Build Simul (2020) 13: 33–50
https://doi.org/10.1007/s12273-019-0559-8

Predictive capability testing and sensitivity analysis of a model for building energy efficiency

G. Kalogeras[1] (✉), S. Rastegarpour[2], C. Koulamas[1], A.P. Kalogeras[1], J. Casillas[3], L. Ferrarini[2]

1. Industrial Systems Institute, ATHENA Research and Innovation Center, Platani-Patras, 26504 Greece
2. Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Via Ponzio 34/5, 20133 Milano, Italy
3. Department of Computer Science and Artificial Intelligence, University of Granada, CITIC-UGR, E-18071, Granada, Spain

Research Article

Abstract

Building energy modelling presents a good tool for estimating building energy consumption. Different modelling approaches exist in literature comprising white-box/physical/calculation-based models, black-box/statistical/measurement-based models or hybrid models combining the former two. Our work presented in this paper deals with a calculation-based quasi-steady-state model for building energy consumption based on the ISO 13790 standard and its implementation in MATLAB/Octave. The model is also well compared to the ISO 52016 standard updating ISO 13790. The model predictive capability is confirmed against both EnergyPlus dynamic simulator results and calculation results of a commercially available relevant tool used as benchmarks. Machine learning techniques are applied to a large dataset of simulated data and a sensitivity analysis is presented narrowing down to the most influential model parameters.

Keywords

building energy modelling,
energy consumption assessment,
model sensitivity analysis,
model predictive capability

Article History

Received: 27 January 2019
Revised: 11 May 2019
Accepted: 16 May 2019

1 Introduction

Rise of CO_2 emissions attributed to buildings in a global scale has experienced an annual rise of nearly 1% since 2010 (Progress ITCE 2017). Growth in building floor area at a global annual percentage of 3% and increasing demand for building energy services have undermined progress in final energy per square meter that decreased at an annual pace of 1.3% in the period between 2010 and 2014. Enforcing energy efficiency policies and preventing energy inefficient building investments is the way to go so as to meet energy and greenhouse gas emission reduction goals.

Building energy consumption modelling is a tool that can help towards this direction. Different modelling approaches appear in literature. A classification of the prevalent approaches is presented in Koulamas et al. (2018). The classification is done according to different criteria, comprising detail of required information, relative hierarchical position of data inputs and building sector, and energy data acquisition approach.

With reference to the detail of required information white-box, black-box and grey-box modelling techniques are used (Foucquier et al. 2013). White-box modelling techniques start from physical laws and their description in the form of mathematical equations. In Rastegarpour et al. (2018) and Ferrarini and Mantovani (2013), authors addressed a control-oriented model of a building together with the heating system modelling to be used for the economic control and energy analysis in building sectors. Building physical models requires deep level of knowledge of the building details and characteristics, as well as a high level of expertise. Adaptation to different buildings requires a significant effort.

On the other hand, black-box techniques utilize large amounts of building data over a significant period of time, after having trained the model with a separate set of training data. Interpretability of the results is more difficult than in white-box models, yet they do not present a need for knowledge in detail of the building characteristics. Adaptation of the resulting models to different buildings is quite difficult.

E-mail: gkalogeras@isi.gr

Building Thermal, Lighting, and Acoustics Modeling

A combination of white-box and black-box modelling approaches produces grey-box modelling. They combine a less detailed physical description of a building with a set of training data over a period of time. In this context they require greater initial effort than black box models, while having better generalization and interpretability.

With reference to relative hierarchical position of data inputs and building sector criterion, approaches are classified as engineering, statistical and hybrid (Swan and Ugursal 2009). Engineering approaches are close to white-box or physical modelling approaches. Statistical approaches resemble black-box approaches. Finally, hybrid approaches combine the characteristics of the other two and match grey-box modelling approaches.

When coming to Energy Data Acquisition approach criterion, modelling approaches are discerned into calculation-based, measurement-based and hybrid (Wang et al. 2012). Again calculation-based approaches resemble white-box or physical or engineering approaches and comprise steady-state method or dynamic simulation. Measurement-based approaches are close to black-box or statistical approaches and comprise bill based or monitoring based techniques. Hybrid approaches comprise calibrated simulation and dynamic inverse modelling. Classification of building modelling approaches is presented in Fig. 1.

Work presented in this paper details a calculation based quasi-steady-state model (QSSM) for building energy consumption and presents its model predictive capability testing and sensitivity analysis. We remark that the sensitivity analysis in the field of building energy performance analysis has been already studied (Tian 2013; Sanchez et al. 2014; Ballarini and Corrado 2012; Rastegarpour et al. 2017). Here in this paper, the validated QSSM model is used as the main data generator in order to drive the machine learning techniques for a sensitivity analysis of the building energy performance.

The rest of the paper is structured as follows. Section 2 presents the quasi-steady-state model based on ISO 13790 standard and its implementation. Section 3 presents a comparison between our model implementation simulation results and ISO 52016 standard. Section 4 presents the model predictive capability testing against EnergyPlus dynamic simulation results used as a benchmark and secondly against results obtained from a commercial software. Section 5 presents a sensitivity analysis of the validated model utilizing to this end the application of machine learning techniques. Finally, conclusions are drawn in Section 6.

2 Calculation based quasi-steady-state model

The International Standard EN13790 2nd edition, 2008-03-01 (Energy performance of buildings—Calculation of energy use for heating and cooling) (ISO 13790 2008) presents a coherent set of calculation methods, at different levels of detail, for the energy use for the space heating and cooling of a building, namely a fully prescribed simple hourly dynamic calculation method, a detailed hourly dynamic calculation method and a monthly quasi-steady-state calculation method. All the three methods presented in the standard can be used for the assessment of the annual energy needed for space heating and cooling in residential or non-residential buildings divided into zones. The standard has been technically

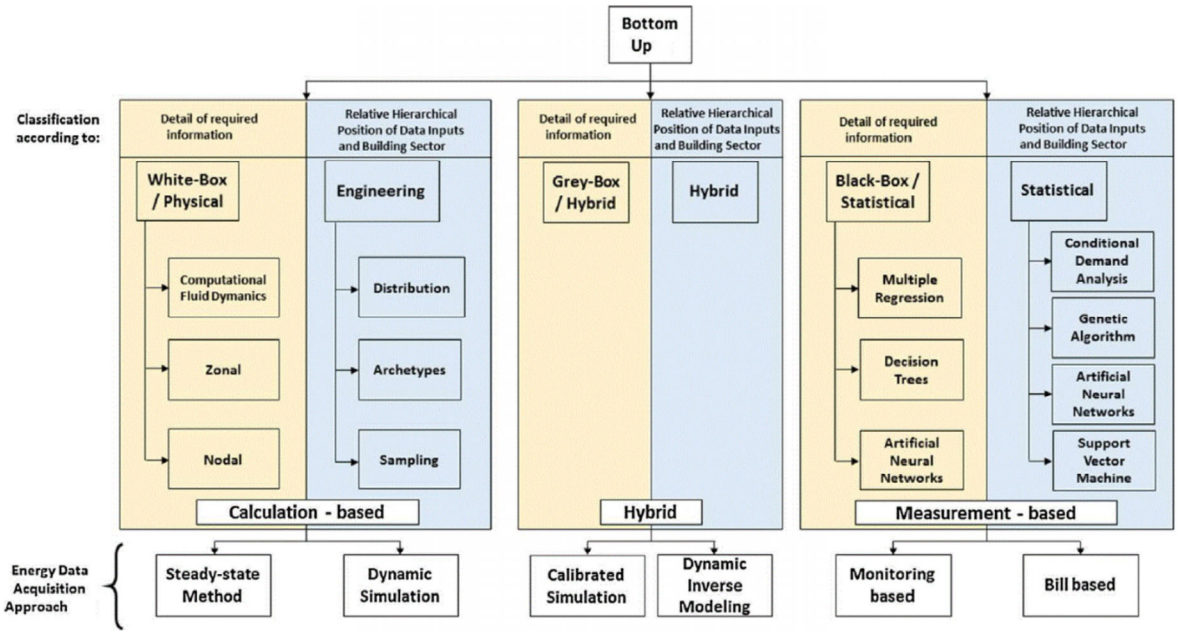

Fig. 1 Classification of building modelling approaches (Koulamas et al. 2018)

revised and superseded by ISO 52016-1 (ISO 52016-1 2017) in 2017, which presents some differences in the monthly calculation method.

The calculation model presented in this paper was created in accordance with the procedures specified for the quasi-steady-state monthly method in ISO 13790. This method was selected as the one providing the best compromise between dynamic and steady-state methods (Moronis et al. 2017).

The goal of the process is to calculate the sensible energy need for heating and cooling per month. The building can be broken down into zones and calculation is performed for each zone independently. The calculation process is based on a monthly balance between heat gains and heat losses, determined in steady state conditions, where dynamic effects are also taken into account by introducing utilization factors (Van Dijk et al. 2005). The calculation of the energy includes: the calculation of the heat transfer by transmission and ventilation of the building zone when heated or cooled to constant internal temperature, the contribution of internal and solar heat gains to the building heat balance, the annual energy needs for heating and cooling to maintain the specified set-point temperatures in the building, and the annual energy needs for heating and cooling of the building, using input from the relevant system standards referred to in ISO 13790:2008 and more specifically in Annex A (ISO 13790 2008). The main equations used for the calculation of the energy need for heating or cooling are presented below:

$$Q_{\mathrm{nd(H)}} = Q_{\mathrm{ht(H)}} - \eta_{\mathrm{gn(H)}} Q_{\mathrm{gn(H)}}$$

$$Q_{\mathrm{nd(C)}} = Q_{\mathrm{gn(C)}} - \eta_{\mathrm{ls(C)}} Q_{\mathrm{ht(C)}}$$

where:

- $Q_{\mathrm{nd(H)}}$ and $Q_{\mathrm{nd(C)}}$ rep resent the monthly energy need for heating and cooling the building, respectively.
- $Q_{\mathrm{ht(H)}}$ and $Q_{\mathrm{ht(C)}}$ represent the total heat transfer for the heating and cooling mode, respectively. The total heat transfer of the building is divided into heat transfer by transmission and ventilation. Heat transfer by transmission is the transfer through all building elements in contact with external air, which is driven by the temperature difference between the inside and outside air of the building. Transfer by ventilation is the temperature-driven heat transfer due to ventilation, including air infiltration.
- $Q_{\mathrm{gn(H)}}$ and $Q_{\mathrm{gn(C)}}$ represent the total heat gains for the heating and cooling mode, respectively. Heat gains are divided into internal and solar heat gains. Internal heat gains represent the heat flow from any internal heat source like occupants, lighting, devices, recoverable losses etc. Solar heat gains represent the total heat flow due to solar irradiance through all external building elements.
- $\eta_{\mathrm{gn(H)}}$ is the gain utilization factor for heating.
- $\eta_{\mathrm{ls(C)}}$ is the loss utilization factor for cooling.

The two utilization factors contribute to the semi-dynamic behaviour of the model, by taking into account the heat balance ratio of the building. They are given by the following equations for heating:

$$\eta_{\mathrm{gn(H)}} = \frac{\left(1-\gamma_{\mathrm{H}}^{\alpha_{\mathrm{H}}}\right)}{\left(1-\gamma_{\mathrm{H}}^{\alpha_{\mathrm{H}}+1}\right)} \quad \text{if } \gamma_{\mathrm{H}} > 0 \text{ and } \gamma_{\mathrm{H}} \neq 1$$

$$\eta_{\mathrm{gn(H)}} = \frac{\alpha_{\mathrm{H}}}{\left(\alpha_{\mathrm{H}}+1\right)} \quad \text{if } \gamma_{\mathrm{H}} = 1$$

$$\eta_{\mathrm{gn(H)}} = \frac{1}{\gamma_{\mathrm{H}}} \quad \text{if } \gamma_{\mathrm{H}} < 0$$

and for cooling mode:

$$\eta_{\mathrm{ls(C)}} = \frac{\left(1-\gamma_{\mathrm{C}}^{-\alpha_{\mathrm{C}}}\right)}{\left(1-\gamma_{\mathrm{C}}^{-(\alpha_{\mathrm{C}}+1)}\right)} \quad \text{if } \gamma_{\mathrm{C}} > 0 \text{ and } \gamma_{\mathrm{C}} \neq 1$$

$$\eta_{\mathrm{ls(C)}} = \frac{\alpha_{\mathrm{C}}}{\left(\alpha_{\mathrm{C}}+1\right)} \quad \text{if } \gamma_{\mathrm{C}} = 1$$

$$\eta_{\mathrm{ls(C)}} = 1 \quad \text{if } \gamma_{\mathrm{C}} < 0$$

In the sets of equations above the terms γ_{H} and γ_{C} represent the heat balance for heating and cooling modes, respectively, as:

$$\gamma_{\mathrm{H}} = Q_{\mathrm{gn(H)}} / Q_{\mathrm{ht(H)}}$$

$$\gamma_{\mathrm{C}} = Q_{\mathrm{ht(C)}} / Q_{\mathrm{gn(C)}}$$

and the terms α_{H} and α_{C} are dimensionless numerical parameters that depend on the time constant of the building zone $\tau_{\mathrm{(H/C)}}$. The dimensionless parameters are given by the following equations:

$$\alpha_{\mathrm{(H/C)}} = 1 + \tau_{\mathrm{(H/C)}} / 15$$

A more detailed analysis of the basic calculation steps is presented in Moronis et al. (2017) and Kalogeras et al. (2018a).

For model implementation purposes the calculation process was broken down into separate procedures for heating and cooling. All the calculations have been implemented in MATLAB/Octave environment. User input and other necessary data for calculations, like weather data and default values of some parameters are imported in the form of MS Excel files. QSSM verification and validation has been performed as presented in Kalogeras et al. (2018b).

3 Model comparison with ISO52016

The ISO 52016-1:2017 standard (Energy performance of

buildings—Energy needs for heating and cooling, internal temperatures and sensible and latent heat loads—Part 1: Calculation procedures) is part of the set of EPB standards that define the methodology for assessing the energy performance of buildings. The calculation methods can be used for entire or part of residential or non-residential buildings and they have been developed for the calculation of the basic energy loads and needs of the building. The standard is applicable to buildings at the design stage, to new buildings after construction and to existing buildings in the use phase.

ISO 52016 proposes some modifications to the calculation procedure, when compared to ISO 13790. The first and most important difference between the two standards is about intermittent heating and cooling. Especially when referring to non-continuous heating the calculations are completely different in philosophy, since ISO 13790 calculates a single coefficient which is applied to the final value of heating need, whereas, ISO 52016 re-calculates the internal set point temperature which then affects all intermediate calculations. More specifically ISO 52016 uses the following formula to calculate the internal setpoint temperature, which is then used in every step of the calculation procedure in place of the actual setpoint temperature:

$$\theta_{\mathrm{int(H)}} = \alpha_{\mathrm{H,red}}(\theta_{\mathrm{set(H)}} - \theta_{\mathrm{e}}) + \theta_{\mathrm{e}}$$

where $\theta_{\mathrm{set(H)}}$ is the actual setpoint temperature, θ_{e} is the monthly mean air temperature of the external environment and $\alpha_{\mathrm{H,red}}$ is the reduction factor introduced by intermittency. This is calculated taking into account the part of time for which the heating system works with a reduced setpoint or is shut down and the temperature difference between regular working conditions and reduced setpoint or switch-off conditions.

For intermittent cooling the new standard considers intermittency only when there is a period of reduced need or switch off greater than 48 hours during the week. If this condition is fulfilled then a reduction factor is applied to the total energy need for cooling as presented below:

$$Q_{\mathrm{nd(C)}} = \alpha_{\mathrm{Cred}}\left(Q_{\mathrm{gn(C)}} - \eta_{\mathrm{ls(C)}} Q_{\mathrm{ht(C)}}\right)$$

where α_{Cred} is calculated taking into account the time fraction of the week where intermittency is applied.

Another difference has to do with controlling the months that actually need energy for heating or cooling. ISO 52016 adds more accurate controls that zero the need for heating or cooling when certain specifications are met. More specifically the energy need for heating in ISO52016 is set to zero when the term γ_{H} has a value greater than 2. For the cooling need the same thing is introduced when the corresponding factor γ_{C} is less than 0.5.

The two standards also differ in the calculation of heat transfer by transmission. In the new standard the heat transfer by elements in contact with the ground is computed separately according to a complex calculation presented in EN ISO 13370 (ISO 13370 2017). For simplification purposes and for the purpose of the predictive capacity testing dealt with in this paper, table values derived from the mentioned standard which are included in Greek national regulations have been used to simulate the difference that this change creates. Results of this simulation are given in the following section.

Another minor difference presented in ISO 52016 is a correction in the gain utilization factor for heating, where an extra control statement is inserted. The differences concerning the heat gain utilization factor are presented below.

$$\eta_{\mathrm{gn(H)}} = \frac{1}{\gamma_{\mathrm{H}}} \quad \text{if } \gamma_{\mathrm{H}} <= 0 \text{ and } Q_{\mathrm{gn(H)}} > 0$$

$$\eta_{\mathrm{gn(H)}} = 1 \quad \text{if } \gamma_{\mathrm{H}} <= 0 \text{ and } Q_{\mathrm{gn(H)}} <= 0$$

Finally, ISO 52016 introduces calculations about the energy need for humidification and dehumidification of a space. For simplification purposes this has not been taken into account in the ISO 52016 compliant version of the QSSM.

In order to perform a mapping of implemented QSSM to ISO 52016 and test its predictive capacity, simulation results have been utilized. To this end a set of interventions has been performed towards having an ISO 52016 compliant version of the model and this latter version simulation results are compared with the original model results.

3.1 Comparison of simulation results

Both models have been used to simulate the energy behaviour of buildings and their results have been compared to each other. To this end a real hospital test case and the single volume building presented in ISO 52016 verification method have been modelled in both approaches. Discussion of the simulation results follows.

The first set of simulations was performed using a real hospital in Sicily, Italy. The hospital in question is an ambulatory health centre situated in one main building which spans a total of 10833,9 m^2 over 5 floors (including the basement and ground floor). The hospital consists of several areas with different usage ranging from operating rooms and medical facilities to kitchens and office spaces. Existing hospital data have been complemented by an audit performed on site. The heating and cooling needs from both models are presented and compared in Fig. 2 and Fig. 3 respectively. It can be clearly seen that the models wield almost identical results.

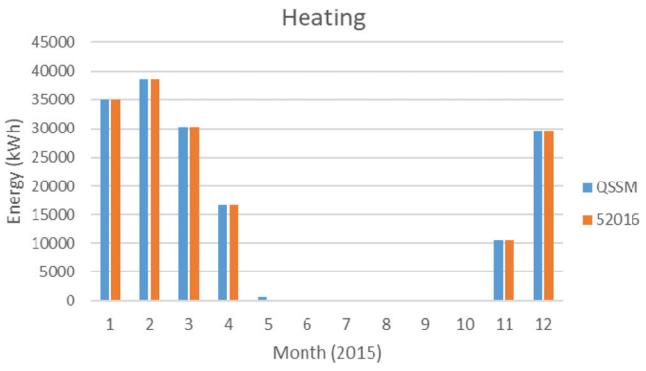

Fig. 2 Sicilian hospital heating need

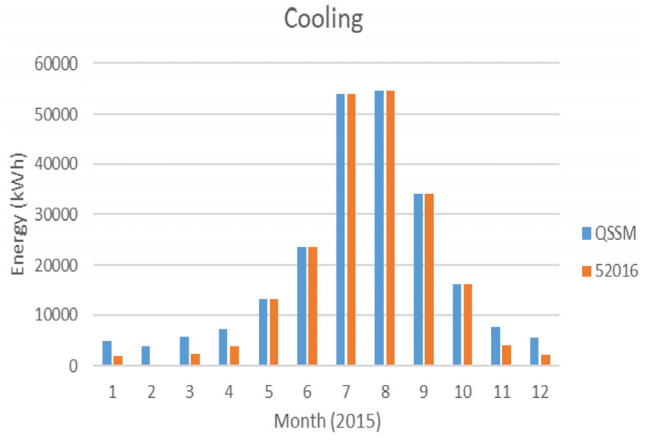

Fig. 3 Sicilian hospital cooling need

To point out the difference between the two standards when the building is under intermittent heating and cooling, another set of simulations has been performed with the following assumptions:

- Heating is on for 12 hours every day except for weekends when in stays off
- Cooling is continuous all weekdays and completely off on weekends

The comparison of results from the two standards is presented in Fig. 4 and Fig. 5. ISO 52016 estimates higher annual heating needs with this intermittency plan by 45% when for cooling the annual error stands at 14% due to the different calculation procedure presented in the new standard.

ISO 52016 also introduces a separate calculation of heat transfer through elements in contact with the ground. This is calculated by altering the heat transmittance of the element in contact with the ground, according to EN ISO 13370. It must be noted that this effect is important only when there is a large part of the building envelope in contact with the ground. If this specification is not met then the effect diminishes. Results for the Sicilian hospital using the updated calculation procedure are presented in Fig. 6 and Fig. 7.

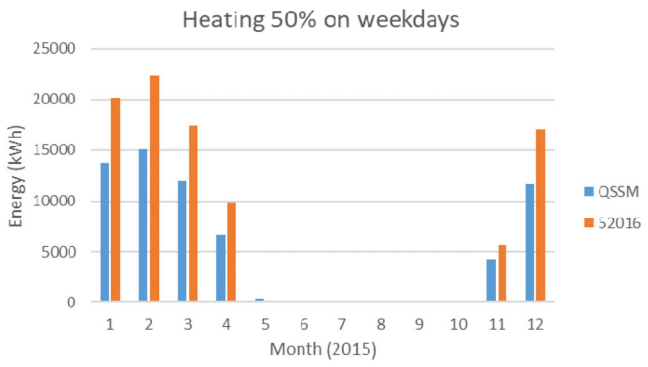

Fig. 4 Heating needs for Sicilian hospital (50% intermittency on weekdays, heating off on weekends)

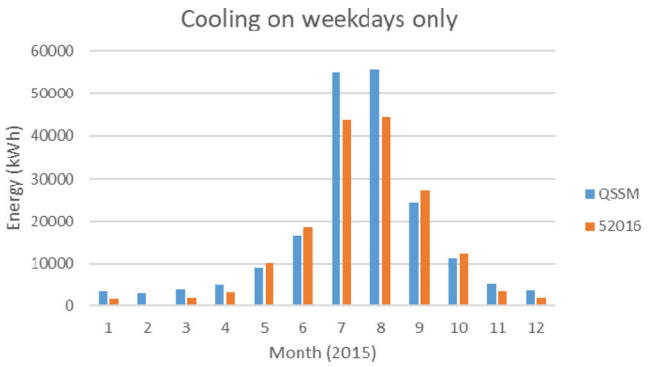

Fig. 5 Cooling need for Sicilian hospital (cooling off on weekends)

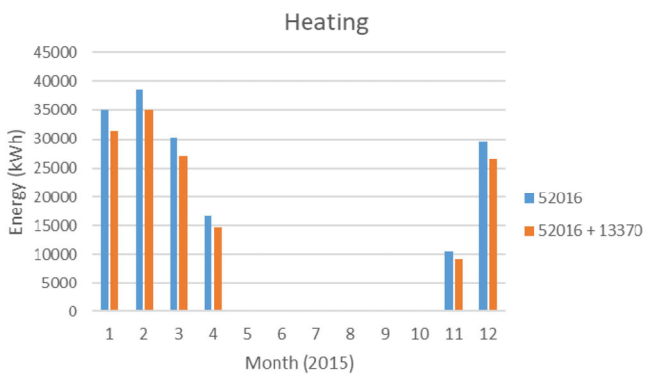

Fig. 6 Heating need with corrected ground factor

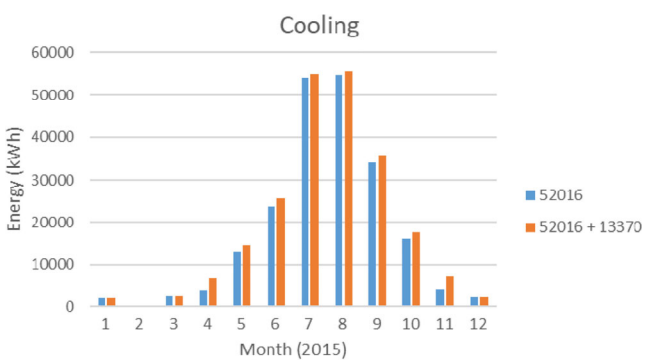

Fig. 7 Cooling need with corrected ground factor

The total annual error for heating introduced when the ground factor is corrected is about 10% when for heating it stands at lower than 7%. These figures highlight the effect of the correction of heat transfer to the ground when combined with the fact that the hospital has almost 30% of its envelope in contact with the ground. These results are obtained when correcting the ground factor for the baseline simulation assumptions, without taking into account then intermittency schedule presented above.

3.2 Comparison against ISO 52016 test cases

Further to the simulation results of a real hospital, test cases related to the single volume space presented in ISO 52016 have been utilized. This single volume space is a 48 m^2 room with two 6 m^2 windows facing south. The height of the room is 2.7 m and all elements are considered as external, i.e. in contact with external air. This also applies to the floor which is decoupled from the ground by increasing the thermal insulation. The geometry of the test room is presented in Fig. 8.

Cases BESTEST 600 and 640 have been used. These cases are described in ANSI/ASHRAE 140 (ASHRAE 2017). Test case 600 refers to a lightweight building construction. The thermophysical properties of all elements are presented in detail by the standard. The internal heat capacity of the zone has been analytically calculated using values from Tables 23 and 24 of the standard. Continuous heating and cooling is used at 20 °C and 27 °C respectively.

Test case 640, which interferes with heating intermittency, is also simulated. In particular it has heating at 20 °C from 07:00 to 23:00 each day and has a night time setback temperature of 10 °C. Cooling is continuous. Just like case 600, case 640 is a lightweight construction. All other properties are the same as in case 600.

In order to provide a verification of the QSSM, simulation results related to the above cases are compared with simulation results associated with the model ISO 52016 version. Furthermore, single volume space has been implemented as a model in EnergyPlus dynamic simulator (EnergyPlus 2015) and relevant simulation results were derived. The single volume created in EnergyPlus is shown in Fig. 8. The results of the simulations of the QSSM, the ISO 52016 model version and the EnergyPlus simulator model are presented in Figs. 9–12.

A cumulative comparison of the results of the two test cases for both ISO 52016 and EnergyPlus simulations is shown in Table 1. The metric used is the percentage difference

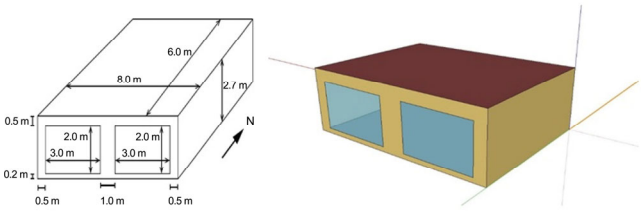

Fig. 8 ISO 52016 test volume geometry and model in EnergyPlus simulator

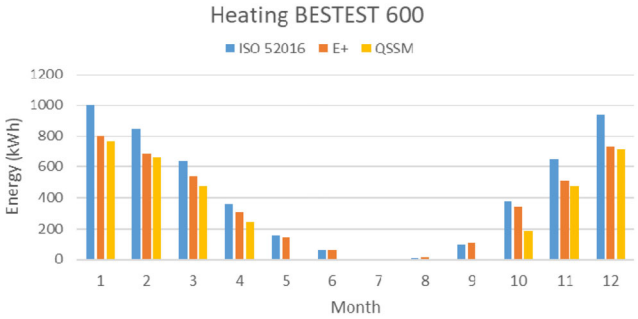

Fig. 9 Heating need for BESTEST 600

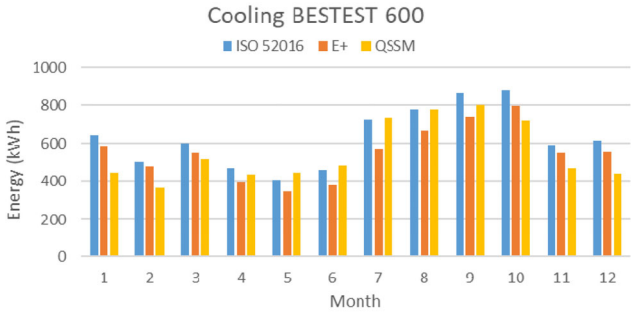

Fig. 10 Cooling need for BESTEST 600

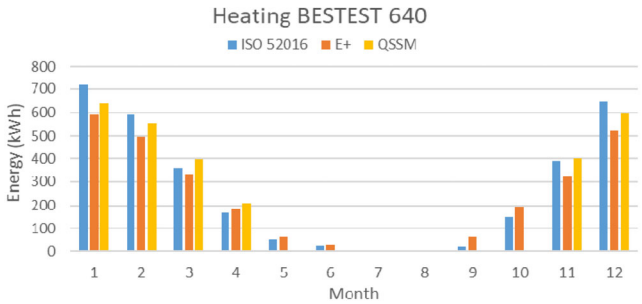

Fig. 11 Heating need for BESTEST 640

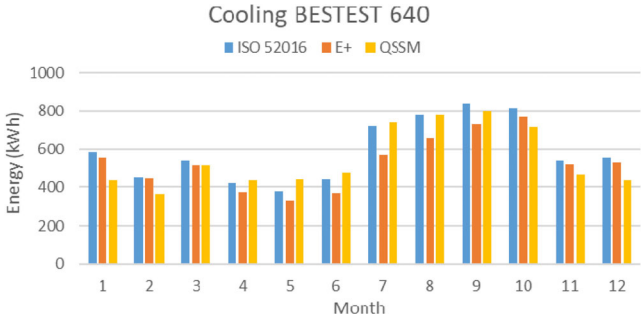

Fig. 12 Cooling need for BESTEST 640

Table 1 Percentage difference of simulations with reference cases

	Percentage difference			
	EnergyPlus		MATLAB/Octave	
ISO 52016 test case	Heating	Cooling	Heating	Cooling
BESTEST 600	−7.0%	−7.2%	−12.7%	−7.1%
BESTEST 640	−3.0%	−6.8%	−3.0%	−4.4%

of the heating or cooling need from the total energy need of the reference case. It may be seen that percentage difference with reference to the reference cases is well below the threshold of 15%, which is defined as acceptable in EN 15265:2007 (EN 15265 2007), further verifying the predictive capacity of the model.

4 Quasi-steady-state model predictive capacity testing

In order to test the predictive capacity of the QSSM, two different paths have been followed. One is related to comparison with simulations performed utilizing EnergyPlus dynamic simulator, while the other is associated with comparison with the results of a commercial software tool being based on the ISO 13790 standard.

4.1 Predictive capacity testing against EnergyPlus simulation results

The predictive capacity of the implemented model is quantitatively assessed against the EnergyPlus dynamic simulation software to better evaluate accuracy and reliability of the QSSM. To do so, the EnergyPlus model of a single volume has been developed, as depicted in Fig. 13. The figure also shows the single volume characteristics. It consists of a room with a single external wall and window. The thermo-physical characteristics of the building elements have been taken from the former EN15265 standard.

Ten different test cases are considered, so as to better evaluate the predictability of the implemented model in presence of different uncertainties. Different building characteristics and weather conditions are altered in each test case. The different test cases and applicable changes may be seen in Table 2. The input variables selected for changing are conductivity, λ (W/(m·K)), which directly affects the thermal transmittance of the element, and density of the material, ρ (kg/m^3), which in turn affects the heat capacity of the element. These changes are only performed on the insulating layer of the element. Other variables are external conditions, which are altered by selecting different location and orientation of the building, which is changed by 90 degrees to cover all possible major orientations.

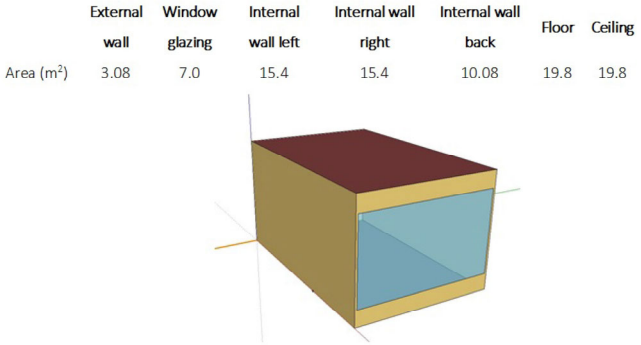

Fig. 13 EnergyPlus model of the EN15265 validation volume and building characteristics

The test cases described above have been simulated in EnergyPlus. Monthly values of external temperature were provided by EnergyPlus, while monthly solar irradiance values were calculated using the hourly values from EnergyPlus weather files.

Simulated data and comparison with results from EnergyPlus can be seen in Table 3 and Fig. 14. Comparison is made using the metric provided in ISO13790, which divides the difference in energy need for heating or cooling with the total energy need of the building.

From the results presented above it may be derived that the average of absolute error for total energy consumption, i.e. heating and cooling together, is about 12.84% over 10 test cases, which decreases to 11.42% for relative error for total energy consumption. More precisely, the average error for cooling is considerably smaller than that for heating. Clearly, the sample cases that have been selected although limited in number, still cover a wide range of typical buildings.

Table 2 Description of test cases

Test case	Conductivity (affecting U value) (W/(m·K))	Density (affecting C value) (kg/m^3)	Weather	Orientation of window
1	0.04	30	TRAPPES - France	West
2	0.04	30	TRAPPES - France	East
3	0.04	30	TRAPPES - France	North
4	0.04	30	TRAPPES - France	South
5	1.5×0.04	30	TRAPPES - France	West
6	0.5×0.04	30	TRAPPES - France	West
7	0.04	1.5×30	TRAPPES - France	West
8	0.04	0.5×30	TRAPPES - France	West
9	0.04	30	PHOENIX – USA	West
10	0.04	30	DULUTH – USA	West

Table 3 Simulated data and comparison with results from EnergyPlus

Test case ID	Month	1	2	3	4	5	6	7	8	9	10	11	12	**Total**	**E+**	**Error**
1	Heating	204.26	143.49	105.14	73.73	16.32	0.00	0.00	0.00	9.43	69.89	157.90	189.47	**969.63**	**761.04**	**22.37%**
	Cooling	0.00	0.00	0.00	2.99	18.10	42.21	80.48	37.54	11.71	0.00	0.00	0.00	**193.03**	**171.47**	**2.31%**
2	Heating	193.00	130.82	73.77	59.74	12.47	0.00	0.00	0.00	6.85	63.62	151.36	184.64	**876.27**	**677.08**	**20.07%**
	Cooling	0.00	0.00	7.60	6.98	31.28	69.17	129.36	64.01	22.17	0.00	0.00	0.00	**330.57**	**315.61**	**1.51%**
3	Heating	205.70	145.88	116.82	78.48	18.35	0.00	0.00	0.00	10.47	73.31	159.70	190.78	**999.49**	**815.56**	**19.84%**
	Cooling	0.00	0.00	0.00	0.00	13.82	33.65	65.86	29.93	9.26	0.00	0.00	0.00	**152.52**	**111.64**	**4.41%**
4	Heating	183.63	121.31	70.96	61.70	14.42	0.00	0.00	0.00	0.00	54.05	142.27	174.03	**822.37**	**563.32**	**33.49%**
	Cooling	0.00	0.00	8.78	6.20	23.55	47.48	95.54	55.31	25.60	3.50	0.00	0.00	**265.96**	**210.16**	**7.21%**
5	Heating	209.16	147.35	108.59	76.46	17.26	0.00	0.00	0.00	10.03	72.41	161.98	194.11	**997.35**	**790.76**	**21.56%**
	Cooling	0.00	0.00	0.00	2.88	17.52	41.22	79.34	36.59	11.28	0.00	0.00	0.00	**188.83**	**167.63**	**2.21%**
6	Heating	194.84	136.10	98.54	68.52	14.57	0.00	0.00	0.00	8.33	65.07	150.06	180.58	**916.61**	**722.09**	**21.63%**
	Cooling	0.00	0.00	0.00	3.23	19.29	44.21	82.71	39.45	12.61	0.00	0.00	0.00	**201.50**	**177.02**	**2.72%**
7	Heating	204.26	143.49	105.14	73.73	16.32	0.00	0.00	0.00	9.43	69.89	157.90	189.47	**969.63**	**761.19**	**22.34%**
	Cooling	0.00	0.00	0.00	2.99	18.10	42.21	80.48	37.54	11.71	0.00	0.00	0.00	**193.03**	**171.71**	**2.29%**
8	Heating	204.26	143.49	105.14	73.73	16.32	0.00	0.00	0.00	9.43	69.89	157.90	189.47	**969.63**	**761.2**	**22.34%**
	Cooling	0.00	0.00	0.00	2.99	18.10	42.21	80.48	37.54	11.71	0.00	0.00	0.00	**193.03**	**171.7**	**2.29%**
9	Heating	36.06	22.24	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	45.59	**103.89**	**47.9**	**2.72%**
	Cooling	3.71	8.70	74.29	147.36	247.91	347.85	374.27	343.22	261.10	124.99	26.52	2.68	**1962.60**	**2011.53**	**−2.38%**
10	Heating	466.02	371.65	276.41	134.46	41.84	0.00	0.00	0.00	31.27	124.81	268.37	435.32	**2150.16**	**1696.99**	**22.49%**
	Cooling	0.00	0.00	0.00	0.00	13.54	60.00	94.95	54.74	7.86	0.00	0.00	0.00	**231.09**	**317.58**	**−4.29%**

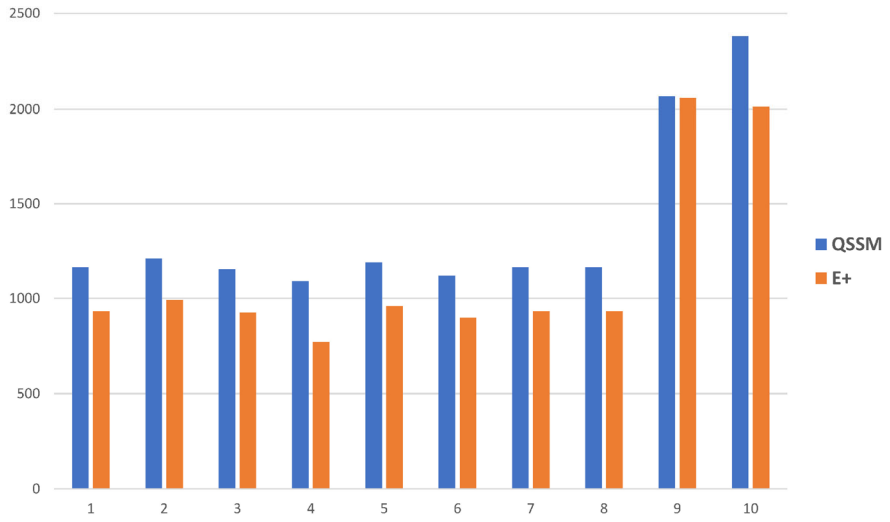

Fig. 14 Total energy need (in kWh) for simulated cases in QSSM and EnergyPlus

The results are definitely satisfactory and line with the expectations. So, in conclusion the predictive capacity of the model is ascertained from the comparative analysis with EnergyPlus simulation package.

4.2 Predictive capacity testing against Edilclima calculation results

In an attempt to test the predictive capability of the implemented QSSM, its simulation results are compared against the results of an available commercial software. The software selected for the testing is Edilclima EC 700 (Edilclima EC 700), version 8.17.49. The software is based on the ISO 13790 standard with a number of enhancements and some modifications, which are presented in the following paragraphs.

Edilclima is a commercial software commonly used in Italy, based on the Italian norm UNI TS 11300 (Le Norme

UNI TS 11300). Since the software is enhanced to calculate other parameters, such as domestic hot water and generation/distribution of energy, only computations related to heating and cooling needs of the hospital (first part of the UNI TS 11300) have been considered.

The two models have been compared using as a test case a private clinic located in Campania, southern Italy, built in 1973, composed of 8 floors, 5 of which partially underground, with a capacity of 108 beds.

The building was modelled using the quasi-steady-state-model implementation and simulated using as QSSM input exactly the same data as the commercial software. Results of the QSSM simulation and comparison to EdilClima calculation results are presented below.

EdilClima software defines a heating season and cooling days to avoid overestimating the total energy consumption of the building. The heating season for the area in question is set by Italian regulations from November 15th to March 31st. The QSSM has been updated to address this issue. The output of the simulation after the adjustment of the QSSM to the same heating season is compared to the results of the commercial software in Fig. 15.

It can be clearly seen that the calculated values for heating are very close for both approaches, with the biggest error standing at 8.5% for the month March.

The cooling need calculations made following the same approach are presented in Fig. 16. QSSM has been modified according to Corrado and Fabrizio (2007) to yield these results. Furthermore, it must be noted that due to EdilClima calculating cooling days and adjusting external conditions using linear interpolation, a procedure which is not part of the ISO 13790 standard, the results are not directly comparable, except for the months July and August where the complete month takes part in the calculation. For these two months the relative error between the two calculations is 5.13% for July and 6.06% for August.

When looking at the comparison of the predictive capacity of the QSSM against EnergyPlus and Edilclima

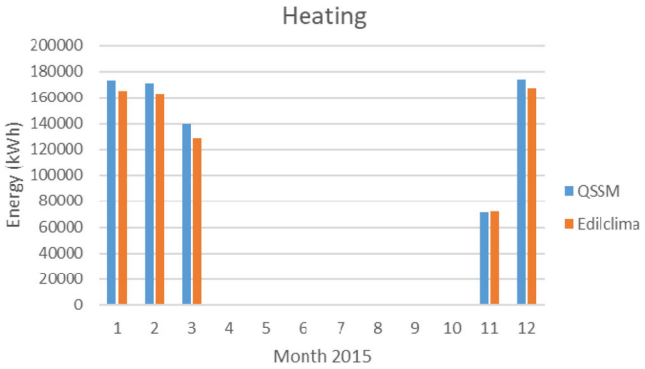

Fig. 15 Heating need comparison between QSSM and Edilclima EC700

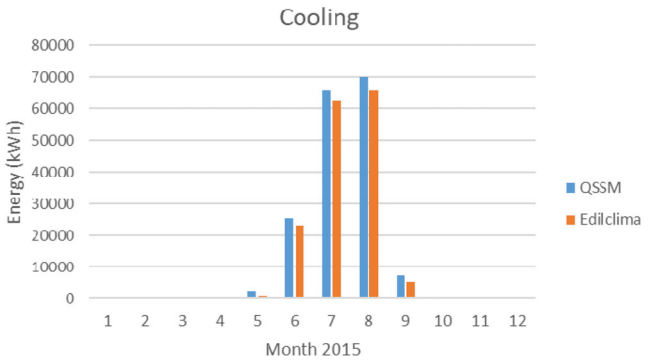

Fig. 16 Cooling need comparison between QSSM and Edilclima EC700

software it is obvious that QSSM results are better comparable against Edilclima. This can be explained by the fact that both QSSM and Edilclima software are static models based on ISO 13790, making them very similar, whereas EnergyPlus is a complex energy simulation program that performs dynamic calculations using a plethora of input parameters.

5 Machine learning

Following the validation of the QSSM presented in Section 4, machine learning techniques have been used in order to perform a sensitivity analysis of the main parameters of the validated model. The QSSM has been used to run a set of simulations. This set of simulated data is then used as input in machine learning techniques which perform data analysis to determine the influence of each variable on the energy consumption of the building. The building selected is the hospital presented in Section 3.1.

In order to get a more accurate picture of variable influence on the final energy consumption of the building an extended dataset has been selected. Selection of variables was based on two criteria. The first was variable importance in the model, with selection focusing on the most important variables in terms of effect on energy consumption. The second criterion was to select variables that can actually be physically manipulated through interventions in the building. Through this selection process the following set of variables was produced:

- *Heating setpoint* (°C): The desired internal temperature during heating period.
- *Cooling setpoint* (°C): The desired internal temperature during cooling period.
- *Occupancy*: This affects the internal heat gains from occupants (W/m^2).
- *Heat recovery system efficiency*: This affects the amount of heat recovered from exhaust ventilated air both in heating and cooling mode. The base building has no heat recovery system.

- *Lighting*: This variable affects the internal gains from lighting (W/m^2) as well as the electrical consumption of the building.
- *Ventilation intermittency*: This refers to the hours per day ratio that the ventilation system is turned on.
- *g window*: Total solar energy transmittance of the transparent part of a window.
- *Ventilation airflow rate* (m^3/s): The airflow rate through the ventilation system.
- *Uf opaque average* (W/(m^2·K)): Thermal transmittance of opaque building envelope elements (the indicator average refers to the average thermal transmittance value of all the different opaque building elements, e.g. concrete, brick wall etc., in contact with external air or the ground).
- *Uf transparent average* (W/(m^2·K)): Thermal transmittance of transparent building envelope elements (the indicator average refers to the average thermal transmittance value of all the different transparent building elements, e.g. window, glass door etc., in contact with external air).
- *Heating system efficiency, Cooling system efficiency*: These variables refer to the total efficiency of the heating and cooling system with the assumption that the building uses heat pumps for temperature control.
- *Heating system working hours per day, Heating system working hours per weekend, Cooling system working days in week ratio*: These variables affect the heating and cooling system intermittency, that is the hours per day that the system is turned on. The base building has continuous heating and cooling.

The value of each of the above variables is modified according to the following index vectors. In vectors with the indication *multiply* variable values are multiplied with each index, whereas when the indication *add* is present the index is added to the original variable value.

- Heating set point = [−2, 0, 2]; add
- Cooling set point = [−2, 0, 2]; add
- Occupancy = [0.5, 1, 1.5]; multiply
- Heat recovery system efficiency = [0, 0.5, 0.8]; add
- Lighting = [0.5, 1, 1.5]; multiply
- Ventilation intermittency = [0.5, 0.75, 1]; multiply
- g window [0.7, 1, 1.3]; multiply
- Ventilation airflow rate [0.6, 0.8, 1, 1.2, 1.4]; multiply
- Uf opaque average, Uf transparent average = [0.5, 0.75, 1, 1.25, 1.5]; multiply
- Heating system efficiency, Cooling system efficiency = [0.3, 0.6, 1, 1.4, 1.7]; multiply
- Heating system working hours per day, Heating system working hours per weekend, Cooling system working days in week ratio = [0.2, 0.4, 0.6, 0.8, 1]; multiply

The above selection of variables and their presented cardinality creates a total of 1,366,875 simulations for use in machine learning techniques. In the machine learning data analysis that follows, the above-mentioned variables are used as inputs versus the total electrical consumption of the building for a year as the variable to predict.

5.1 Exploratory data analysis

Firstly, an exploratory data analysis is performed to see the influence of each individual variable in the energy consumption. This can be accomplished by plotting the univariate probability density function (by kernel density estimation) of energy consumption with different values of the variables. To do so, we have obtained three probability density functions corresponding to the minimum, median and maximum value in each variable as depicted in Fig. 17. These graphs allow us to understand the univariate influence of each variable in the energy consumption. It is clear to see that "g window" has a low influence on the energy consumption (the distribution of energy consumption depending of the value that takes this variable are almost identical). This can be explained by the fact that for this particular building simulation the total glazed area is only 13.6% of the total building envelope, which diminishes the importance of the changes in the variable. On the contrary, "Heating system efficiency" and "Cooling system efficiency" are very relevant to determine the energy consumption. We can also observe the indirect relation with energy consumption in variables as "Cooling setpoint", "Heat recovery system efficiency" or "Heating system working hours", where lower values of these variables imply a higher energy consumption.

This exploratory analysis can be completed in two dimensions by visualizing with heatmaps the different combinations of variables. The matrix is shown in Fig. 18. The cells represent pairs of variables (column and row) and its heatmap shows the average energy consumption from the available dataset for each combination of values. For example, the heatmap of the cell of the first row, second column, represents as abscissa (columns of the heatmap) the three different values of "Cooling setpoint" (from 23.5 to 27.5) while the ordinate (rows of the heatmap) represents the "Heating setpoint" (from 19 to 23). In this example, we may observe that the highest energy consumption is produced, as expected, when the cooling setpoint is the minimum (23.5 °C) and the heating setpoint the maximum (23 °C) at the left-top corner of this heatmap. It is remarkable the great influence of "Heating system efficiency" (or its corresponding correlated variable "Cooling system efficiency").

Apart from the expected behaviour, we can discover other aspects from this bivariate visualization. For instance, "Heating working hours" has an important influence on energy consumption (the heatmaps of the rows and columns where this variable is used show a higher contrast of the colours) while "g window" has little influence on the energy

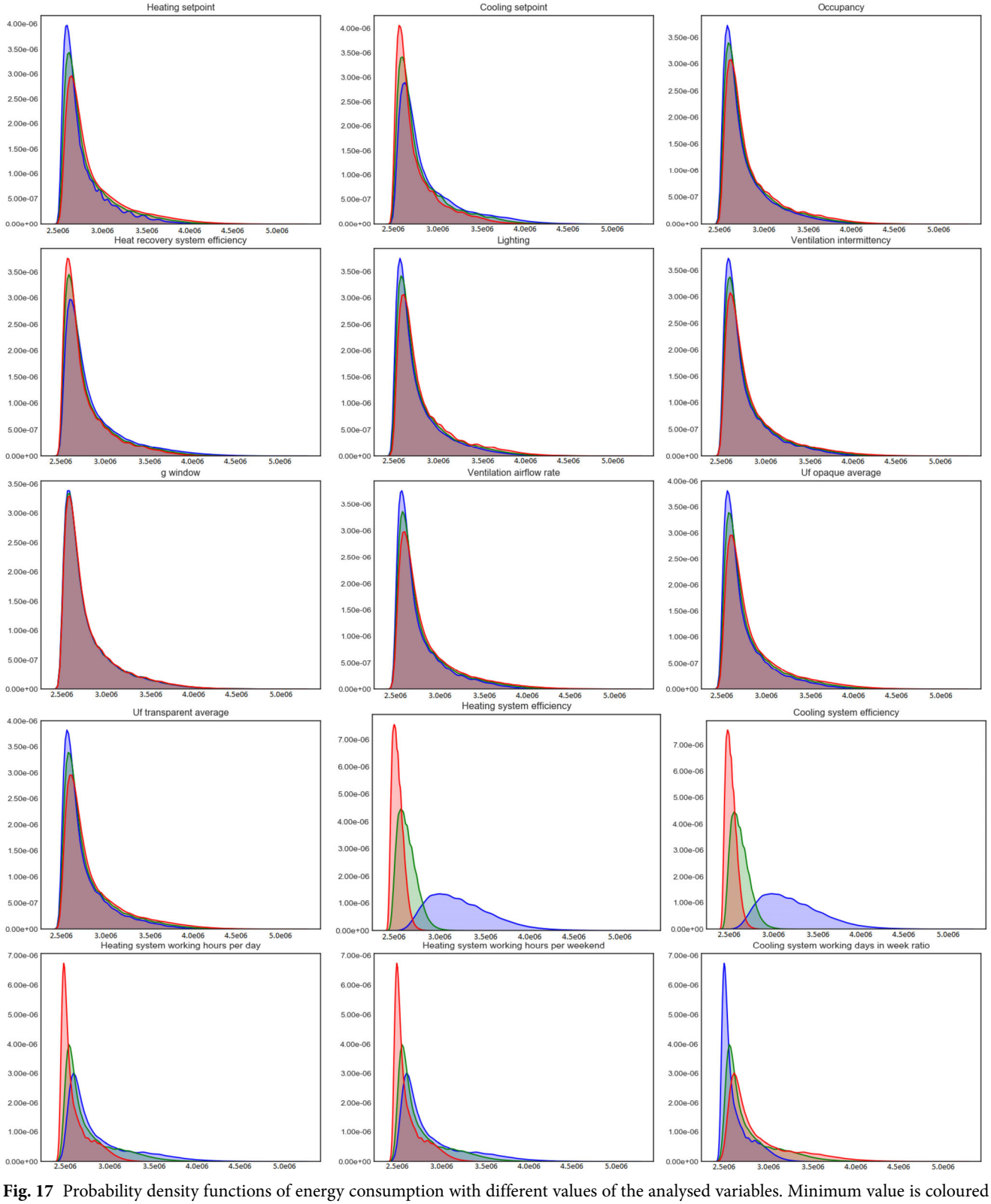

Fig. 17 Probability density functions of energy consumption with different values of the analysed variables. Minimum value is coloured in blue, median in green and maximum in red. In each plot, the abscissa axis represents the energy consumption and the ordinate represents the probability

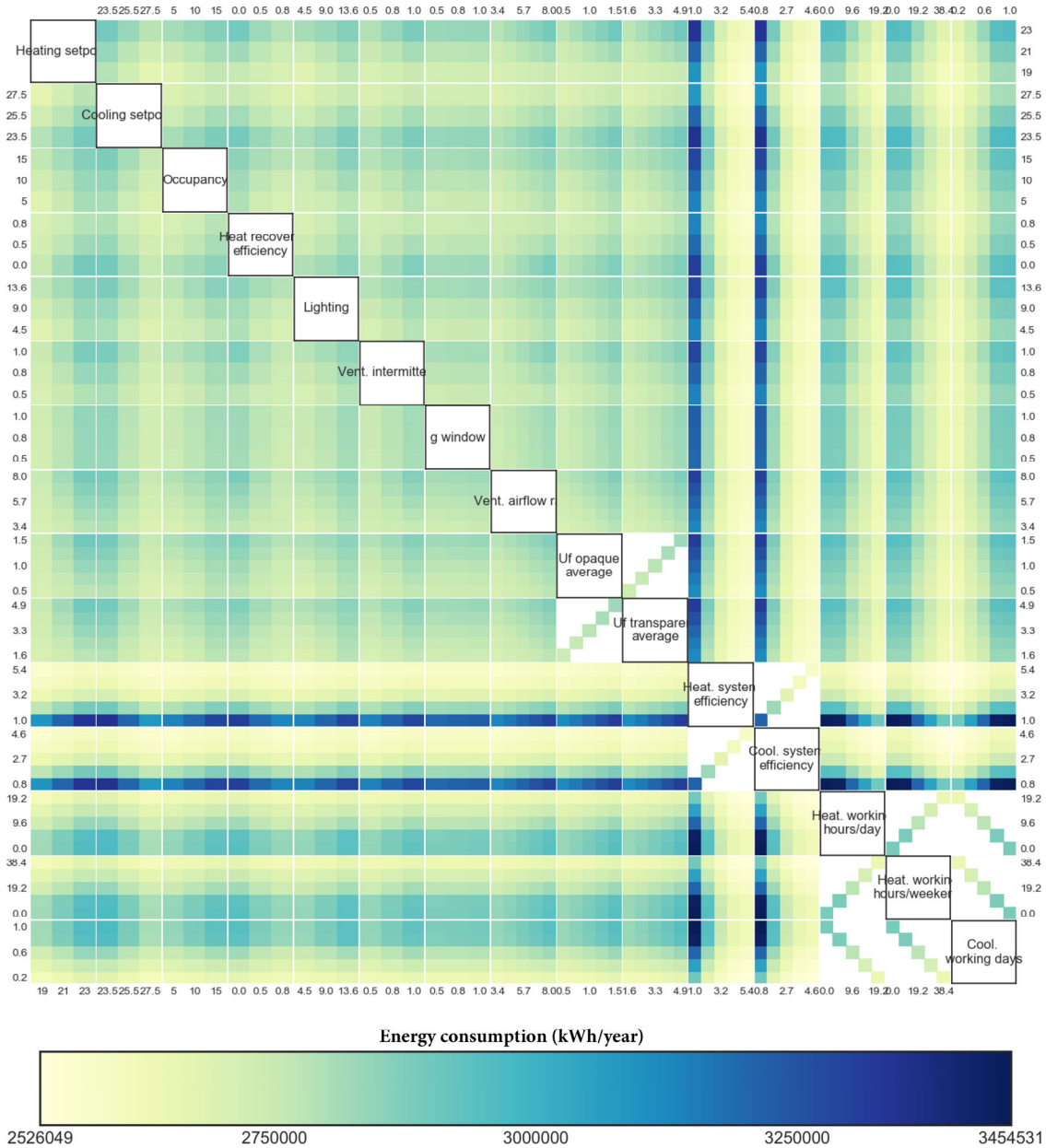

Fig. 18 Heat maps matrix of average energy consumption for the different combinations of pair of variables

consumption (its rows and columns does not generate significant colour contrasts).

5.2 Machine learning to obtain highly accurate models

We have applied the well-known, state-of-the-art, LightGBM (a gradient boosting decision tree) algorithm (Ke et al. 2017) on the available dataset. This algorithm, based on ensemble learning with decision trees as base repressors, is characterized by its great accuracy degree, so it will generate models that faithfully imitate the simulation used to generate the dataset. However, the interpretability of these complex models is very poor, so an analysis of the use of the variables in the models will be performed to get some idea about its behaviour. LightGBM is run with 300 models (estimators) and default values for the remaining parameters. Five-fold cross validation is used to obtain the test results, which are gathered in Table 4 where the averaged correlation coefficient (R^2) and root mean squared error (RMSE) are included. The prediction (real vs. predicted value of energy consumption in kWh per year) is also plotted in Fig. 19. As we can observe, the models obtained by LightGBM are highly accurate.

To get further insight about the importance of the variables in the models obtained by LightGBM, we have analysed the levels and frequency where the different variables appear in the 300 decision trees generated by the algorithm.

Table 4 Test error obtained by LightGBM with 5-fold cross validation

	R^2	RMSE (kWh/year)
LightGBM	0.99928	7742.9600

Figure 20 shows these results. From it, we can conclude that the most relevant variable (the one with the highest discriminant power) is "Heating system efficiency", which coincides with what was observed in the exploratory data analysis. Other variables as "Heating setpoint" and "Uf opaque average" have also a great relevance. Among the least relevant variables, we find "Cooling working days" and "g window". There are three variables that have a null importance because they are purely correlated with some other ones (as shown in Fig. 18) and so they can be inferred from them. The three pairs of correlated variables are "Heating system efficiency" and "Cooling system efficiency", "Heating working hours per day" and "Heating working hours per weekend"

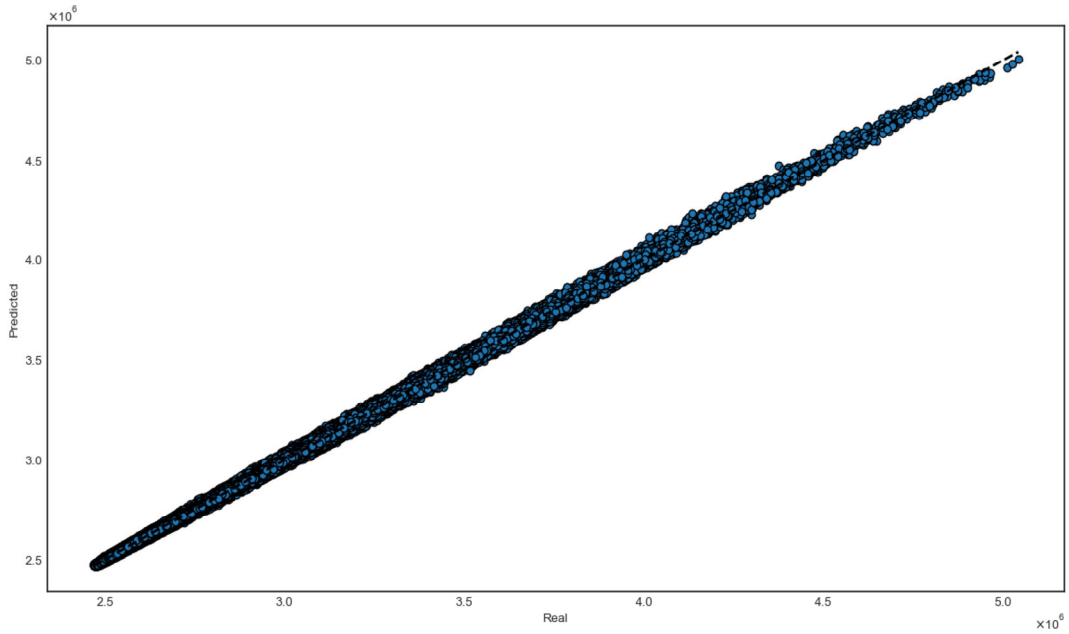

Fig. 19 Real vs. predicted values of energy consumption (kWh/year) with LightGBM

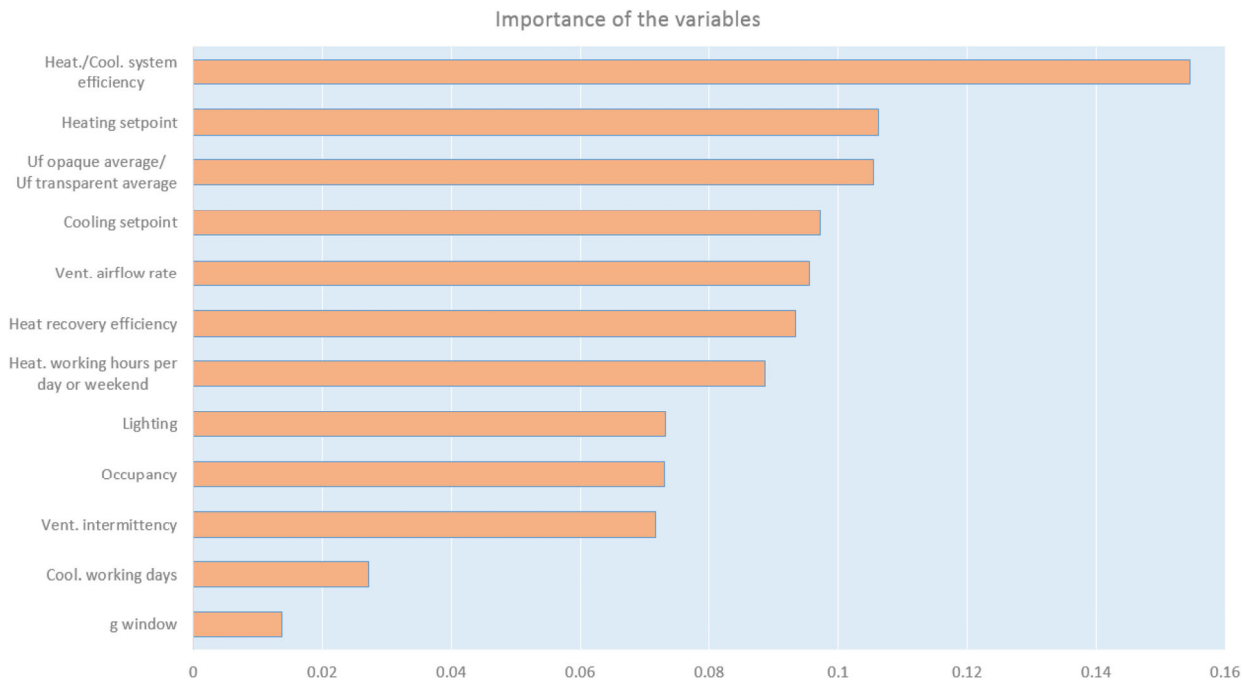

Fig. 20 Importance of the variables from the models generated by LightGBM

and finally "Uf opaque average" and "Uf transparent average".

5.3 Machine learning to obtain highly interpretable models

Since we are dealing with simulated data, the accuracy is not so relevant as to understand what exactly makes the energy consumption being higher or lower. Therefore, we will apply now another machine learning algorithm that ensures a good interpretability at the expense of a lower accuracy. To do so, decision trees where the leaves are lineal regressions have been generated. This approach is equivalent to a piecewise linear regression, so the model is highly legible. To ensure a not too complex decision tree, the depth has been limited by pruning when less than 7000 samples are obtained when splitting the node. M5P algorithm (Quinlan 1992; Wang and Witten 1997) is used in this case. In this algorithm, the model is expressed through a binary decision tree with linear regression functions in the terminal nodes (leaves) that establishes a relationship of the dependent variable with the independent ones. The tree is built by the divide-and-conquer method. The division criterion is based on the standard deviation of the values of the subset of data that reaches a node as a measure of the error value in that node, thus choosing the attribute that maximizes the expected reduction in the error.

Table 5 collects the obtained results (R^2, RMSE and root relative squared error (RRSE)) while Fig. 21 depicts the decision tree generated by M5P. The leaves of this tree are the different linear models (LM) generated in each situation. In these leaves, the values within brackets are the number of samples covered (over 1,366,875 data) and the approximate error. Note that almost half of the samples are covered by leaves 17, 18 and 19 with an approximate error lower than 7%. There are other areas, however, where the error is significantly greater, especially when the heating system efficiency is very low (subtree A). The overall error (R^2=0.9744 as shown in Table 5) is, in any case, acceptable considering its good interpretability. Table 6 includes the coefficient of the linear equations of the leaves. From these regressors, it is possible to compute the relative contribution of each variable in order to analyse its importance in each case. To do so, the variables have been previously normalized. Table 7 shows the results.

The last row of Table 7 contains the average relevance of each variable weighted by the number of samples of each leaf (positive value means that the variable influences directly on the energy consumption, while negative ones means it influences indirectly). The variables with a negative importance meaning that they indirectly influence on energy consumption (i.e., low values increase the energy consumption) are, as expected, "Cooling setpoint", "Heat recovery system efficiency", "Heating system efficiency" and "Heating system working hours per day".

We can observe that "Heating setpoint", "Cooling setpoint", "Heat recovery system efficiency" and "Uf opaque average" are the most important variables but now we are able to see the importance degree depending on which combination of values of the variables we are considering. For example, "Heating setpoint" gest its higher importance in LM1 (which represents the 2.7% of the total number of samples), i.e., when "Heating system efficiency" is less than or equal to 1.44, "Heating system working hours per day" is less than or equal to 7.2 and "Cooling setpoint" is less than or equal to 24.5. A second example is LM4 (1.3% of samples), where "Uf opaque average" gets the higher importance to determine the energy consumption when "Heating system efficiency" is less than or equal to 1.44, "Heating system working hours per day" is in the interval (7.2, 12] and "Heating setpoint" is greater than 22. However, these two linear regressions LM1 and LM4 have a high RRSE error (47% and 54%, respectively) so these conclusions cannot be considered statistically robust.

On the contrary, LM17, LM18 and LM19 concentrate 48% of samples and cover the cases either where "Heating system working hours per day" is greater than or equal to 14.4 and "Heating system efficiency" is greater than or equal to 3.2 (LM19), or where "Heating system working hours per day" is less than or equal to 9.6 and "Heating system efficiency" is greater than or equal to 4.32 (LM17 and LM18). In these three cases the error is low (RRSE below 7%) and the most influential variable is "Heating system efficiency". This conclusion coincides with what we got in Fig. 19 for LightGBM algorithm but now we can go beyond by understanding better when and how this variable influence on energy consumption.

Table 5 Test error obtained by M5 with 5-fold cross validation

	R^2	RMSE (kWh/year)	RRSE
M5P	0.97440	64,632.6939	22.47%

6 Conclusion

Having accurate estimations of the energy consumption of buildings is quite significant in taking decisions on their renovation or retrofitting and thus in increasing their energy efficiency. Building energy modelling is thus very important, providing the necessary tool to estimate building energy consumption. Models replicable to different buildings and relevant methodologies are thus quite useful, especially when they make possible narrowing down to a number of the most significant and influential parameters, whose values need

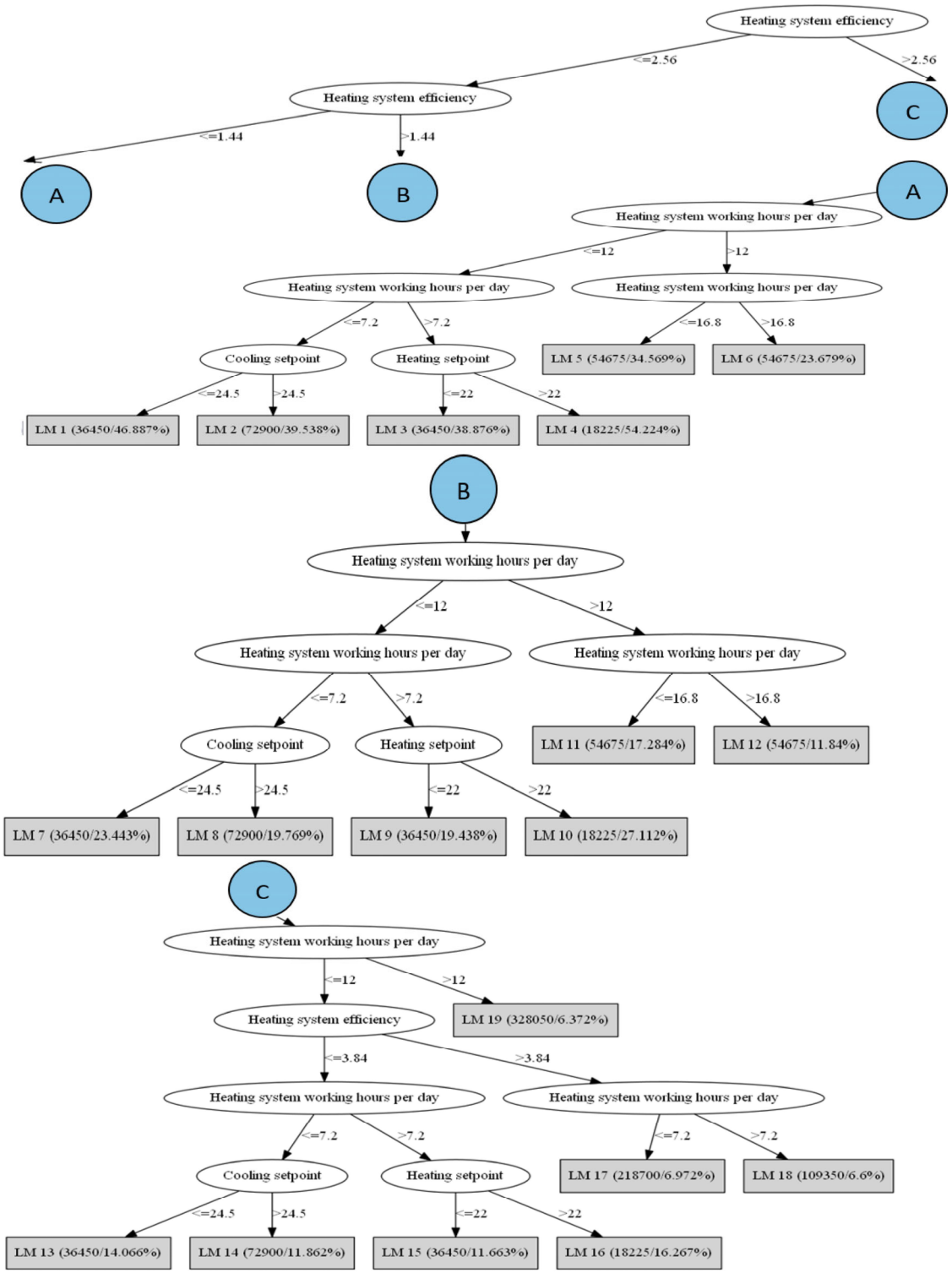

Fig. 21 Decision tree with linear regressions in the leaves generated by M5

to be provided.

Our approach leads in fact to such a result, obtaining the validated implementation of a calculation-based quasi-steady-state physical model based on the ISO 13790 standard. The model is successfully compared to ISO 52016 standard, updating ISO 13790. Model predictive capability is ascertained using as a benchmark the simulation results of EnergyPlus dynamic simulator, as well as the calculation results of a commercially available software. According to the results, the predictive capability of the model with respect to EnergyPlus

Table 6 Linear regressions of the leaves. Each cell includes the coefficient of the corresponding variable in the linear combination (with error term being the independent coefficient)

Rule/ leaf	Heating setpoint	Cooling setpoint	Occupancy	Heat recovery system efficiency	Lighting	Ventilation intermittency	Ventilation airflow rate	Uf opaque average	Heating system efficiency	Heating system working (h/day)	*Error term*
LM1	74,764.91	−56.62	22,912.56	−351,367.75	27,535.87	3.45	60,668.41	308,755.44	−25.41	−8.00	1,072,872.36
LM2	74,764.91	−80,926.79	20,338.87	−249,171.76	24,004.11	3.45	42,769.69	215,274.69	−25.41	−8.00	3,155,549.16
LM3	59,538.49	−63,327.67	7.27	−184,607.55	17,614.31	3.45	17.97	161,089.69	−25.41	−11.53	3,323,481.37
LM4	81.14	−63,327.67	7.27	−322,479.55	21.58	3.45	17.97	303,686.94	−25.41	−11.53	4,784,288.98
LM5	52,337.75	−51,011.17	2.83	−175,050.09	12,299.05	3.45	7.14	159,864.77	−25.41	−14.76	3,063,065.91
LM6	38,949.58	−38,704.12	2.83	−115,951.12	9,181.57	3.45	7.14	106,801.96	−25.41	−14.76	2,931,244.91
LM7	37,384.08	−30.06	11,456.68	−175,689.66	13,768.40	3.45	30,335.20	154,382.87	−25.41	−4.82	1,759,141.92
LM8	37,384.08	−40,465.14	10,169.83	−124,591.67	12,002.52	3.45	21,385.84	107,642.49	−25.41	−4.82	2,800,480.32
LM9	29,770.87	−31,665.59	4.03	−92,309.56	8,807.62	3.45	9.98	80,549.99	−25.41	−6.59	2,884,446.43
LM10	42.19	−31,665.59	4.03	−161,245.56	11.26	3.45	9.98	151,848.62	−25.41	−6.59	3,614,850.23
LM11	26,170.50	−25,507.33	1.81	−87,530.83	6,149.99	3.45	4.57	79,937.53	−25.41	−8.20	2,754,238.70
LM12	19,476.41	−19,353.81	1.81	−57,981.34	4,591.25	3.45	4.57	53,406.13	−25.41	−8.20	2,688,328.20
LM13	22,430.00	−16.80	6,873.97	−105,413.28	8,261.00	11.67	18,201.03	92,629.17	−8.33	−2.07	2,033,583.02
LM14	22,430.00	−24,278.67	6,101.84	−74,753.46	7,201.44	11.67	12,831.24	64,584.01	−8.33	−2.07	2,658,407.01
LM15	17,861.92	−18,998.75	2.15	−55,383.55	5,284.44	11.67	5.29	48,327.96	−8.33	−3.13	2,708,788.36
LM16	24.11	−18,998.75	2.15	−96,746.53	6.44	11.67	5.29	91,108.57	−8.33	−3.13	3,147,045.34
LM17	14,608.58	−17,183.78	4,141.77	−55,343.79	4,920.35	72,825.30	9,522.88	48,152.72	−39,426.71	−1.08	2,748,114.02
LM18	12,544.73	−12,373.35	0.92	−45,052.35	3,088.84	59,350.93	7,766.26	40,765.21	−30,020.31	−1.43	2,656,637.24
LM19	10,509.64	−10,328.51	0.28	−33,501.53	2,472.92	44,218.25	5,789.68	30,700.16	−29,856.09	−7,456.95	2,761,902.48

Table 7 Relative contribution of each variable in each linear regression

Rule/ leaf	% data	Heating setpoint	Cooling setpoint	Occupancy	Heat recovery system efficiency	Lighting	Ventilation intermittency	Ventilation airflow rate	Uf opaque average	Heating system efficiency	Heating system working (h/day)	*Error term*
LM1	2.7%	6.3%	0.0%	4.8%	−5.9%	5.2%	0.0%	5.9%	6.6%	0.0%	0.0%	65.2%
LM2	5.3%	6.2%	−6.7%	4.2%	−4.2%	4.5%	0.0%	4.1%	4.5%	0.0%	0.0%	65.5%
LM3	2.7%	5.8%	−6.2%	0.0%	−3.6%	3.9%	0.0%	0.0%	4.0%	0.0%	0.0%	76.5%
LM4	1.3%	0.0%	−5.9%	0.0%	−6.0%	0.0%	0.0%	0.0%	7.2%	0.0%	0.0%	80.8%
LM5	4.0%	5.5%	−5.3%	0.0%	−3.7%	2.9%	0.0%	0.0%	4.2%	0.0%	0.0%	78.4%
LM6	4.0%	4.5%	−4.5%	0.0%	−2.7%	2.4%	0.0%	0.0%	3.1%	0.0%	0.0%	82.8%
LM7	2.7%	4.2%	0.0%	3.2%	−3.9%	3.5%	0.0%	3.9%	4.3%	0.0%	0.0%	77.1%
LM8	5.3%	4.1%	−4.5%	2.8%	−2.8%	3.0%	0.0%	2.7%	3.0%	0.0%	0.0%	77.1%
LM9	2.7%	3.6%	−3.9%	0.0%	−2.3%	2.4%	0.0%	0.0%	2.5%	0.0%	0.0%	85.3%
LM10	1.3%	0.0%	−3.8%	0.0%	−3.8%	0.0%	0.0%	0.0%	4.6%	0.0%	0.0%	87.8%
LM11	4.0%	3.3%	−3.3%	0.0%	−2.2%	1.8%	0.0%	0.0%	2.6%	0.0%	0.0%	86.8%
LM12	4.0%	2.6%	−2.6%	0.0%	−1.6%	1.4%	0.0%	0.0%	1.8%	0.0%	0.0%	89.9%
LM13	2.7%	2.9%	0.0%	2.2%	−2.7%	2.4%	0.0%	2.7%	3.0%	0.0%	0.0%	84.2%
LM14	5.3%	2.8%	−3.1%	1.9%	−1.9%	2.1%	0.0%	1.9%	2.1%	0.0%	0.0%	84.2%
LM15	2.7%	2.4%	−2.6%	0.0%	−1.5%	1.6%	0.0%	0.0%	1.7%	0.0%	0.0%	90.2%
LM16	1.3%	0.0%	−2.5%	0.0%	−2.6%	0.0%	0.0%	0.0%	3.1%	0.0%	0.0%	91.8%
LM17	16.0%	1.8%	−2.1%	1.3%	−1.3%	1.4%	1.1%	1.3%	1.5%	−5.4%	0.0%	82.9%
LM18	8.0%	1.6%	−1.6%	0.0%	−1.2%	0.9%	1.0%	1.2%	1.3%	−4.4%	0.0%	86.8%
LM19	24.0%	1.3%	−1.3%	0.0%	−0.8%	0.7%	0.7%	0.8%	1.0%	−4.1%	−4.4%	84.9%
Average importance		**2.7%**	**−2.7%**	**1.0%**	**−2.1%**	**1.8%**	**0.4%**	**1.3%**	**2.3%**	**−2.2%**	**−1.1%**	**82.4%**

is quite better in summer rather than winter. For sure, one of the reasons for this is that EnergyPlus is a dynamic simulator and in winter, due to the temperature difference between outside and inside ambient, and due to the many transients during a typical day, transients have a strong impact on consumption.

As the building data availability part is a weak point of the building energy modelling approach, a sensitivity analysis of the model is performed utilizing machine learning techniques. This approach derives as most influential parameters for the building energy consumption estimation "Heating setpoint", "Cooling setpoint", "Heat recovery system efficiency" and "Uf opaque average". Our methodological approach is well transferable to different buildings and has a high degree of interpretability.

The developed QSSM provides an energy modelling solution of low computational cost and complexity when compared to other available software. Data input requirements are kept at the minimum level necessary for model calculations, thus minimizing the required user effort and enhancing user friendliness while targeting a broader audience with little or no experience in building energy modelling. Future directions include the extension of the analysis to different buildings, possibly with real data coming from the field, as well as the development of semi-automatic procedures to tune as many parameters as possible based on little amount of information.

Acknowledgements

The work presented in this paper has been funded in the framework of "Support Tool for Energy Efficiency Programmes in medical centres – STEER" project, Grant Agreement 655694, Research and Innovation Staff Exchange (RISE), H2020 – MSCA – RISE – 2014, European Commission.

References

ASHRAE (2017). ASHRAE Standard 140. Standard Method of Test for the Evaluation of Building Energy Analysis Computer Programs (ANSI Approved). Atlanta, GA, USA: American Society of Heating, Refrigerating, and Air Conditioning Engineers.

Ballarini I, Corrado V (2012). Analysis of the building energy balance to investigate the effect of thermal insulation in summer conditions. *Energy and Buildings*, 52: 168–180.

Corrado V, Fabrizio E (2007). Assessment of building cooling energy need through a quasi-steady state model: Simplified correlation for gain-loss mismatch. *Energy and Buildings*, 39: 569–579.

Edilclima (2017). Edilclima EC 700. Available at https://www.edilclima.it/software-termotecnica/prog-termotecnica-energetica/scheda/700

EN 15265 (2007). Energy performance of buildings. Calculation of energy needs for space heating and cooling using dynamic methods. General criteria and validation procedures. BSI British Standards.

EnergyPlus (2015). U.S. Department of Energy's (DOE). Building Technologies Office (BTO) and National Renewable Energy Laboratory (NREL).

Ferrarini L, Mantovani G (2013). Modeling and control of thermal energy of a large commercial building. In: Proceedings of IEEE International Workshop on Intelligent Energy Systems (IWIES), Vienna, Austria, pp. 149–154

Foucquier A, Robert S, Suard F, Stéphan L, Jay A (2013). State of the art in building modelling and energy performances prediction: A review. *Renewable and Sustainable Energy Reviews*, 23: 272–288.

Garcia Sanchez D, Lacarrière B, Musy M, Bourges B (2014). Application of sensitivity analysis in building energy simulations: Combining first- and second-order elementary effects methods. *Energy and Buildings*, 68: 741–750.

ISO 13790 (2008). Energy performance of buildings—Calculation of energy use for space heating and cooling. International Standard. International Organization for Standardization.

ISO 13370 (2017). Thermal performance of buildings—Heat transfer via the ground—Calculation methods. International Organization for Standardization.

ISO 52016-1 (2017). Energy performance of buildings—Energy needs for heating and cooling, internal temperatures and sensible and latent heat loads. International Organization for Standardization.

Kalogeras G, Koulamas C, Kalogeras A, Moronis A (2018a). Precision robustness testing of a simulation model for energy use in buildings. In: Proceedings of the 23rd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Turin, Italy, pp. 932–939.

Kalogeras G, Koulamas C, Kalogeras A, Moronis A (2018b). Verification and validation of a simulation model for energy use in buildings. In: Proceedings of the 14th IEEE International Workshop on Factory Communication Systems (WFCS), Imperia, Italy.

Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, Ye Q, Liu T-Y (2017). Lightgbm: A highly efficient gradient boosting decision tree. In: Proceedings of the Neural Information Processing Systems Conference (NIPS), Long Beach, CA, USA, pp. 3146–3154.

Koulamas C, Kalogeras AP, Pacheco-Torres R, Casillasc J, Ferrarini L (2018). Suitability analysis of modeling and assessment approaches in energy efficiency in buildings. *Energy and Buildings*, 158: 1662–1682.

Le Norme UNI TS 11300 (2011). La normativa tecnica di riferimento sul risparmio energetico e la certificazione energetica degli edifici. (in Italian)

Moronis A, Koulamas C, Kalogeras A (2017). Validation of a monthly quasi-steady-state simulation model for the energy use in buildings. In: Proceedings of the 22nd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Limassol, Cyprus.

Progress ITCE (2017). Energy Technology Perspectives 2017 Excerpt Informing Energy Sector Transformations. Paris: International Environment Agency.

Quinlan JR (1992). Learning with continuous classes. In: Proceedings of the 5th Australian Joint Conference on Artificial Intelligence, Hobart, Australia.

Rastegarpour S, Ferrarmi L, Pacheco-Torres R, Kalogeras A, Koulamas C (2017). Sensitivity analysis of medical centers energy consumption with EnergyPlus. In: Proceedings of the 22nd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Limassol, Cyprus.

Rastegarpour S, Ghaemi M, Ferrarini L (2018). A predictive control strategy for energy management in buildings with radiant floors and thermal storage. In: Proceedings of the SICE International Symposium on Control Systems (SICE ISCS), Tokyo, Japan.

Swan LG, Ugursal VI (2009). Modeling of end-use energy consumption in the residential sector: A review of modeling techniques. *Renewable and Sustainable Energy Reviews*, 13: 1819–1835.

Tian W (2013). A review of sensitivity analysis methods in building energy analysis. *Renewable and Sustainable Energy Reviews*, 20: 411–419.

Van Dijk H, Spiekman M, De Wilde P (2005). A monthly method for calculating energy performance in the context of European building regulations. In: Proceedings of the 9th International IBPSA Building Simulation Conference, Montreal, Canada, pp. 255–262.

Wang S, Yan C, Xiao F (2012). Quantitative energy performance assessment methods for existing buildings. *Energy and Buildings*, 55: 873–888.

Wang Y, Witten IH (1997). Inducing model trees for continuous classes. In: Proceedings of the 9th European Conference on Machine Learning, Prague, Czech Republic.