

Data-Driven Building Energy Modelling – Generalisation Potential of Energy Signatures Through Interpretable Machine Learning

Massimiliano Manfren – University of Southampton, UK – M.Manfren@soton.ac.uk

Maria Cristina Tommasino – ENEA, Italy – cristina.tommasino@enea.it

Lamberto Tronchin – University of Bologna, Italy – lamberto.tronchin@unibo.it

Abstract

Building energy modeling based on data-driven techniques has been demonstrated to be effective in a variety of situations. However, the question about its limits in terms of generalization is still open. The ability of a machine-learning model to adapt to previously unseen data and function satisfactorily is known as generalization. Apart from that, while machine-learning techniques are incredibly effective, interpretability is required for a "human-in-the-loop" approach to be successful. This study develops and tests a flexible regression-based approach applied to monitored energy data on a Passive House building. The formulation employs dummy (binary) variables as a piecewise linearization method, with the procedures for producing them explicitly stated to ensure interpretability. The results are described using statistical indicators and a graphic technique that allows for comparison across levels in the building systems. Finally, suggestions are provided for further steps toward generalization in data-driven techniques for energy in buildings.

1. Introduction

Data-driven building energy modeling methods that use machine-learning techniques have been shown to be useful in a variety of applications (Hong et al., 2020), from design (Westermann & Evins, 2019) to operation (Manfren et al., 2020). As a result, they have the potential to become a key tool for accelerating the ongoing process of building stock decarbonisation (Norton et al., 2021; Tronchin, & Knight, 2016) as well as an integral part of innovative services and technologies (Farina et al., 1998; Manfren et al., 2021a). However, the question of whether data-driven approaches can be generalized is still being debated.

The ability of a machine-learning model to adapt to previously unknown data and perform reasonably well, given specified performance criteria, is referred to as generalization. A simple example of generalization is a model trained on building energy consumption data over a period of time and then used to estimate energy consumption during a successive period of time. This is the counterfactual approach used in Measurement and Verification (M&V) protocols, which uses statistical indicators as model acceptability criteria during the calibration phase. A more ambitious form of generalization would be that of using data-driven methods on energy modeling problems involving sets of building with homogeneous characteristics. Using data-driven methods on energy modeling problems involving different sets of buildings with relatively similar characteristics would be a more ambitious type of generalisation.

In this research, we use regression models trained and tested on building energy signatures as a tool for addressing the generalization problem. In fact, methods based on energy signature (ISO, 2013) (i.e., energy divided by the number of operating hours in the time interval of the analysis, corresponding to an average power) are scalable (temporally and spatially) (Manfren et al., 2021b; Tronchin, 2021; Tronchin et al., 2018), can work with unstructured data (using clustering) (Pistore et al., 2019; Westermann et al., 2020) and provide results that are weather normalized (Fazeli et al., 2016). Additionally, energy signatures can be scaled according to the building's size (Pistore et al., 2019; Tronchin et al., 2016), to produce a performance comparison that is independent of the size. Further, regression-based approaches are considered interpretable machine learning techniques (ISO, 2020) because it is possible

to predict how the model output will change in response to a change in input data or algorithmic parameters (i.e., the rationale behind model output and the algorithmic logic can be easily understood in human terms) (Fabbri et al., 2014 and 2021). In this study, we employ this technique to analyze monitored data from a Passive House building in the Province of Forlì-Cesena in northern Italy, with the goal of improving formulations at the state of the art while also considering generalization and interpretability issues (Tronchin, 2005 and 2021).

2. Methods

The regression model proposed in this research is a reformulation of the variable-based degree-days regression, originally proposed by Kissock et al. (2003) in their Inverse Modeling Toolkit (IMT), which has been included in ASHRAE 14:2014 (ASHRAE, 2014) and has been evolving steadily with different algorithmic formulations. Essentially, interpretable regression-based methods can be based on general piecewise linearization methods (Lin et al., 2013) and use dummy (binary variables) to handle non-linearities. Model formulation and calibration criteria are reported hereafter (Tronchin et al., 2021a and 2021b).

2.1 Model Formulation

Separate sub-models (heating, base load and cooling), indicated in Table 1, are combined into a single model (which is the sum of the individual sub-models) by introducing additional variables (dummy, 0-1 binary variables) to the original datasets using rules, indicated in Table 2. The binary variables multiply the original variables and act as interaction terms. Two types of models are tested, type 1 and type 2. In the first case, the independent variable is outdoor air temperature, while, in the second case, the independent variables are outdoor air temperature and total solar radiation on horizontal surface. The dependent variables are the energy signatures calculated from the monitored data described in Section 3.

The rules provided hereafter in Table 2 can be applied both manually and in an automated way to the

dataset, for example, using ranges of balance-point temperature (change-points) for heating and cooling (Manfren et al., 2019).

Table 1 – Regression model formulation

Mode Demand		Sub-models
Type 1	Heating	$q_h = a_0(X_h) + a_1(X_h\theta_e) + \varepsilon_h$ (1)
	Base load	$q_b = b_0(X_b) + b_1(X_b\theta_e) + \varepsilon_b$ (2)
	Cooling	$q_c = c_0(X_c) + c_1(X_c\theta_e) + \varepsilon_c$ (3)
Type 2	Heating	$q_h = a_0(X_h) + a_1(X_h\theta_e) + a_2(X_hI_{sol}) + \varepsilon_h$ (4)
	Base load	$q_b = b_0(X_b) + b_1(X_b\theta_e) + b_2(X_bI_{sol}) + \varepsilon_b$ (5)
	Cooling	$q_c = c_0(X_c) + c_1(X_c\theta_e) + c_2(X_cI_{sol}) + \varepsilon_c$ (6)

Table 2 – Rules for dummy variable creation

Rule	Description	Variables
1	If the energy demand is greater than 0 for the corresponding sub-model (e.g., heating, cooling or base load), then the dummy variable is equal to 1.	X_h , X_b , X_c
2	If the outdoor air temperature is lower than balance point temperature for heating (i.e., heating base temperature), the dummy variable for heating is equal to 1.	X_h
3	If the outdoor air temperature is greater than balance point temperature for cooling (i.e., cooling base temperature), the dummy variable for cooling is equal to 1.	X_c
4	All the dummy variables (that partition heating, cooling and base load demands) should be coherent with the schedules of operation for building services (i.e., months of heating and cooling system operation).	X_h , X_b , X_c
5	The dummy variables for base load are assumed to be 1 in all the months (i.e., electricity and hot water demand are always present).	X_b

2.2 Model Calibration Criteria

Following the indications proposed by state-of-the-art Measurement and Verification (M&V) protocols, such as ASHRAE 14:2014 (ASHRAE, 2014), Efficiency Value Organization (EVO) IPMVP (EVO,

2003), and Federal Energy Management Program (FEMP) (FEMP, 2008), the thresholds of acceptability for regression models as calibrated with monthly data are reported in Table 3.

Table 3 – Model calibration criteria

Data interval	Metric	ASHRAE Guidelines 14		
		IPMVP	FEMP	
Monthly	<i>NMBE</i>	±5	±20	±5
	<i>Cv(RMSE)</i>	15	-	15

3. Case Study

The case study is Passive House building, located in the northern Italian province of Forlì-Cesena; the essential building data are reported in Table 4.

The building was monitored for three years and electric and thermal demand data were split by end use, as indicated in Table 5.

The modeling workflow pursued to test the models reported in Table 1 incrementally is set out in the following steps:

1. Initial training, year 1, 2.
2. Testing, year 3 (model created in step 1).
3. Retraining, year 1, 2 and 3.

The results obtained are reported hereafter in Section 4.

Table 4 – Building design data

Group	Type	Unit	Design
Geometry	Gross volume	m ³	1557
	Net volume	m ³	1231
	Heat loss surface area	m ²	847
	Net floor area	m ²	444
	Glazed area/total wall area ratio	%	22.5
	Surface/volume ratio	1/m	0.54
Envelope	U value external walls	W/(m ² K)	0.18
	U value roof	W/(m ² K)	0.17
	U value transparent components	W/(m ² K)	0.83
HVAC and DHW	Ground-source heat pump (GSHP) - Brine/Water Heat Pump (B0/W35)*	kW	8.4
	Borehole heat exchanger (2 double U boreholes)	m	100
On-site energy production	Building Integrated Photovoltaic (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

* EN 14511 test condition in heating mode, brine at 0 °C and water 35 °C with supply-return temperature difference $\Delta t = 10$ °C.

Table 5 – Dataset used for modeling

Data	Enduse	Interval	Monitoring period
Electric	Total	Monthly	3 years
	HVAC, DHW	Monthly	3 years
	Appliances and lighting	Monthly	3 years
Thermal energy	Heating	Monthly	3 years
	Cooling	Monthly	3 years

4. Results and Discussion

In this section, the results of the model training and testing process are reported, indicating both numerical results represented by statistical indicators (Table 3) and visualization of energy signatures for the different models fitted.

4.1 Numerical Results of Regression Models

The results obtained are split with respect to the two types of models considered, namely type 1 and type 2.

4.1.1 Model Type 1

It can be clearly seen that model type 1, after 2 years, obtains values for the indicators *NMBE* and *Cv(RMSE)* that make them acceptable as calibrated (Table 3), with the exception of electricity demand for HVAC and DHW and thermal demand for cooling. In these cases, the values are higher than 15 % for *Cv(RMSE)* but lower than 20 %.

Table 6 – Model type 1 – Initial training (year 1 and 2)

Data	End-use	<i>EN(M)</i> <i>EN(P)</i>		<i>R</i> ²	<i>NMBE</i>	<i>Cv(RMSE)</i>
		kWh	kWh			
Electric	Total	23812	23510	84.77	-1.27	12.59
	HVAC, DHW	8611	8499	91.05	-1.30	17.99
	Appliances	15201	15202	70.27	0.01	10.12
Thermal	Heating	17761	17757	98.05	-0.03	8.87
	Cooling	5054	5004	93.50	-0.99	16.45

The model trained for the period indicated in Table 6 (first 2 years) is then tested for the third year of monitoring. In this case, we can see how, for model type 1, the statistical indicators in the testing phase are larger (i.e., the model performance is lower in terms of goodness of fit).

Table 7 – Model type 1 – Testing (year 3)

Data	End-use	<i>EN(M)</i> <i>EN(P)</i>		<i>R</i> ²	<i>NMBE</i>	<i>Cv(RMSE)</i>
		kWh	kWh			
Electric	Total	11318	10167	65.37	-10.17	20.15
	HVAC, DHW	3659	3181	90.78	-13.07	20.76
	Appliances	7659	7013	26.83	-8.44	22.13
Thermal	Heating	6029	5460	85.28	-9.53	21.17
	Cooling	1784	1841	63.23	3.19	20.21

Finally, models are retrained with the entire 3-year dataset, obtaining results that are slightly better compared to the ones presented in Table 6, but not largely different.

Table 8 – Model type 1 – Retraining (year 1, 2 and 3)

Data	End-use	<i>EN(M)</i> <i>EN(P)</i>		<i>R</i> ²	<i>NMBE</i>	<i>Cv(RMSE)</i>
		kWh	kWh			
Electric	Total	35130	34819	84.41	-0.88	12.12
	HVAC, DHW	12270	12139	91.23	-1.07	17.35
	Appliances	22860	22868	67.34	0.04	10.56
Thermal	Heating	23790	23795	95.91	0.02	12.42
	Cooling	6838	6735	90.91	-1.51	17.77

4.1.2 Model Type 2

The same workflow presented for model type 1 in Section 4.1.1 is repeated here for model type 2. The results are reported in Tables 9, 10 and 11, respectively. In this case, we can see a moderate improvement for the model training (Table 9) and retraining (Table 11), but a much better performance of the models in the testing phase (Table 10). In general, model type 2 performs better than type 1.

Table 9 – Model type 2 – Initial training (year 1 and 2)

Data	End-use	$EN(M)$	$EN(P)$	R^2	$NMBE$	$Cv(RMSE)$
		kWh	kWh	%	%	%
Electric	Total	23812	23518	88.36	-1.23	11.05
	HVAC, DHW	8611	8534	91.74	-0.90	17.16
	Appliances	15201	15204	82.19	0.02	7.84
Thermal	Heating	17761	17759	98.95	-0.01	6.50
	Cooling	5054	5229	93.54	3.46	18.34

Table 10 – Model type 2 – Testing (year 3)

Data	End-use	$EN(M)$	$EN(P)$	R^2	$NMBE$	$Cv(RMSE)$
		kWh	kWh	%	%	%
Electric	Total	11318	10551	85.96	-6.78	12.53
	HVAC, DHW	3659	3248	91.75	-11.24	18.80
	Appliances	7659	7329	66.66	-4.30	12.04
Thermal	Heating	6029	5828	94.13	-3.59	13.52
	Cooling	1784	1964	85.58	10.10	16.66

Table 11 – Model type 2 – Retraining (year 1, 2 and 3)

Data	End-use	$EN(M)$	$EN(P)$	R^2	$NMBE$	$Cv(RMSE)$
		kWh	kWh	%	%	%
Electric	Total	35130	34847	89.75	-0.81	10.16
	HVAC, DHW	12270	12200	92.41	-0.57	16.56
	Appliances	22860	22870	84.05	0.05	7.60
Thermal	Heating	23790	23798	98.32	0.03	7.96
	Cooling	6838	7002	91.61	2.39	17.04

4.2 Energy Signature Visualization

In this section, energy signatures are visualized for both electric and thermal data and divided by end use. Section 4.2.1 focuses on electricity data, while Section 4.2.2 focuses on thermal data.

4.2.1 Energy Signatures – Electricity

The energy signature shapes shown in Fig. 1 for total electricity, Fig. 2 for electricity for HVAC and DHW, Fig. 3 for appliances and lighting, are substantially similar to a 5p or nearly 4p model according to the classification proposed by ASHRAE14:2014. The charts are used to compare models type 1 and 2 at multiple levels in the building

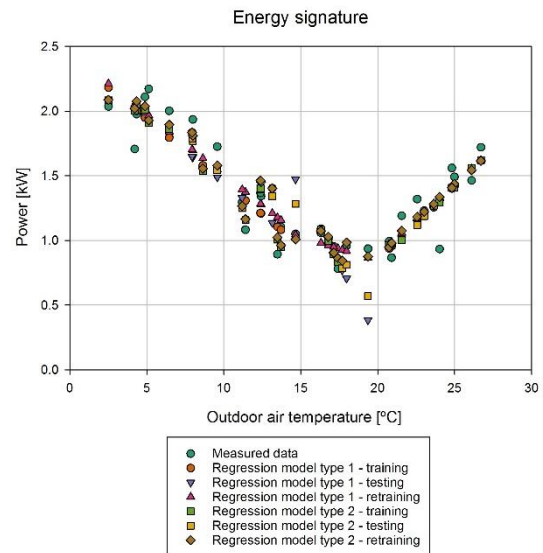


Fig. 1 – Energy signatures, measured data and regression models types 1 and 2 – Total electricity

The spread of data around the trend line is quite limited in Fig. 1 (total electricity) and Fig. 2 (electricity for HVAC and DHW), while it is more pronounced for Fig. 3 (appliances and lighting). In the latter, there is a lower temperature dependence (steepness of the trend line), compared to the other cases. However, the dependence is actually on daylight hours, which are correlated to temperature (lower temperatures correspond to winter condition where daylight hours are less and electric consumption for lighting is higher) and on the actual operation pattern, whose variability also determines the larger spread of values.

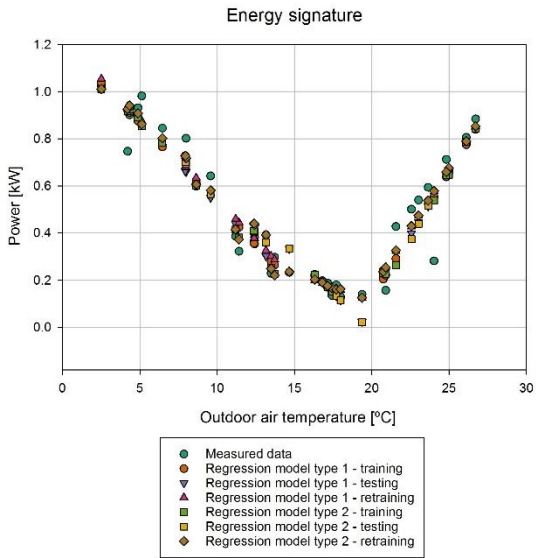


Fig. 2 – Energy signatures, measured data and regression models types 1 and 2 – Electricity for HVAC and DHW

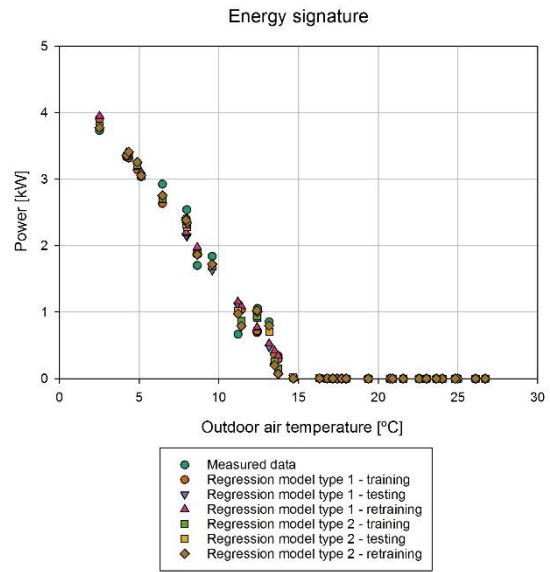


Fig. 4 – Energy signatures, measured data and regression models types 1 and 2 – Thermal energy for heating

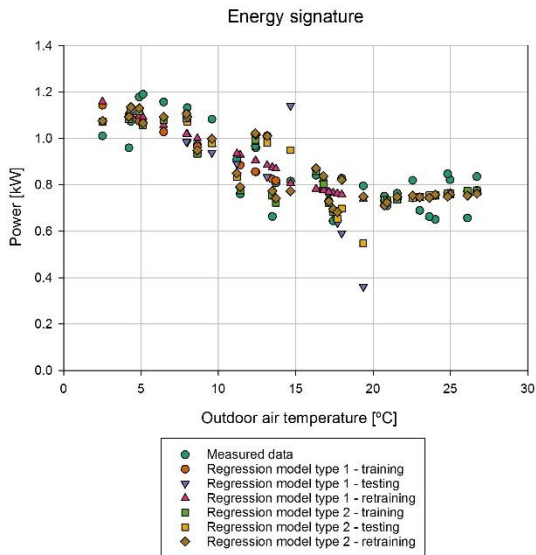


Fig. 3 – Energy signatures, measured data and regression models types 1 and 2 – Electricity for appliances and lighting

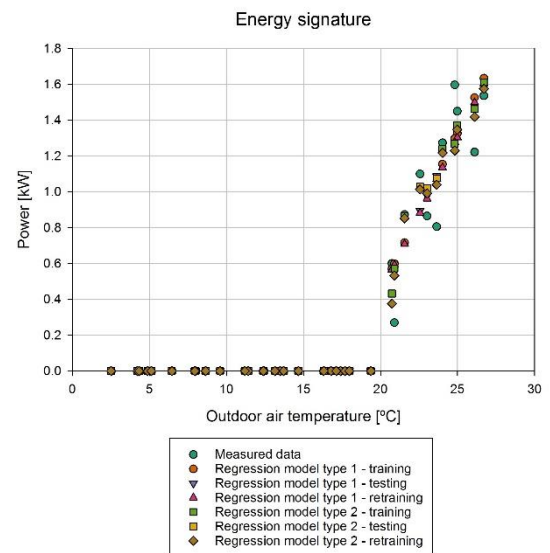


Fig. 5 – Energy signatures, measured data and regression models types 1 and 2 – Thermal energy for cooling

4.2.2 Energy Signatures – Thermal Energy

The energy signature shapes shown in Fig. 4 for thermal demand for heating and in Fig. 5 for thermal demand for cooling are essentially similar to a 3p model according to the classification proposed by ASHRAE14:2014. There are clearly some months when the technical systems do not produce either heating or cooling. Finally, it is possible to identify graphically, in an approximated way, the balance-point temperature for heating and cooling, around 14 °C and 19 °C, respectively.

5. Conclusion

Machine learning-based building energy modeling techniques have proved to be effective in a range of applications. However, problems such as generalization and interpretability must be considered in order to enable the widespread adoption of these techniques. A piecewise linear regression model (interpretable) was proposed to analyze monitored data from a Passive House building, located in the northern Italian province of Forlì-Cesena. The

building was monitored for three years, and this technique requires at least two years of monthly interval data to be used effectively. Nonetheless, the formulation provided is quite simple and flexible; the visualization of energy signatures can also help understanding the actual spread of data around the trend line, which represent outdoor air temperature dependence.

Further efforts involving the categorization of building data according to archetypes could be considered to address the generalization issue effectively. Finally, interpretability is extremely relevant because of the necessity to promote a “human-in-the-loop” approach when using machine learning tools and the transparent link between regression model formulation and other analytical techniques at the state-of-the-art could represent an interesting research area from the perspective of future studies (Tronchin et al., 2020a and 2020b).

Nomenclature

Symbol	Quantity	Unit
a_0, b_0, c_0	regression coefficients, intercept	kW
a_1, b_1, c_1	regression coefficients, temperature dependence term	kW/K
a_2, b_2, c_2	regression coefficients, solar radiation dependence term	m ²
$Cv(RMSE)$	coefficient of variation of RMSE	-
$EN(M)$	measured energy	kWh
$EN(P)$	predicted energy	kWh
I_{sol}	total solar radiation on horizontal surface (direct and diffuse) average hourly value on monthly base	kW/m ²
$NMBE$	normalized mean bias error (expressed in percentage)	-
q^h	energy signature heating	kW
q^b	energy signature base load	kW

q^c	energy signature cooling	kW
R^2	determination coefficient (expressed in percentage)	-
X_h	dummy variable (binary 0-1) heating	-
X_b	dummy variable (binary 0-1) base load	-
X_c	dummy variable (binary 0-1) cooling	-
θ_e	outdoor air temperature	°C
ε_h	error term heating	kW
ε_b	error term base load	kW
ε_c	error term cooling	kW

References

- ASHRAE. 2014. *ASHRAE Guideline 14-2014: Measurement of Energy, Demand, and Water Savings*. American Society of Heating, Refrigerating and Air-Conditioning Engineers: Atlanta, GA, US.
- EVO. 2003. *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, US.
- Fabbri, K., L. Tronchin, and V. Tarabusi. 2014. “Energy Retrofit and Economic Evaluation Priorities Applied at an Italian Case Study”. *Energy Procedia* 45: 379-384. doi: 10.1016/j.egypro.2014.01.041.
- Fabbri, K., L. Tronchin, and F. Barbieri. 2021. “Cocoon fibre insulators: The hygrothermal behaviour in the case of green roofs”. *Construction and Building Materials* 266: 1-9. doi: 10.1016/j.conbuildmat.2020.121026.
- Farina, A., L. Langhoff, and L. Tronchin. 1998. “Acoustic characterisation of “virtual” musical instruments: using MLS technique on ancient violins”. *Journal of new music research* 27(4): 359-379.

- Fazeli, R., M. Ruth, and B. Davidsdottir. 2016. "Temperature Response Functions for Residential Energy Demand – A Review of Models." *Urban Climate* 15: 45–59. doi: <https://doi.org/https://doi.org/10.1016/j.uclim.2016.01.001>
- FEMP. 2008. *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, US.
- Hong, T., Z. Wang, X. Luo, & W. Zhang. 2020. "State-of-the-Art on Research and Applications of Machine Learning in the Building Life Cycle." *Energy and Buildings* 212: 109831. doi: <https://doi.org/https://doi.org/10.1016/j.enbuild.2020.109831>
- ISO. 2013. *ISO 16346:2013, Energy Performance of Buildings – Assessment of Overall Energy Performance*.
- ISO. 2020. *ISO/IEC TR 29119-11:2020(En) Software and Systems Engineering – Software Testing – Part 11: Guidelines on the Testing of AI-Based Systems*.
- Kissock, J. K., J. S. Haberl, and D. E. Claridge. 2003. "Inverse Modeling Toolkit: Numerical Algorithms." *ASHRAE Transactions* 109: 425.
- Lin, M.-H., J. Gunnar Carlsson, D. Ge, J. Shi, and J.-F. Tsai. 2013. "A Review of Piecewise Linearization Methods." *Mathematical Problems in Engineering* 2013: 101376. <https://doi.org/10.1155/2013/101376>
- Manfren, M., B. Nastasi, E. A. Piana, and L. Tronchin. 2019. "On the link between energy performance of building and thermal comfort: An example". In *AIP Conference Proceedings* 2123: 1-9. doi: <https://doi.org/10.1063/1.5116993>
- Manfren, M., B. Nastasi, and L. Tronchin. 2020. "Linking Design and Operation Phase Energy Performance Analysis Through Regression-Based Approaches." *Frontiers in Energy Research* 8: 288. doi: <https://doi.org/10.3389/fenrg.2020.557649>
- Manfren, M., B. Nastasi, L. Tronchin, D. Groppi, and D. Astiaso Garcia. 2021a. "Techno-Economic Analysis and Energy Modelling as a Key Enablers for Smart Energy Services and Technologies in Buildings." *Renewable and Sustainable Energy Reviews* 150: 111490. doi: <https://doi.org/https://doi.org/10.1016/j.rser.2021.111490>
- Manfren, M., M. Sibilla, and L. Tronchin. 2021b. "Energy Modelling and Analytics in the Built Environment—A Review of Their Role for Energy Transitions in the Construction Sector." *Energies* 14(3): 679. doi: <https://doi.org/10.3390/en14030679>
- Norton, B., W. B. Gillett, and F. Koninx. 2021. "Decarbonising Buildings in Europe: A Briefing Paper." *Proceedings of the Institution of Civil Engineers - Energy* 174(4): 147-155. doi: <https://doi.org/10.1680/jener.21.00088>
- Pistore, L., G. Pernigotto, F. Cappelletti, A. Gasparella, and P. Romagnoni. 2019. "A Stepwise Approach Integrating Feature Selection, Regression Techniques and Cluster Analysis to Identify Primary Retrofit Interventions on Large Stocks of Buildings." *Sustainable Cities and Society* 47: 101438. doi: <https://doi.org/https://doi.org/10.1016/j.scs.2019.101438>
- Tronchin, L., M. Manfren, and L. C. Tagliabue. 2016. "Optimization of Building Energy Performance by Means of Multi-Scale Analysis – Lessons Learned from Case Studies." *Sustainable Cities and Society* 27: 296–306. doi: <https://doi.org/https://doi.org/10.1016/j.scs.2015.11.003>
- Tronchin, L. 2005. "Modal analysis and intensity of acoustic radiation of the kettledrum". *The Journal of the Acoustical Society of America* 117(2): 926-933. doi: <https://doi.org/10.1121/1.1828552>.
- Tronchin, L. 2021. "Variability of room acoustic parameters with thermo-hygrometric conditions". *Applied Acoustics* 177: 1-14. doi: <https://doi.org/10.1016/j.apacoust.2021.107933>
- Tronchin, L., K. Fabbri, and C. Bertolli. 2018. "Controlled Mechanical Ventilation in Buildings: A Comparison between Energy Use and Primary Energy among Twenty Different Devices". *Energies* 11: 1-20. doi: <https://doi.org/10.3390/en11082123>
- Tronchin, L., and D. J. Knight. 2016. "Revisiting Historic Buildings through the Senses. Visualising Aural and Obscured Aspects of San Vitale, Ravenna". *International Journal of*

- Historical Archaeology* 20: 127-145. doi:
<https://doi.org/10.1007/s10761-015-0325-2>
- Tronchin, L., M. Manfren, and V. Vodola. 2020a. "The carabattola - vibroacoustical analysis and intensity of acoustic radiation (IAR)". *Applied Sciences* 10(2): 641 doi:
<https://doi.org/10.3390/app10020641>
- Tronchin, L., M. Manfren, and V. Vodola. 2020b. "Sound characterization through intensity of acoustic radiation measurement: A study of persian musical instruments". *Applied Sciences* 10(2): 633. doi:
<https://doi.org/10.3390/app10020633>
- Tronchin, L., F. Merli, and M. Manfren. 2021a. "On the acoustics of the Teatro 1763 in Bologna". *Applied Acoustics* 172: 1-9. doi:
<https://doi.org/10.1016/j.apacoust.2020.107598>
- Tronchin, L., F. Merli, and M. Dolci. 2021b. "Virtual acoustic reconstruction of the Miners' Theatre in Idrija (Slovenia)". *Applied Acoustics* 172: 1-9. doi:
<https://doi.org/10.1016/j.apacoust.2020.107595>
- Westermann, P., C. Deb, A. Schlueter, and R. Evins. 2020. "Unsupervised Learning of Energy Signatures to Identify the Heating System and Building Type Using Smart Meter Data." *Applied Energy* 264: 114715. doi:
<https://doi.org/https://doi.org/10.1016/j.apenergy.2020.114715>
- Westermann, P., and R. Evins. 2019. "Surrogate Modelling for Sustainable Building Design – A Review." *Energy and Buildings* 198: 170–86. doi:
<https://doi.org/https://doi.org/10.1016/j.enbuild.2019.05.057>