**Data-driven load profile modelling for advanced Measurement and Verification (M&V) in a fully electrified building**

Benedetto Nastasia,[[1]](#footnote-1)\*, Massimiliano Manfrenb, Daniele Groppia, Mario Lamagnac, Francesco Mancinia, Davide Astiaso Garciaa

a Department of Planning, Design & Technology of Architecture, Sapienza University of Rome, Via Flaminia 72 00196 Rome, Italy

b Faculty of Engineering and Physical Sciences, University of Southampton, Boldrewood Campus – SO16 7QF Southampton, United Kingdom

c Department of Astronautical, Electrical and Energy Engineering, Sapienza University of Rome, Via Eudossiana 18 00184 Rome, Italy

# Abstract

The process of decarbonising stock will result in a considerable shift in consumption away from fossil fuels and toward electricity. The growing trend of building electrification necessitates a thorough examination from the standpoint of end-use efficiency and dynamic behaviour in order to fully understand the potential for grid flexibility. The problem of accurately representing dynamic behaviour (e.g. electric load profiles) while retaining simple and easy to use modelling approaches (i.e. supporting a "human in the loop" approach to data-driven methodologies) is a challenging task, especially when operating conditions are very variable. For these reasons, we used an interpretable (regression-based) technique called Time Of Week a Temperature (TOWT) to predict the dynamic electric load profiles before, during, and after the COVID lockdown (for nearly 4 years) of a public office building in Southern Italy, the Procida City Hall. TWOT models perform reasonably well in most conditions, and their application allowed for the detection of changes in energy demand patterns, critical aspects to consider when tuning them, and areas for improvement in algorithmic formulation and data visualisation, which will be the focus of future research.

**Keywords:** Data-driven methods, building energy demand, Regression-based approaches, Energy Management, Measurement and Verification, Energy Analytics, M&V 2.0.

**Highlights:**

* An office building monitored for nearly 4 years, before, during and after COVID-19 lockdown.
* TWOT modelling technique is used to predict hourly load profiles in different periods.
* The electricity consumption before, during and after the COVID-19 lockdown is assessed.
* The anomalies in load profile patterns are highlighted visually and numerically.
* A criterion is defined to subset time series and improve models’ predictive performance.
* Possible improvements from the point of view of algorithmic formulation and data visualization are identified.

**Table of contents**

[Abstract 1](#_Toc103959831)

[1 Nomenclature 4](#_Toc103959832)

[2 Introduction 4](#_Toc103959833)

[3 Literature review 6](#_Toc103959836)

[4 Methodology 8](#_Toc103959837)

[4.1 Model formulation 8](#_Toc103959838)

[4.2 Model calibration criteria 10](#_Toc103959839)

[4.3 Modelling workflow 12](#_Toc103959840)

[5 Case study description 12](#_Toc103959841)

[6 Results and discussion 14](#_Toc103959887)

[6.1 Results of initial models’ fitting 15](#_Toc103959888)

[6.2 Discussion of results 17](#_Toc103959889)

[6.2.1 Visual analysis before models’ training 17](#_Toc103959890)

[6.2.2 Visual analysis after models’ training 22](#_Toc103959973)

[6.3 Results of model retraining on subsets of data 27](#_Toc103959974)

[6.4 Limitations and further research 29](#_Toc103959975)

[7 Conclusions 30](#_Toc103959976)

[References 32](#_Toc103959977)

# Nomenclature

The nomenclature used in the paper are reported hereafter.

*Table 1: Nomenclature*

|  |  |  |
| --- | --- | --- |
| **Variables and parameters** | | |
| **Symbol** | **Quantity** | **Unit** |
| *aj* | regression coefficients, *j* is the hour of the week | kW |
| *bk* | regression coefficients, *k* is the temperature segment | kW/K |
| *co* | regression coefficient, intercept |  |
| *CV(RMSE)* | coefficient of variation of *RMSE* (expressed in percentage) | - |
| *Locc* | electric load when the building is occupied | kW |
| *Lunocc* | electric load when the building is not occupied | kW |
| *NMBE* | normalized mean bias error (expressed in percentage) | - |
| *RMSE* | root mean square error | kW |
| *R2* | determination coefficient (expressed in percentage) | - |
| *T(i)k* | outdoor air temperature value for the segment *k* at time interval *i* | ºC |
| *tow,j* | time of week binary variable | - |

# Introduction

The need to accelerate building stock decarbonisation [1] necessitates a significant shift in consumption from fossil fuels to electricity. On the other hand, the increased electrification of building stock [2] necessarily requires a more careful analysis of electricity consumption because heating and cooling demands, as well as electricity demand for lighting, appliances and electric vehicles [3], must be met. Further, there are multiple issues to be considered in the process of increasing building stock energy efficiency, namely refurbishment options [4], appropriate design and sizing of technical systems [5] and neighbourhood [6] or community scale solutions that require adequate policies [7], which are part of the current scientific debate on a global scale [8]. On the one hand, as shown for the United States [1], the potential for energy savings and flexibility defined by targeted actions can be highly significant at the national level. As a result, these approaches can make a significant contribution to the grid's deep decarbonisation [2]. On the other hand, recent research programmes such as the IEA EBC Annex 79 "Occupant-Centric Building Design and Operation" [3] have focused on giving meaningful insights on user behaviour and preferences utilising dynamic data.

Despite the fact that the notion that user behaviour is the most significant factor on energy usage [4] must be critically questioned, it is clear that a significant information gap [5] exists in many circumstances, even among those who manage buildings. In recent years, advanced measurement and verification (M&V) techniques [6] (commonly referred to as M&V 2.0) have received a lot of attention in the industry for quantifying energy efficiency and flexibility potential [7]. The dynamics of energy consumption and grid interaction, especially when on-site renewable power generation is present, are critical to monitor [10] and handle from a control perspective. Additionally, a proper length for the monitoring period [12] and an adequate granularity of data is crucial to be able to evaluate the real dynamic energy behaviour and extract useful insights. Indeed, energy bills with monthly data resolution are the simplest form of datasets that can be obtained and analysed [13], but this is not enough if the goal is promoting a better integration of renewables [14] and new business models in the energy market [15], where the time scale of the minutes matters [16]. Possible ways to reconstruct dynamic electricity load profiles from utility bills have been proposed recently for example by Lamagna et al. [18], but monitoring tools that allow to record consumption data on higher resolution [21] are preferrable. Greater access to energy data and information (i.e. some data analytics) is part of a fair energy transition [22], and while sophisticated "black box" models such as deep learning show promising results [23], simpler techniques that provide adequate performance but are easier to understand are important to promote a “human in the loop” approach to data driven methods and increase the level of energy literacy. For these reasons, we selected in this research a regression-based approach called Time Of Week a Temperature (TOWT) [24] to monitor the dynamic electric load profiles before, during and after the COVID-19 lockdown of a public office building, the Procida City Hall, located in Southern Italy. The aim of the research is to monitor the building's energy behaviour over a long period of time (nearly 4 years) and to test the applicability of TOWT in critical conditions. Understanding the actual impact of measures during and after COVID-19 lockdown, identifying critical aspects when tuning the models, and, finally, identifying potential areas of improvement (with a focus on interpretability) are the specific objectives of monitoring and testing activities. Section 2 goes into greater detail about why the TOWT regression-based approach was chosen.

# Literature review

The modeling strategies presented in this study can be framed within a research area that make use of the fundamentals of data-driven methods [24], machine learning [25], M&V 2.0 [26], and building energy calibration techniques [27] to create data-driven workflows that can be applied at various stages of the building life cycle, from early design stage optimization [28] to automated identification of best-fitting models [29] and operational patterns [30]. Indeed, top-down statistical analysis of building energy performance can be linked to these modeling techniques [31,32]. In this regard, various types of (data-driven) models, when properly designed for interoperability and combined into systems [33], can make a significant contribution to the digital transformation of the built environment [34]. In simpler terms, the presence of multiple models that can work in synergy (being designed with similar rules and standards) and use both short-term/high-frequency measurements (i.e. daily, hourly, sub-hourly) and long-term/low-frequency measurements (i.e. monthly, over multiple years) in a harmonised manner can greatly improve the robustness of the performance assessment and create multiple feedback loops and continuous improvement, ideally following a Deming cycle or PDCA (Plan-Do-Check-Act), as indicated in standardization of energy management [35]. The increasing use of machine learning (ML) techniques is causing a rapid evolution in the broad research area of data-driven methods in the energy energy sector, but it is necessary to reflect on the concepts of interpretability and explainability. Interpretability is defined as the "level of understanding how the underlying (AI) technology works" in the ISO/IEC TR 29119-11:2020 standard [36] for Artificial Intelligent (AI) software testing. The provided definition does not refer to the extent to which the internal mechanics of the machine learning algorithm can be explained in human terms, indicated as explainability, but rather to the degree to which a cause and effect can be observed within a system. Given this causality stance, interpretability becomes a problematic concept [37], but it is still useful in terms of allowing the user to predict what will happen when the model input is changed. The more "transparent" the model is for the user, the more it can promote a "human-in-the-loop approach" that "black-box" (non-interpretable) techniques cannot. Techniques based on linear multivariate regression and regression trees are considered interpretable, whereas others, such as random forests or neural networks (or other ML methods), are "black-box," even if the mechanics of their computation are explainable. As an example of largely diffuse interpretable approach, variable-based degree-days regression (indicated frequently as change-point methods), originally proposed by Kissock et al. in the Inverse Modeling Toolkit (IMT) [38], has been included in ASHRAE 14:2014 [39] and has been evolving steadily over time with the introduction of algorithmic techniques for the selection of base temperatures [40], up to the explicit solution for the three parameters case [41]. In general, as evidenced by recent research, interpretable regression-based methods are versatile enough to be used for a variety of purposes throughout the building life cycle [42]. In this study the Time Of Week and Temperature (TOWT) model [43–45] has been used, among the possible interpretable regression-based methods, which was implemented in the software RMV2.0 [46] using the R programming language. The model input consists of an hourly or sub-hourly time series of temperature and load profiles. The time stamp of the series determines the weekday, and additional variables are included in the regression model's construction (that is, differentiating each day and hour of the week to capture specific recurring weekly patterns of operation). Finally, the temperature dependence of electric load is taken into account by dividing load data into temperature bins. It can handle as well Non Routine Events (NRE), i.e unpredicted changes in occupancy and operation strategies, in its software implementation in RMV2.0 [46], using methods developed by Killick [47,48]. In terms of model performance (i.e., the model's ability to approximate real behavior), the selected statistical performance indicators are namely Normalized Mean Bias Error, *NMBE*, and Coefficient of Variation of Root Mean Square Error, *CV(RMSE)*, as used by state-of-the-art protocols for Measurement and Verification (M&V), ASHRAE 14:2014 [39], Efficiency Value Organization (EVO) [49], Federal Energy Management Program (FEMP) [50]. In addition, the coefficient of determination *R2* as defined in ISO 50006:2014 [51] has been included. In the next Section the TOWT model characteristics are described in more detail.

# Methodology

The model formulation (Section 4.1) and model calibration criteria (Section 4.2) are discussed in greater detail below, providing the necessary background to understand the key characteristics of the trained models as well as the criteria used to assess their acceptability.

## Model formulation

In this study, the Time Of Week and Temperature [43–45] model formulation in its implementation of the software RMV2.0 [46] has been used. The model was chosen first for its interpretability, which is determined by a regression-based formulation with the use of additional variables defined based on a weekly schedule (that is, differentiating each day and hour of the week to capture specific recurring weekly patterns of operation) and temperature, to account for electric loads determined by heating and cooling services (i.e. temperature dependent component of energy consumption). The model input is very simple based on this formulation, as it requires simply a hourly or sub-hourly dataset with timestamp, temperature and energy (the time of week is determined by the time stamp of the series and the temperature data are binned according to the algorithm implementation). In particular the model is formulated as follows:

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  |  |

While the model formulation seems quite clearly understandable (the first part of the formulas 1 and 2 is the time of week component, while the other is the temperature components, using binned temperature data), there may be differences in the implementation in different software tools and peculiar characteristics to be highlighted. For example, the criterion to separate the model between occupied and unoccupied hours (i.e. formula 1 or 2) depends on an occupancy threshold factor that separate the hours with greater consumption from the ones with lower consumption, identified as unoccupied. The threshold can be left as default or changed, depending on the software implementation chosen. The time of week variable (*tow,j*) is a binary variable (or dummy variable, interaction factor, ecc.), *n-1* is the number of hours of the week (i.e. 168-1=167), the last term (168th) is included in the intercept term *c0*. Temperature variable *T(i)k* is a continuous variables with arbitrary temperature scale (clearly the chosen temperature scale influences the dimension of the coefficient *bk* estimated by the model, in our case we use degree celsius), *m* is the number of segments chosen when binning the temperature data (i.e. *m-1* change point sfor the piecewise linear function used to represent the temperature response component). In order to ensure continuity in the temperature response component of the function (piecewise linear), temperature segments are calculated using the algorithm described in [44]; the number of temperature segment in the original implementation was 6 , but more recent implementations give the possibility to automate the choice and to subset temperature intervals with equal width or equal number of data points. Alternatively, the user could specify the change-points for the piecewise linear function as an input to the algorithm, but this should be done manually. A potential advantage of TOWT over change-point methods is that the user doesn’t need to specify the change-point of the piecewise linear function. However, if reformulated TOWT temperature response could become comparable, at least for the temperature response component, to the one provided by a change-point method, if similar change-points values are specified. Finally, only one hyper-parameter for model tuning is present in RMV2.0 implementation, represented by the time scale of the weighting function (i.e. the number of days nearby the predicted day that are used for weighting). The weighting approach is not reported for simplicity in the formulas 1 and 2.

## Model calibration criteria

As introduced in Section 3, M&V approaches [39,49,50] use statistical indicators such as Normalized Mean Bias Error, *NMBE*, and Coefficient of Variation of Root Mean Square Error, *CV(RMSE)* to provide acceptability thresholds for models, which are summarized in Table 3. In this case, hourly resolutions is used for data for having the corresponding value. The coefficient of determination *R2* (defined in the range of 0-100 percent, or 0-1) is considered, as reported in ISO 50006:2014 [51], but acknowledging its limitations. Its limitations stem from the fact that *R2* is inherently related to the model's slope (i.e. the dependence on input variables). Model with higher slope value have a higher *R2* value even when the variance of the predicted variable is the same. In general, a *R2* value greater than 80 % indicates a good model fit, while 75 % is indicated in IPMVP Guidelines for Assessing Uncertainty as a reference value [52].

*Table 2: Thresholds of acceptability for M&V models as calibrated with monthly and hourly data*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Interval** | **Metric** | **ASHRAE Guidelines 14** | **IPMVP** | **FEMP** |
| **Monthly** | ***NMBE*** | ±5 | ±20 | ±5 |
|  | ***Cv(RMSE)*** | 15 | - | 15 |
| **Hourly** | ***NMBE*** | ±10 | ±5 | ±10 |
|  | ***Cv(RMSE)*** | 30 | 20 | 30 |

The formulas for the calculation of the statistical indicators are reported hereafter. The first indicator is the Normalized Mean Bias Error (*NMBE*) reported in Equation 1. Its positive value entails an overestimation of electricity consumption while a negative value its underestimation. It is the negative total sum of the error *E* in the time intervals, divided by the sum of the measured electricity consumption *M* expressed as a percentage.

|  |  |
| --- | --- |
|  |  |

The second indicator is the Coefficient of Variation of Root Mean Squared Error (*CV(RMSE))* reported in Equation 2. It expresses the normalized measure of the error as the ratio between the Root Mean Squared Error (*RMSE)* computed in Equation 5 as the sample deviation of the differences among measured and predicted values divided by *A* computed in Equation 6 as the average measured electricity consumption. The lower the cvRMSE value is, the more calibrated the model [15].

|  |  |
| --- | --- |
|  |  |
|  |  |
|  |  |

## Modelling workflow

In this section, the workflow is explained with the aid of flow charts and diagrams, showing the main steps of the modelling process and introducing the work performed for the case study.

*Figure 1: Flow chart of modelling approach*

# Case study description

The case study selected for this work is the Procida City Hall. It is a fully electrified building monitored for almost 4 years from February 2018 to November 2021. This period, as depicted in Table 2, corresponds to (i) its Business As Usual behaviour in terms of energy demand (period 1) followed by (ii) the very low profile due to the lockdown occurred in Italy due to the pandemic (period 2) and (iii) the new operation driven by smart working procedures with reduced occupancy (period 3).

*Table 3: Monitoring period subdivision*

|  |  |  |  |
| --- | --- | --- | --- |
| **Period** | | **Dates** | **Notes** |
| **1** | Before COVID-19 lockdown | From 01/02/2018  to 07/03/2020 | All employees work in presence according to the usual schedule |
| **2** | During COVID-19 lockdown | From 08/03/2020  to 07/10/2020 | No presence in the first 6 weeks of the period due to hard lock-down while few officers were allowed to enter to the building after that. |
| **3** | After COVID-19 lockdown | From 08/10/2020  to 01/11/2021 | General rule of remote working applies to this period with an average ratio of 50% of the employees, reduced to 25% in the last 10 weeks. |

Data are on 15-minute resolution for the electricity consumption collected by the electricity meter, later used summed to build the hourly profile to obtain the same time interval as the temperature time series. The rated power of the building electricity meter is 60 kW. The building features and its equipment have been surveyed to understand more precisely their energy use rather then their characteristics from a building physics perspective.

The building is equipped by individual electric heaters for the winter season activated by each employee in his room at his need and by few reversible split air conditioners for the summer one with the same control strategy. The climate is a Mediterranean one with mild winter and hot summer. However, the summer operation is reduced by the holiday season. Therefore, the impact of high temperature is not so strong on the building energy consumption helped also by the natural ventilation thanks to the location in front of the bay. The key characteristics of the end-uses of electricity in the building are summarized in Table 4.

*Table 4: Building electricity end-uses characteristics*

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **End-use** | **Power** | **Operation** | **Monday** | **Tuesday** | **Wednesday** | **Thursday** | **Friday** | **Saturday** | **Sunday** |
|  | **kW** | **-** | **-** | **-** | **-** | **-** | **-** | **-** | **-** |
| **Lift** | 13 | Year | On request | On request | On request | On request | On request | On request | On request |
| **Other** | 2.5 | Year | On request | On request | On request | On request | On request | On request | On request |
| **Lighting** | 6.8 | Year | 08:00-14:35 | 08:00-14:35 | 08:00-14:35 | 08:00-17:30 | 08:00-14:35 | 08:00-18:00 | 08:00-18:00 |
| **Printers** | 12.3 | Year | 08:00-14:35 | 08:00-14:35 | 08:00-14:35 | 08:00-17:30 | 08:00-14:35 | 08:00-18:00 | 08:00-18:00 |
| **Computers** | 14.7 | Year | 08:00-14:35 | 08:00-14:35 | 08:00-14:35 | 08:00-17:30 | 08:00-14:35 | 08:00-18:00 | 08:00-18:00 |
| **Split conditioners** | 8.8 | Year | 08:00-14:35 | 08:00-14:35 | 08:00-14:35 | 08:00-17:30 | 08:00-14:35 | 08:00-18:00 | 08:00-18:00 |
| **Electric water heater** | 3.6 | Year | On request | On request | On request | On request | On request | On request | On request |
| **Electric heater** | 49.5 | Winter | 08:00-14:35 | 08:00-14:35 | 08:00-14:35 | 08:00-17:30 | 08:00-14:35 | 08:00-18:00 | 08:00-18:00 |
| **Fans** | 0.7 | Summer | 08:00-14:35 | 08:00-14:35 | 08:00-14:35 | 08:00-17:30 | 08:00-14:35 | 08:00-18:00 | 08:00-18:00 |

# Results and discussion

The results of visual and numerical data analysis are reported and discussed in this Section to indicate insights that can both improve the model's current implementation and indicate future developments at the level of algorithmic formulation and graphical representation of results. First, the results of the model fitting process are reported in Section 6.1, comparing measured and predicted energy consumption for the various monitoring periods listed in Section 4 in Table 3. The impact of the hyper-parameter (time scale of the weighting function, explained in Section 4.1) choice on model goodness of fit is discussed in the same section, using the indicators listed in Table 2. The graphical interpretation of results is then proposed in Section 6.2 for the various monitoring periods (Section 4), focusing on exploratory data analysis before model fitting (Section 6.2.1) and after model fitting (Section 6.2.2). (Section 6.2.1).

In this way, the insights discovered during the exploratory phase are distinguished from those related to the model in its current implementation. The models are then subsetted and retrained, with the results discussed in Section 6.3, taking into account the model performance summarised in Section 6.1 and the visual interpretation of model fitting results in Section 6.2. Finally, Section 6.4 discusses the limitations and future research opportunities.

## Results of initial models’ fitting

The results of model testing for the different periods, indicated in Table 3, are reported in Table 5 considering energy indicators (measured and predicted energy consumption) and statistical indicators (*R2*, *NMBE*, and *CV(RMSE)* introduced in Section 4.2). It can be seen how the models fit the measured data reasonably well, with predicted electric energy consumption differing by only a few percentage points from the measured value in every monitored period. However, by examining the statistical performance (and therefore the acceptability according to M&V approaches), the impact of hyper-parameters can be seen. As Mathieau et al. [44] stated in their paper presenting the initial implementation of the TOWT algorithm, the model can easily overfit the data, reducing its ability to generalise. As a result, it is critical to investigate the effect of the hyperparameter on the model's goodness of fit. Only one hyperparameter is present in the RMV2.0 implementation, represent the time scale of weighting function, i.e. the number of days nearby the predicted day that are used for weighting. We considered three values for the hyper-parameter, namely 15 days (2 weeks), 7 days (1 week) and 1 day.

*R2* decreases significantly in period 2 and compared to period 1 if a model with hyper-parameters equal to 15 and 7 days is considered. After the lockdown (in period 3), *R2* value is still lower than period 1, despite being higher than in period 2. The values of *NMBE* are acceptable in all the periods (near 0, but not really significant we judging the performance of these type of models), but the values of *CV(RMSE)* are above the threshold for acceptability (30 %) reported in Table 3 for model calibration in most of the cases with hyper-parameters equal to 15 and 7 days, with only expection of period 1 with 7 days weighting. The highest value of *CV(RMSE)* (42.30 %) is found in period 3 with 15 days weighting, which confirm the initial guess of a less predictable behaviour, hypothesized due to the remote working introduction.

*Table 5: Results of TOWT model testing in the different monitoring periods – Hourly interval data*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Hyper-parameter** | **Energy indicators** | | **Statistical indicators** | | |
| **Monitoring**  **Period** | **Days** | **Energy measured** | **Energy predicted** | ***R2*** | ***NMBE*** | ***CV(RMSE)*** |
|  | d | kWh | kWh | % | % | % |
| **1** | 15 | 126882 | 123915 | 75.74 | 0.04 | 32.60 |
|  | 7 | 126882 | 123951 | 85.36 | 0.06 | 27.06 |
|  | 1 | 126882 | 123953 | 98.53 | 0.01 | 8.58 |
| **2** | 15 | 22363 | 22086 | 60.25 | 0.12 | 36.07 |
|  | 7 | 22363 | 22127 | 69.87 | -0.07 | 31.42 |
|  | 1 | 22363 | 22128 | 96.78 | -0.08 | 10.28 |
| **3** | 15 | 53587 | 53582 | 69.25 | -0.34 | 42.30 |
|  | 7 | 53587 | 53516 | 79.53 | -0.22 | 34.52 |
|  | 1 | 53587 | 53422 | 98.07 | -0.04 | 10.59 |

Table 6 compares average monthly electricity consumption and average power across different monitoring periods to understand the magnitude of variation in energy consumption determined by COVID-19 lockdown first and then by the new operational regime after COVID-19. The results of monitoring indicate a significant reduction (-41.2 %) of average energy consumption in period 2 (as expected due to lockdown) and a modest reduction (-16.8 %) in period 3. However, the results reported in Table 5 are not weather normalized, but only obtained by considering an equivalent average month to create a simple comparison.

*Table 6: Comparison of average monthly electricity consumption and average power across different periods*

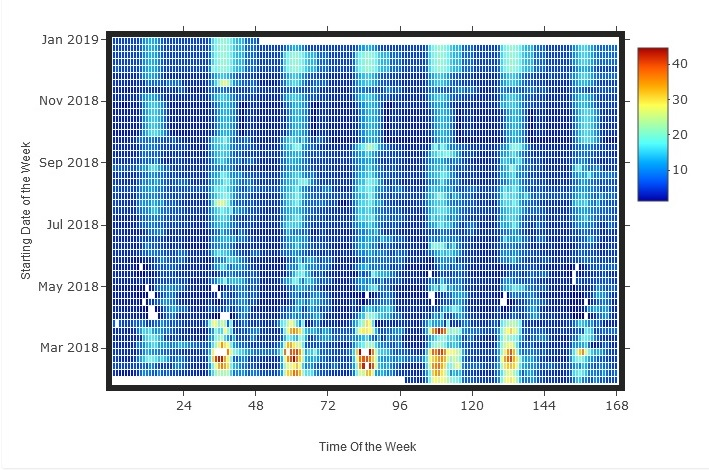
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Monitoring**  **Period** | **Description** | **Average monthly consumption** | **Average power** | **Variation with respect to baseline** |
|  |  | kWh | kW | **%** |
| **1** | Before COVID-19 lockdown | 5032 | 6.9 | 0 (baseline) |
| **2** | During COVID-19 lockdown | 2961 | 4.1 | -41.2 |
| **3** | After COVID-19 lockdown | 4189 | 5.7 | -16.8 |

## Discussion of results

Following the analysis of energy and statistical indicators reported in Section 6.1 for the various periods and models fitted, the visual interpretation of results is discussed in the following sections, which are divided into two phases, before and after model training. The heatmaps of weekly electric load patterns are reported in Section 6.2.1, along with a scatterplot aimed at identifying the temperature dependence of load. Then, the time series of measured electric load profiles is plotted against predicted data and measured temperatures in Section 6.2.2. In this case, the error term (residuals) is plotted against temperature to highlight the conditions in which models are unable to accurately estimate the temperature dependence.

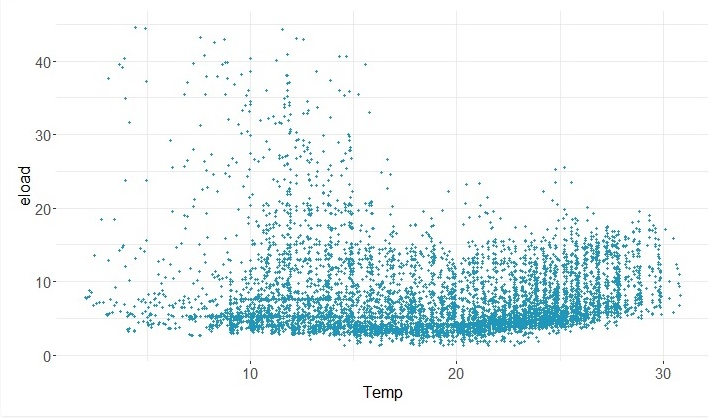
### Visual analysis before models’ training

In this section weekly operation patterns (time of week dependence) and outdoor air temperature dependence of electric load are analysed graphically before training TOWT models. First, period 1 is considered (from 01/02/2018 to 07/03/2020), and, in order to keep the amount of data in the visualization comparable to periods 2 and 3 (i.e. with a similar length of the period and, consequently, a similar size of the weekly pattern diagram), the period from February 2018 to January 2019 is plotted. For this interval, Figure 2 depicts a heatmap of the electric load (eload) expressed in kW with respect to the time of week (hours from the start of the week, plotted as multiples of 24 hours) and week starting date. In this figure, it can be noted how peak conditions are generally occurring during the central hours of the day in winter conditions (February/March 2018), whereas it is more uniform throughout the year, but with higher load values during the central hours, as expected based on the type of end-use of the building. At first glance, the patterns appear simple to predict, at least for a portion of the year (excluding the February/March 2018 peaks).



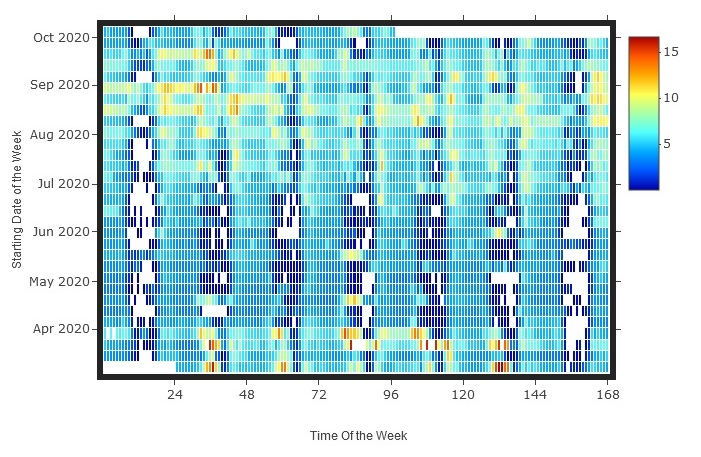
*Figure 2: Heatmap of electric load (kW) as a function of time week and starting day of the week – Period 1*

As a result, in Figure 3, the potential patterns of correlation between load (eload, expressed in kW) and outdoor air temperature (temp, expressed in °C) has been examined. As explained in Section 2 in relation to change-point models, when there is a significant portion of electric load due to heating and/or cooling services, outdoor air temperature is often the most influential variable. However, in this case, the heating demand is only sporadic, even though it determines some relevant peaks in the electric load. In fact, some load peaks with values up to around 40 kW occur at lower temperatures (as shown in Figure 2 for February/March 2018). The highest concentration of load data is in the range of 3 to 8 kW, with a relevant number of data points reaching up to 15 kW. Finally above 20 °C of outdoor air temperature, the load pattern shows a slight increase trend.



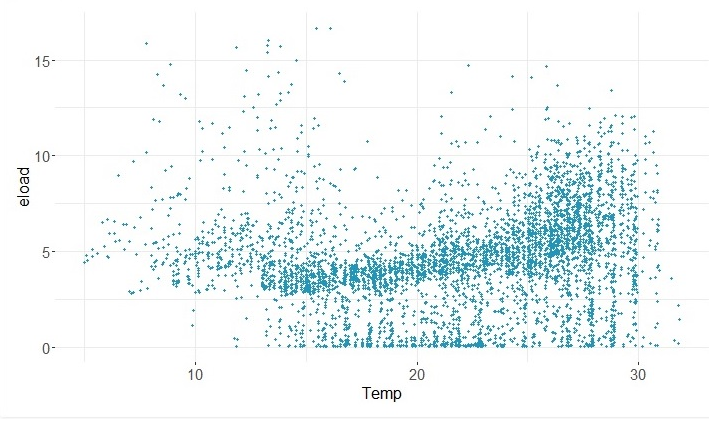
*Figure 3: Scatterplot of electric load (kW) with respect to outdoor air temperature (°C) – Period 1*

Moving to period 2 corresponding to the COVID-19 pandemic lockdown, what can be seen in Figure 4 is that there is much less variability in the electric load ranges (as expected due to the end-use of the building), but the pattern appears less uniform and less predictable. However, in this case as well, the peak data are concentrated in the middle of the day.

**

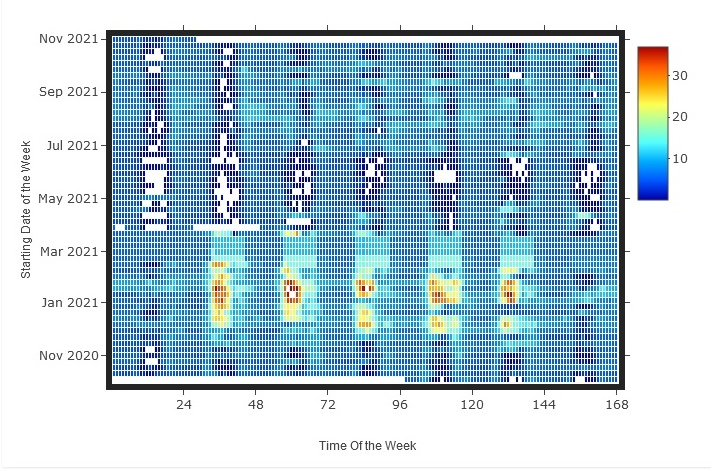
*Figure 4: Heatmap of electric load (kW) as a function of time week and starting day of the week – Period 2*

Again, by plotting the load data against the outdoor air temperature in Figure 5, we can see that the majority of the data points are in the range of approximately 2.5 kW to around 7 kW, with a modest upward trend of consumption in relation to temperatures, but no clear pattern can be detected. The spread of the cloud of data points is smaller in Figure 5 compared to Figure 3 and it is distributed in a different way in relation to temperatures. Peak conditions in the pre-pandemic period were around 40 kW, but here they are much lower, with highest values around 15 kW.



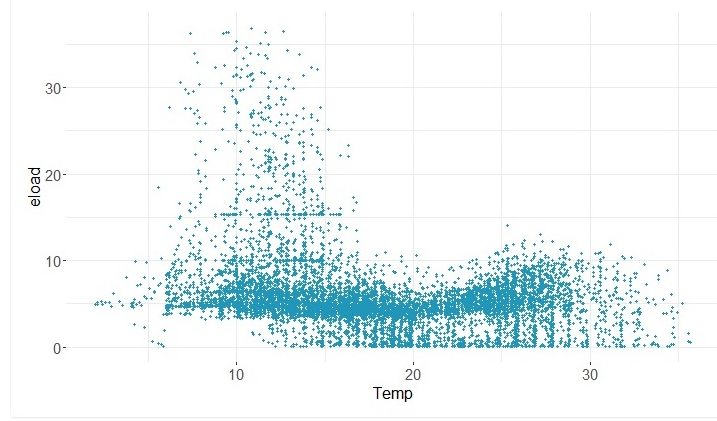
*Figure 5: Scatterplot of electric load (kW) with respect to outdoor air temperature (°C) – Period 2*

The monitoring period following COVID-19 lockdown, denoted as period 3 in Table 3 and depicted in Figure 6, shows how load peaks in winter (specifically from December 2020 to February 2021) appear during the central hours of the day during weekdays, while load is lower throughout the rest of the period. Nonetheless, unlike Period 1 (shown in Figure 1), load patterns appear less uniform and, as a result, more difficult to predict.



*Figure 6: Heatmap of electric load (kW) as a function of time week and starting day of the week – Period 3*

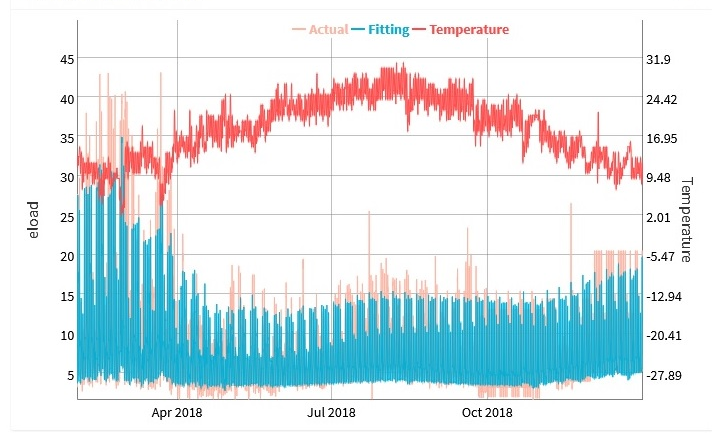
In Figure 7, the potential patterns of correlation between load and outdoor air temperature has been examined once more. The highest concentration of load data (cloud of data points) is again in the 3 to 8 kW range and shows a modest increase trend above 20 °C of outdoor air temperature. Peak loads (above 30 kW) are more evident at lower temperatures in this period than in Period 1. In other words, the behavioural anomaly identified in Figure 3 is accentuated in this case.



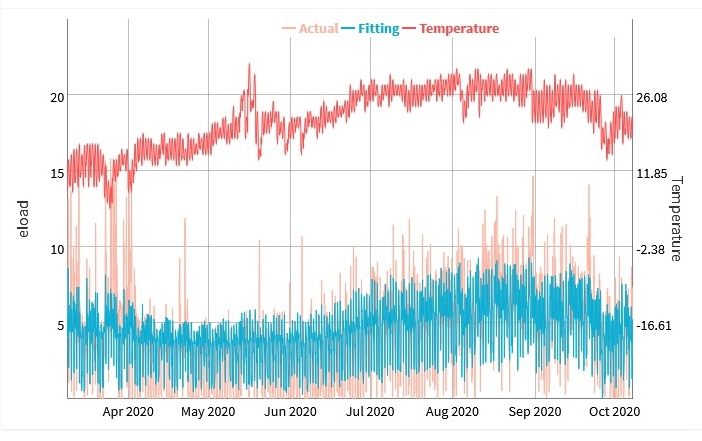
*Figure 7: Scatterplot of electric load (kW) with respect to outdoor air temperature (°C) – Period 3*

### Visual analysis after models’ training

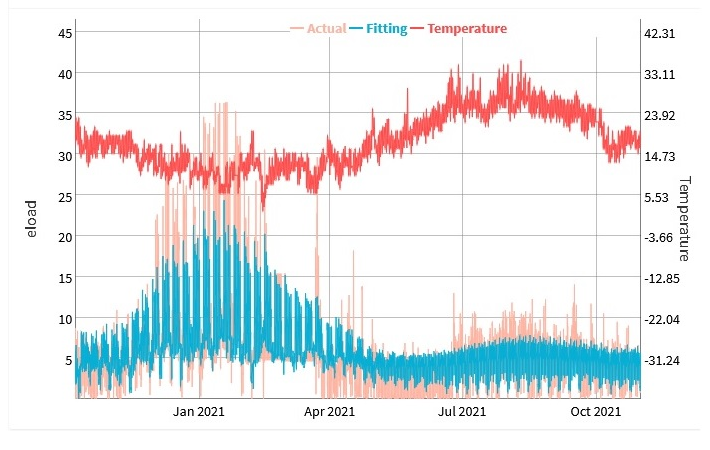
The visual analysis started in Section 6.2.1 is continued here by plotting the actual measured values with respect to model predictions as time series in Figures 9, 10, and 11, for periods 1, 2, and 3, respectively. The outside air temperature is also reported in the same charts. We consider the results of model fitting with a 15-day hyper-parameter in these figures because they highlight the differences between measured data and model predictions more clearly. TOWT models are unable to correctly fit the peak values in February/March 2018 for period 1 and from December 2020 to February 2021 for period 3, essentially the periods with the lowest temperatures, as shown in Figures 8 and 10, confirming the anomalies highlighted before model training in Figures 3 and 7 in Section 6.2.1, respectively. In general, the models seem to fit the data relatively well for non-peak periods, but the patterns of load variation are not very regular and this increases the difficulty in prediction.



*Figure 8: Time series of measured and predicted electric load (kW) with TOWT algorithm – Period 1*



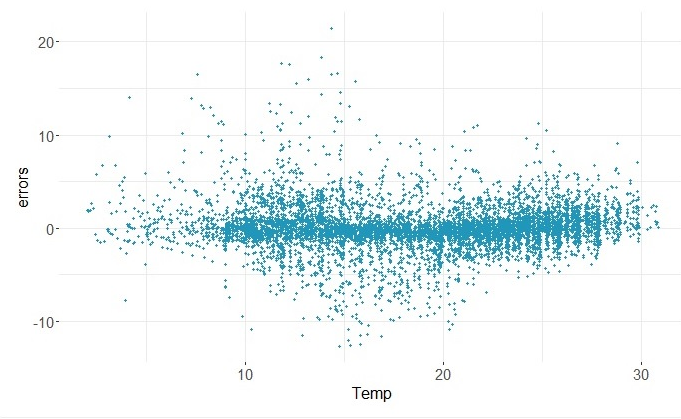
*Figure 9: Time series of measured and predicted electric load (kW) with TOWT algorithm – Period 2*



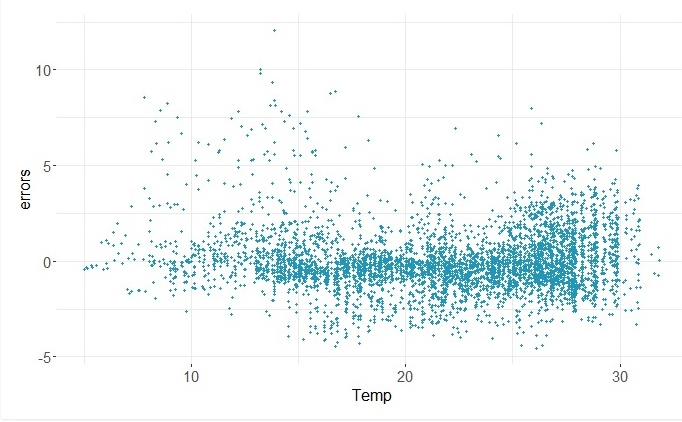
*Figure 10: Time series of measured and predicted electric load (kW) with TOWT algorithm – Period 3*

After having inspect the time series, the errors terms (residuals, differences between measured and predicted data, expressed in kW) were plotted with respect to outdoor air temperature (expressed in °C) as the final step of visual analysis. Figures 11, 12, and 13 show the results for the first, second, and third periods, respectively. What is immediately noticeable is that in period 1, the distribution is relatively uniform (as it should be when a regression model fits data correctly, residual should be homoscedastic), despite the fact that higher values are present at lower temperatures, as predicted by the visual analysis of the time series in Figure 10.

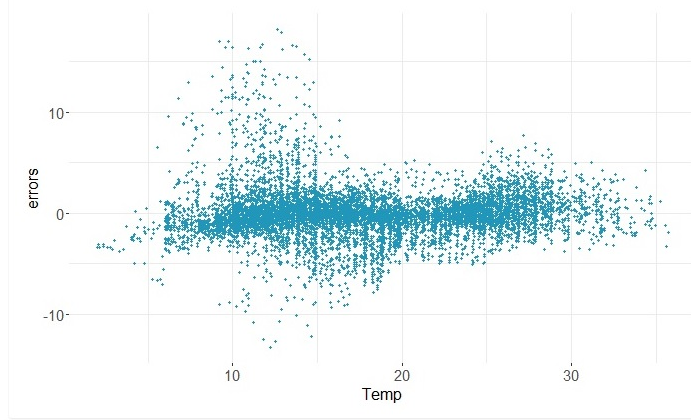
Instead, there is a concentration of errors at higher temperatures in period 2 (the cloud of data points appears to have a less uniform spread of residuals). Finally, at lower temperatures, essentially below 15 °C, error terms are more pronounced in period 3. Overall, the patterns of error terms appear to be quite different depending on the period under consideration.



*Figure 11: Scatterplot of error term (kW) with respect to outdoor air temperature (°C) – Period 1*



*Figure 12: Scatterplot of error term (kW) with respect to outdoor air temperature (°C) – Period 2*



*Figure 13: Scatterplot of error term (kW) with respect to outdoor air temperature (°C) – Period 3*

Considering the non-uniform spread of error terms for periods and 1 and 3 and the impossibility to achieve model calibration using the thresholds in Table 2 for a 15- and 7-days hyper-parameter, the analysis was re-run by sub setting the datasets for the two periods. The results are reported in Section 6.3 hereafter. Finally, a potential improvement over the current model visualisations could be the use of control charts [53] as in Statistical Process Control (SPC) [54]. This visualization can be used to analyse performance anomalies by plotting the error term against the number of standard deviations (1, 2, or 3 usually) from the mean (ideally equal to zero). The same visualisation can also make use of cumulative differences (CUSUM) [21] and additional input variables (covariates) [55] to investigate, respectively a rapid change in conditions and the impact of covariates time series anomalies.

## Results of model retraining on subsets of data

Following the analysis of fitted models and error terms in Section 6.2, which show a higher spread of data in months with lower temperatures, the dataset has been partitioned as shown in Table 7, essentially providing two subsets (corresponding to lower/higher temperatures) for monitoring periods 1 and 3 respectively. Period 2 has not been portioned further because it corresponds essentially to a warmer period, characterised by higher temperatures.

*Table 7: Monitoring period subdivision – subsets selected for model retraining*

|  |  |  |  |
| --- | --- | --- | --- |
| **Monitoring**  **Period** | | **Subset** | **Dates** |
| **1** | Before COVID-19 lockdown | Lower temperatures | From 01/02/2018  to 31/03/2018 |
|  |  | Higher temperatures | From 01/04/2018  to 30/11/2018 |
| **3** | After COVID-19 lockdown | Lower temperatures | From 01/12/2020 to 28/02/2021 |
|  |  | Higher temperatures | From 01/13/2021  to 01/11/2021 |

The TOWT models were then retrained for the various time periods, yielding the energy and statistical indicators shown in Table 8. Overall, partitioning the dataset resulted in much higher *R2* values for all subsets, greater than 75 %, even with hyper-parameter equal to 15 days, with the exception of period 3-higher temperatures, where *R2* decreases sensibly (due to the lower regularity of load patterns) to 56.65 %. Again, the *NMBE* is very small (near 0) in all the cases (therefore not really significant to judge the goodness of fit for this type of models), and *CV(RMSE)* improves slightly in all cases. The period with the highest *CV(RMSE)* is period 3-higher temperatures (in the range 34.67-9.35 %); in all other cases, the values are very close to the calibration threshold (30 % reported in Table 2 for hourly data), and the model can be accepted as calibrated in one case, period 1-higher temperatures, if we consider the hyper-parameter value of 15 days, while calibrated in all the cases with hyper-parameter equal to 7 days. As already shown in Section 6.1, with hyper-parameter equal to 1 day the models overfit data in all periods. Finally, in terms of predicted energy for the corresponding subset, all models can produce estimates that are very close to the actual measured value.

*Table 8: Results of TOWT model testing in the different periods – Subsets of the original datasets – Hourly interval*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | **Hyper-parameter** | | **Energy indicators** | | **Statistical indicators** | | |
| **Monitoring**  **Period** | **Subset for model retraining** | | **Days** | **Energy measured** | **Energy predicted** | ***R2*** | ***NMBE*** | ***CV(RMSE)*** |
|  |  | | d | kWh | kWh | **%** | **%** | **%** |
| **1** | Lower temperatures | | 15 | 15811 | 15864 | 86.34 | -0.38 | 30.93 |
|  |  | | 7 | 15811 | 15847 | 89.25 | -0.27 | 27.44 |
|  |  | | 1 | 15811 | 15806 | 98.97 | -0.01 | 8.51 |
|  | Higher temperatures | | 15 | 38281 | 37990 | 78.96 | 0.11 | 25.10 |
|  |  | | 7 | 38281 | 38008 | 83.98 | 0.06 | 21.9 |
|  |  | | 1 | 38281 | 38033 | 98.35 | 0.01 | 7.03 |
| **3** | Lower temperatures | | 15 | 21056 | 21101 | 79.42 | -0.25 | 31.82 |
|  |  | | 7 | 21056 | 21079 | 84.56 | -0.15 | 27.55 |
|  |  | | 1 | 21056 | 21052 | 98.63 | -0.02 | 8.21 |
|  | Higher temperatures | | 15 | 26511 | 26131 | 56.65 | 0.20 | 34.67 |
|  |  | | 7 | 26511 | 26197 | 69.17 | -0.06 | 29.24 |
|  |  | | 1 | 26511 | 26194 | 96.85 | -0.05 | 9.35 |

## Limitations and further research

In general, the results of visual and numerical analysis of time series and TWOT models applied in different periods show that the pre-COVID-19 (period 1) load profiles were more predictable than the post-COVID-19 (period 3) ones. The COVID-19 lockdown (period 2) was an outlier in itself, and we expected less predictable conditions and lower energy demand as a result of remote working and shorter operating hours. Due to irregular operating schedules, the period with less predictable operational profiles was the one post-COVID-19 (period 3) with higher outdoor air temperatures. In terms of visual analysis, statistical process control charts (SPC) are a simple but useful technique for understanding anomalies because they can be used to plot the error term (residuals) in the time series and check the distance from the mean (ideally equal to zero for residuals) in terms of standard deviations. The ability to include additional input variables in a SPC control chart could greatly improve the model's ability to identify anomalies in a visual but also automated way.

Although the TWOT model was chosen for its simplicity and ease of use, its applicability and performance could be improved also by automating further the process of hyper-parameter tuning and subsetting of the original time series, in case of relevant changes in the operating schedules of the building. This is relatively simple to do by after a visual inspection of data, but more difficult to automate because the models must be retrained iteratively and criteria or thresholds have to be defined for model acceptability. Indeed, the choice of hyper-parameter value was shown to be extremely relevant and the risk of overfitting conditions that are too specific is evident, as already indicated in the first implementation of the method by Mathieau et al. [44].

Overall, by improving the method's interpretability, it may be possible to access two fundamental components of building behaviour visually, the time of week (schedule of operation) and the temperature response (piecewise linear function). This necessitates rewriting the algorithm in order to obtain an approximated physical interpretation for the model parameters that can be visualised and easily understood in human terms. In this way, the user can see whether the model parameters are overfitting a specific condition or useful for generalising model findings.

Furthermore, the algorithm could be reformulated so that at least some of the model parameters are comparable to those obtained using other methods, such as change-point methods (e.g. by defining comparable criteria for the definition of temperature segments, as discussed in Section 4.1). Finally, as previously demonstrated in research on change-point methods, additional variables (other than the time series stamp and temperature as in its current formulation) can be included to improve performance while retaining interpretability and ease of use (i.e. only a few hyper-parameters to be tuned).

# Conclusions

The goal of this study was to put the Time Of Week and Temperature (TOWT) modelling approach to the test on a case study building that was monitored before, during, and after the COVID-19 lockdown. The Procida City Hall was chosen as the case study building. It is a completely electrified building that has been monitored for nearly 4 years, from February 2018 to November 2021. The COVID-19 pandemic clearly affected the building's occupancy levels dramatically, with periods characterised by the total absence of employees to periods characterised by a fixed percentage of occupants.

The case study is thus challenging, and the TWOT technique was chosen for its simplicity (2 input variables, time stamp and temperature, 1 output variable, electric load) and ease of use, with only one hyper-parameter for model tuning, represented by the time scale of weighting function (i.e. the number of days nearby the predicted day that are used for weighting). Furthermore, the TOWT technique was chosen for its interpretability, which is determined by the ability to represent the electric load on a weekly schedule basis (i.e. differentiating each day and hour of the week to capture specific recurring weekly patterns of operation) and temperature, which accounts for electric loads determined by heating and cooling services (i.e. temperature dependent component of energy consumption).

The specific objectives of the research included testing model fitting performance in challenging conditions (i.e. during and immediately after the COVID-19 pandemic period in particular), understanding critical aspects for model fitting (i.e. hyper-parameter tuning and dataset subsetting), and understanding areas of improvement for visualisation and algorithmic formulation, with interpretability as the focus, for the reasons outlined in the introduction and literature review.

Using the initial dataset subdivision (before, during and after lockdown), the results of model fitting, presented in Section 6.1, revealed that calibration according to M&V protocols is not possible if a hyper-parameter of 15 days is considered, and only possible with 7 days for the pre-pandemic period, which is characterised by more predictable operating schedules. Using one day as the hyper-parameter causes the model to overfit for all periods and, consequently, reduce its potential for generalisation.

Following the initial model fitting process, we visually analysed the data in Section 6.2 and defined a criterion to subset the original datasets based on periods with higher and lower temperatures. Following that, in Section 6.3, the datasets were subset based on lower and higher temperature periods, and the models were retrained, yielding significantly better results. Following data sub setting, all models are calibrated with a hyper-parameter of 7 days. Using one day as the hyper-parameter causes the models to overfit for all periods in this case as well. Data analysis for the different time frames suggested numerous improvements, both visually and algorithmically. In terms of visualisation, a simple improvement could be visualising the error term in the time series using statistical process control (SPC) charts, which could immediately flag performance anomalies.

The ability to automate the process of hyper-parameter tuning and time series sub setting would be extremely beneficial from an algorithmic standpoint, ideally guiding the user in the search for the best compromise between fitting performance and generalisation.

Following that, by improving the method's interpretability (i.e., improving the method's "human-in-the-loop" capabilities), it may be possible to visually understand two fundamental components of building behaviour, the time of week (schedule of operation) and the temperature response (piecewise linear function).

This necessitates a partial reformulation of the algorithm, which could then be done in such a way that some of the model parameters are at least comparable to those obtained using other methods, such as those specified in Section 6.4. Finally, additional variables can be added to improve performance while maintaining interpretability and usability. As a result, all of the suggested improvements can improve the applicability and performance of the TOWT approach while retaining its simplicity and ease of use and will be included in future research.

# References

[1] B. Norton, W.B. Gillett, F. Koninx, Decarbonising Buildings in Europe: A Briefing Paper, Proc. Inst. Civ. Eng. - Energy. 0 (n.d.) 1–18. https://doi.org/10.1680/jener.21.00088.

[2] P.C. Slorach, L. Stamford, Net zero in the heating sector: Technological options and environmental sustainability from now to 2050, Energy Convers. Manag. 230 (2021) 113838. https://doi.org/https://doi.org/10.1016/j.enconman.2021.113838.

[3] F. Calise, F.L. Cappiello, M. Dentice d’Accadia, M. Vicidomini, Smart grid energy district based on the integration of electric vehicles and combined heat and power generation, Energy Convers. Manag. 234 (2021) 113932. https://doi.org/https://doi.org/10.1016/j.enconman.2021.113932.

[4] M. Manfren, N. Aste, F. Leonforte, C. Del Pero, M. Buzzetti, R.S. Adhikari, L. Zhixing, Parametric energy performance analysis and monitoring of buildings-HEART project platform case study, Sustain. Cities Soc. (2020) 102296.

[5] M. Dongellini, C. Naldi, G.L. Morini, Influence of sizing strategy and control rules on the energy saving potential of heat pump hybrid systems in a residential building, Energy Convers. Manag. 235 (2021) 114022. https://doi.org/https://doi.org/10.1016/j.enconman.2021.114022.

[6] M. Ala-Juusela, T. Crosbie, M. Hukkalainen, Defining and operationalising the concept of an energy positive neighbourhood, Energy Convers. Manag. 125 (2016) 133–140. https://doi.org/https://doi.org/10.1016/j.enconman.2016.05.052.

[7] A. Kona, P. Bertoldi, Ş. Kılkış, Covenant of Mayors: Local Energy Generation, Methodology, Policies and Good Practice Examples, Energies . 12 (2019). https://doi.org/10.3390/en12060985.

[8] Ş. Kılkış, G. Krajačić, N. Duić, M.A. Rosen, M. Ahmad Al-Nimr, Accelerating mitigation of climate change with sustainable development of energy, water and environment systems, Energy Convers. Manag. 245 (2021) 114606. https://doi.org/https://doi.org/10.1016/j.enconman.2021.114606.

[9] S. Zhang, X. Ma, Q. Cui, Assessing the Impact of the Digital Economy on Green Total Factor Energy Efficiency in the Post-COVID-19 Era , Front. Energy Res. . 9 (2021) 808. https://www.frontiersin.org/article/10.3389/fenrg.2021.798922.

[10] H. ur Rehman, T. Korvola, R. Abdurafikov, T. Laakko, A. Hasan, F. Reda, Data analysis of a monitored building using machine learning and optimization of integrated photovoltaic panel, battery and electric vehicles in a Central European climatic condition, Energy Convers. Manag. 221 (2020) 113206. https://doi.org/https://doi.org/10.1016/j.enconman.2020.113206.

[11] L. Zhang, H. Su, E. Zio, Z. Zhang, L. Chi, L. Fan, J. Zhou, J. Zhang, A data-driven approach to anomaly detection and vulnerability dynamic analysis for large-scale integrated energy systems, Energy Convers. Manag. 234 (2021) 113926. https://doi.org/https://doi.org/10.1016/j.enconman.2021.113926.

[12] National Institute of Building Sciences (NIBS) New buildings Institute, national environmental balancing bureau. Data needs for achieving high performance buildings: high-performance building data collection initiative. Washington, DC (2011), (n.d.). http://www.businessperformance.org/sites/default/files/NIBS\_DataCollectionReport.pdf.

[13] F. Mancini, G. Lo Basso, How Climate Change Affects the Building Energy Consumptions Due to Cooling, Heating, and Electricity Demands of Italian Residential Sector, Energies . 13 (2020). https://doi.org/10.3390/en13020410.

[14] E. Fabrizio, M. Filippi, J. Virgone, An hourly modelling framework for the assessment of energy sources exploitation and energy converters selection and sizing in buildings, Energy Build. 41 (2009) 1037–1050. https://doi.org/https://doi.org/10.1016/j.enbuild.2009.05.005.

[15] S. Wang, Y. Zang, W. Ge, A. Wang, D. Li, J. Tang, Data-Driven Real-Time Pricing Strategy and Coordinated Optimization of Economic Load Dispatch in Electricity Market , Front. Energy Res. . 9 (2021) 434. https://www.frontiersin.org/article/10.3389/fenrg.2021.714951.

[16] J. Zhang, H. Zhang, S. Ding, X. Zhang, Power Consumption Predicting and Anomaly Detection Based on Transformer and K-Means , Front. Energy Res. . 9 (2021) 681. https://www.frontiersin.org/article/10.3389/fenrg.2021.779587.

[17] N. Fumo, P. Mago, R. Luck, Methodology to estimate building energy consumption using EnergyPlus Benchmark Models, Energy Build. 42 (2010) 2331–2337. https://doi.org/https://doi.org/10.1016/j.enbuild.2010.07.027.

[18] M. Lamagna, B. Nastasi, D. Groppi, M.M. Nezhad, D.A. Garcia, Hourly energy profile determination technique from monthly energy bills, Build. Simul. 13 (2020) 1235–1248. https://doi.org/10.1007/s12273-020-0698-y.

[19] E. Zaidan, A. Ghofrani, E. Dokaj, Analysis of Human-Building Interactions in Office Environments: to What Extent Energy Saving Boundaries can be Displaced? , Front. Energy Res. . 9 (2021) 450. https://www.frontiersin.org/article/10.3389/fenrg.2021.715478.

[20] I. Gaetani, P.-J. Hoes, J.L.M. Hensen, Occupant behavior in building energy simulation: Towards a fit-for-purpose modeling strategy, Energy Build. 121 (2016) 188–204. https://doi.org/https://doi.org/10.1016/j.enbuild.2016.03.038.

[21] A. Fichera, R. Volpe, E. Cutore, Energy performance measurement, monitoring and control for buildings of public organizations: Standardized practises compliant with the ISO 50001 and ISO 50006, Dev. Built Environ. 4 (2020) 100024. https://doi.org/https://doi.org/10.1016/j.dibe.2020.100024.

[22] C. Fan, M. Chen, X. Wang, J. Wang, B. Huang, A Review on Data Preprocessing Techniques Toward Efficient and Reliable Knowledge Discovery From Building Operational Data , Front. Energy Res. . 9 (2021) 77. https://www.frontiersin.org/article/10.3389/fenrg.2021.652801.

[23] H. Zhao, J. Zhao, T. Shu, Z. Pan, Hybrid-Model-Based Deep Reinforcement Learning for Heating, Ventilation, and Air-Conditioning Control , Front. Energy Res. . 8 (2021) 412. https://www.frontiersin.org/article/10.3389/fenrg.2020.610518.

[24] B. Grillone, G. Mor, S. Danov, J. Cipriano, A. Sumper, A data-driven methodology for enhanced measurement and verification of energy efficiency savings in commercial buildings, Appl. Energy. 301 (2021) 117502. https://doi.org/https://doi.org/10.1016/j.apenergy.2021.117502.

[25] T. Hong, Z. Wang, X. Luo, W. Zhang, State-of-the-art on research and applications of machine learning in the building life cycle, Energy Build. 212 (2020) 109831. https://doi.org/https://doi.org/10.1016/j.enbuild.2020.109831.

[26] C. V Gallagher, K. Leahy, P. O’Donovan, K. Bruton, D.T.J. O’Sullivan, Development and application of a machine learning supported methodology for measurement and verification (M&V) 2.0, Energy Build. 167 (2018) 8–22. https://doi.org/https://doi.org/10.1016/j.enbuild.2018.02.023.

[27] A. Chong, Y. Gu, H. Jia, Calibrating building energy simulation models: A review of the basics to guide future work, Energy Build. 253 (2021) 111533. https://doi.org/https://doi.org/10.1016/j.enbuild.2021.111533.

[28] H.J. Kang, Development of an Nearly Zero Emission Building (nZEB) Life Cycle Cost Assessment Tool for Fast Decision Making in the Early Design Phase, Energies . 10 (2017). https://doi.org/10.3390/en10010059.

[29] S. Song, C.G. Park, Alternative Algorithm for Automatically Driving Best-Fit Building Energy Baseline Models Using a Data—Driven Grid Search, Sustain. . 11 (2019). https://doi.org/10.3390/su11246976.

[30] I. Ridwana, N. Nassif, W. Choi, Modeling of Building Energy Consumption by Integrating Regression Analysis and Artificial Neural Network with Data Classification, Build. . 10 (2020). https://doi.org/10.3390/buildings10110198.

[31] S. Ha, S. Tae, R. Kim, Energy Demand Forecast Models for Commercial Buildings in South Korea, Energies . 12 (2019). https://doi.org/10.3390/en12122313.

[32] E. Burman, S.-M. Hong, G. Paterson, J. Kimpian, D. Mumovic, A comparative study of benchmarking approaches for non-domestic buildings: Part 2 – Bottom-up approach, Int. J. Sustain. Built Environ. 3 (2014) 247–261. https://doi.org/https://doi.org/10.1016/j.ijsbe.2014.12.001.

[33] L.A. Bollinger, C.B. Davis, R. Evins, E.J.L. Chappin, I. Nikolic, Multi-model ecologies for shaping future energy systems: Design patterns and development paths, Renew. Sustain. Energy Rev. 82 (2018) 3441–3451. https://doi.org/https://doi.org/10.1016/j.rser.2017.10.047.

[34] M. Manfren, M. Sibilla, L. Tronchin, Energy Modelling and Analytics in the Built Environment—A Review of Their Role for Energy Transitions in the Construction Sector, Energies . 14 (2021). https://doi.org/10.3390/en14030679.

[35] ISO 50001:2018, Energy management systems - Requirements with guidance for use, (2018).

[36] ISO/IEC TR 29119-11:2020(en) Software and systems engineering — Software testing — Part 11: Guidelines on the testing of AI-based systems, (n.d.).

[37] Z.C. Lipton, The Mythos of Model Interpretability: In machine learning, the concept of interpretability is both important and slippery., Queue. 16 (2018) 31–57.

[38] J.K. Kissock, J.S. Haberl, D.E. Claridge, Inverse modeling toolkit: numerical algorithms, ASHRAE Trans. 109 (2003) 425.

[39] ASHRAE, ASHRAE Guideline 14-2014: Measurement of Energy, Demand, and Water Savings; American Society of Heating, Refrigerating and Air-Conditioning Engineers: Atlanta, GA, USA, 2014., (2014).

[40] M.T. Paulus, D.E. Claridge, C. Culp, Algorithm for automating the selection of a temperature dependent change point model, Energy Build. 87 (2015) 95–104. https://doi.org/https://doi.org/10.1016/j.enbuild.2014.11.033.

[41] M.T. Paulus, Algorithm for explicit solution to the three parameter linear change-point regression model, Sci. Technol. Built Environ. 23 (2017) 1026–1035.

[42] M. Manfren, B. Nastasi, L. Tronchin, Linking Design and Operation Phase Energy Performance Analysis Through Regression-Based Approaches, Front. Energy Res. 8 (2020) 288. https://doi.org/10.3389/fenrg.2020.557649.

[43] P. Price, Methods for analyzing electric load shape and its variability, Lawrence Berkeley National Laboratory Report LBNL-3713E, 2010.

[44] J.L. Mathieu, P.N. Price, S. Kiliccote, M.A. Piette, Quantifying changes in building electricity use, with application to demand response, IEEE Trans. Smart Grid. 2 (2011) 507–518.

[45] S.D. Borgeson, Targeted efficiency: Using customer meter data to improve efficiency program outcomes, University of California, Berkeley, 2013.

[46] RMV2.0 - LBNL M&V2.0 Tool (https://lbnl-eta.github.io/RMV2.0/), (n.d.).

[47] R. Killick, P. Fearnhead, I.A. Eckley, Optimal Detection of Changepoints With a Linear Computational Cost, J. Am. Stat. Assoc. 107 (2012) 1590–1598. https://doi.org/10.1080/01621459.2012.737745.

[48] R. Killick, K. Haynes, I.A. Eckley, {changepoint}: An {R} package for changepoint analysis, J. Stat. Softw. 58 (2014) 1–19. http://www.jstatsoft.org/v58/i03/.

[49] EVO, 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, (2003).

[50] FEMP, 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, (2008).

[51] ISO 50006:2014, Energy management systems — Measuring energy performance using energy baselines (EnB) and energy performance indicators (EnPI) — General principles and guidance, (2014).

[52] Uncertainty Assessment for IPMVP, International Performance Measurement & Verification Protocol, Efficiency Valuation Organization (EVO), (n.d.).

[53] M. Flores, S. Naya, R. Fernández-Casal, S. Zaragoza, P. Raña, J. Tarrío-Saavedra, Constructing a Control Chart Using Functional Data, Math. . 8 (2020). https://doi.org/10.3390/math8010058.

[54] L.C. Braga, A.R. Braga, C.M.P. Braga, On the characterization and monitoring of building energy demand using statistical process control methodologies, Energy Build. 65 (2013) 205–219. https://doi.org/https://doi.org/10.1016/j.enbuild.2013.05.002.

[55] F. Centofanti, A. Lepore, A. Menafoglio, B. Palumbo, S. Vantini, Functional Regression Control Chart, Technometrics. 63 (2021) 281–294. https://doi.org/10.1080/00401706.2020.1753581.

1. \* Corresponding author:

   e-mail: benedetto.nastasi@outlook.com [↑](#footnote-ref-1)