment – Part 1: General framework and procedures (draft). 2017.

[86] Cipriano X, Gamboa G, Danov S, Mor G, Cipriano J. Developing indicators to improve energy action plans in municipalities: An accounting framework based on the fund-flow model. Sustainable Cities and Society. 2017;32:263-76.

[87] Aste N, Buzzetti M, Caputo P, Manfren M. Local energy efficiency programs: A monitoring methodology for heating systems. Sustainable Cities and Society. 2014;13:69-77.

[88] Server F, Kissock JK, Brown D, Mulqueen S. Estimating industrial building energy savings using inverse simulation. 2011.
[89] Abushakra B, Reddy A, Singh V. ASHRAE Research Project Report 1404-RP, Measurement, Modeling, Analysis and Reporting Protocols for Short-term M&V of Whole Building Energy Performance, Arizona State University, USA. 2012.
[90] Lin G, Claridge DE. A temperature-based approach to detect abnormal building energy consumption. Energy and Buildings. 2015;93:110-8.
[91] Mantesi E, Hopfe CJ, Cook MJ, Glass J, Strachan P. The modelling gap: Quantifying the discrepancy in the representation of thermal mass in building simulation. Building and Environment. 2018;131:74-98.
[92] Kicsiny R. Multiple linear regression based model for solar collectors. Solar Energy. 2014;110:496-506.
[93] Kicsiny R. Improved multiple linear regression based models for solar collectors. Renewable Energy. 2016;91:224-32.
[94] Oberkampf WL, Roy CJ. Verification and validation in scientific computing: Cambridge University Press, 2010.

This work was carried out within the research project n.201594LT3F which is funded by PRIN (Programmi di Ricerca Sci-entifica di Rilevante Interesse Nazionale) of the Italian Ministry ofEducation, University and Research.