Dipartimento di Ingegneria Industriale DII

Corso di Laurea magistrale in Ingegneria Energetica

Analysis of data-driven models for energy consumption in buildings

Relatore: Prof. Angelo Zarrella
Correlatore: Roberto Garay-Martinez

Studente: Borgato Nicola
Matricola: 2028906

Anno accademico 2022/2023

Abstract

Building sector in Europe is responsible for the 40% of the total energy consumption and it produces the 36% of all energy-related greenhouse gas emissions. Improving the efficiency of all types of buildings (residential ones, schools, hotels, offices…) is a crucial point for reaching goals established by EU in terms of reduction of energy consumption and CO_2. This thesis contributes to improve the possibility of studying this sector, providing a method for obtaining a large amount of information about the structure analysed from the analysis of its energy consumption. Results obtained from the application of methodology proposed in this study can be used to improve knowledge about building's behaviour. This fact produces benefits to users, house companies and manager of district heating networks.

The methodology proposed consists of a set of different tools useful to analyse the building's behaviour in terms of energy consumption. Starting just from hourly measured value of energy demand and hourly values of outdoor temperature, this algorithm can obtain a large set of information about building's activity. Energy demand over a year is studied divided by hour (creating 24 different analyses, one per each hour of the day) and all together. There is the creation of the energy signature curve (ESC) with the consequent individuation of the change point temperature (CPT), coefficient of global heat losses (Q_{tot}), losses in DHW circuit. There is the splitting of global consumption into the component related to space heating (SH) and domestic hot water (DHW) heat loads. In addition to that, also a method to obtain information about thermal inertia of the structure is proposed.

Furthermore, there is the study of boxplots of consumption considering average yearly profile of the entire week and Saturdays, Sundays, and workdays separately. There is the analysis of summer consumption, where only DHW heat load is present. About DHW profile, a detailed analysis in summer period is performed. It was possible to find similarities among different days and to individuate hours of minimum, maximum and peak load, calculating boundary values in terms of demand between these three zones. A method for improving accuracy of DHW profile in winter using the detailed analysis performed in summer is proposed and applied.

All the analysis is implemented in MATLAB, and everything is tested into two different case studies. The first one is about 51 buildings connected in a district heating network (DHN) in Tartu. In this dataset every building presents a complete yearly demand measurements (8760 values) and no other information are known. The second case study examines 26 buildings belonging to a DHN in historic centre of Verona. The particularity of this second case is that every building has a partial dataset, where consumptions of just few months in winter season are available. For this reason, it was necessary to develop a method to also analyse this case.

In the final part there is the presentation of results of three buildings from Tartu's dataset and two from the Verona's one. They are the most significant cases among all analysed structures. In that chapter there are, also, comments and considerations about all possible information that can be obtained. In general, thanks to this methodology it is possible to obtain a large set of useful data about the building. In the appendix it is reported the used code. It is completely automatic and can be applied to every building.

Index

Chapter 2: Introduction.

This work is focused on building sector. Energy savings in buildings are critical for the European Union's (EU) power system, as they have significant economic and environmental implications, as emerged from the study conducted by Thonipara A. et al. Currently, buildings account for 40% of the total energy consumption in the EU, according to the performance of buildings directive redacted by the European Commission. This means that they are the primary energy consumer in Europe, representing the largest share of energy consumption.

Continuing talking about numbers, according to a Eurostat's study from 2018, residential buildings alone account for around 25% of total energy consumption and about 20% of greenhouse gas emissions. If all buildings in Europe are taken into consideration, the percentage of energy-related greenhouse gas emissions raises up to 36%. As a result, achieving a highly energy-efficient building stock is one of the main goals of the current energy policies in the EU.

Among all the technical systems in buildings, space heating (SH) and domestic hot water (DHW) are the most significant energy consumers. According to the research work of Pezzutto S. et al., SH and DHW together account for over 20% of the total energy utilization in the EU, with SH being responsible for approximately 85% of the heat demand in the EU and DHW accounting for the remaining 15%. According to another study from the European Commission from 2019, heating, cooling, and domestic hot water account for 80% of the energy consumed by citizens. Therefore, improving energy efficiency in SH and DHW systems is crucial for achieving the EU's energy targets. From these considerations it was born the idea to focus the attention of the work proposed in this thesis on the buildings sector.

In addition to that, there is the problem that most buildings in the EU are currently energy inefficient. In fact, around 35% of the building stock is over 50 years old, and almost 75% of buildings are not energy efficient, according to the study about energy efficiency in buildings, conducted by European Parliament in 2021. Additionally, only a small percentage of buildings (approximately 1%) are renovated each year, indicating a slow rate of progress in achieving energy efficiency targets.

From these data it was decided to give a contribution on the analysis of building sector, in particular existing buildings. The goal is to provide tools and to develop a method that starting from few easily obtainable data is able to give detailed information about building's behaviour and consumption. The other aim is to guarantee a large field of applicability, avoiding creating a method too specific for a certain type of building.

A big part of the analysis is about obtaining, and study domestic hot water and space heating heat loads for every building. The information obtained from the code developed in this thesis can be used to identify potential opportunities for further improvements of building sector, to help house companies, DHN manager and users to administrate building in a better way and can be useful to provide insights into the key factors affecting the adoption of energy-efficient practices in this sector.

The contribution of this thesis to the state of art regarding these issues regards various aspects. First, there is the union of different tools and method developed for different buildings and climates than the ones analysed in the proposed case studies. Second, there is the development of a methodology to improve the existing model for splitting SH and DHW using an innovative approach to DHW profile. Third, there is the adaptation and application of these methods to a partial database, a case not considered in literature. Fourth, there is a wide discussion of all possible outputs and information obtainable from the tools applied. Fifth, there are other smaller improvements, highlighted in the following chapters.

The thesis is articulated into 9 chapters. The first and the second ones are dedicated to the abstract and the general introduction of the work. Here there are presented some information about building sector and the general aim of the work.

In chapter 3 there is the literature review. Here there is the review and the presentation of scientific articles and studies that talk about tools used to build up the code for the analysis. This chapter is divided into two main parts. The first one provides a general brief about distinct types of approach in modelling heat load of the building, focusing the attention on the statistical one, used in this thesis. The second part presents articles that talk about every single tool used in this analysis.

Chapter 4 is dedicated to the presentation of the methodology followed for the analysis of building with an entire year database. There are explained all equations and mathematical tools used for the analysis. Chapter 5 is about the same topic, but in case there is a building with incomplete database, where only few months of measurements are available.

In chapter 6 there is a concise description of case studies analysed, Tartu and Verona. Here some general information about the district heating and the meters used to obtain data are given.

Results and comments are present in chapter 7. In this site there are the presentations of three buildings of Tartu and two of Verona. There is the explanation of all outputs of the code, and it was provided an interpretation of results for every building. Decided to present just some buildings. The decision was to show one standard building for both databases and to present some exceptional cases worthy of attention.

Chapter 8 and 9 are dedicated to a brief show of future possibility and to conclusions of the work. After them there is the appendix, where it is possible to find every function and, also, the general code implemented in MATLAB to perform the analysis. They are subdivided into two categories: the first one refers to the complete database code and the second one to the partial database.

Chapter 3: Literature review

In this chapter it will be presented the state of art of tools for buildings' energy consumptions analysis.

In this work it was presented a relevant number of tools and methodologies for studying thermal energy demand in buildings and it is necessary to present a global overview to understand the starting point and to appreciate the improvement done in this work.

In the following paragraphs there is a literature review for every tool used in the thesis. In the first section it will be a discussion about models for estimating buildings' thermal energy consumption. In the second section it enters more in detail about the aspects studied in this thesis.

3.1 Models for estimating building's thermal energy consumption.

The use of models is a fundamental activity in the study of buildings when we talk about energy consumption. There are many types of models, built for studying every aspect of a building: from energy efficiencies to the phase of construction, passing through the detection of faults or the energy consumption. In this work the interest is about thermal energy consumption of buildings, so the attention is focused on this topic.

According to the work of review of Tanveer Ahmad et al. over 50% of the studies are focused on the models for estimating the entire energy consumption of the building (electricity, heating, cooling and others) and only 35% of works are about heating and cooling energy demand. The explanation of this phenomenon is because there is more commercial interest in models which can forecast the entire demand of a building than a partial models.

Having a dedicated model for heating and cooling, however, is better in all cases when there is a necessity, for example, to study the behaviour of an envelope in case of an operation of retrofitting. With the constant growth of ZEB or NZEB buildings, which present a lower demand for space heating compared to domestic hot water demand or respect to other types of building, it is important to have a model for the analysis of the energy demand specific for domestic hot water production.

In this work the models implemented are all used for analysis of data, but the results obtained can be used also for forecasting building behaviours in years different from the one considered or as shown in the partial database analysis section (chapter 5), to forecast energy demand and building behaviour in the same year but in a different season. So, in simple terms, all models and tools implemented and discussed in this thesis are building thermal energy prediction models.

Simulation tools have been developed since the 1990s to estimate building energy usage and these different methods can be subdivided into: engineering; Artificial Intelligence based; hybrid and data-driven approaches.

When it comes to forecasting building loads, the regression model is commonly used due to its ease of development and interpretation. Existing literature provides a basis for

categorizing models based on type, forecast horizon, and scale. There are many interesting works about this topic: the article of Catalina et al. is only one of the many examples of regression model for the evaluation and the prediction of heating energy demand based on the main factors that influence a building's heat consumption. In this thesis it was used a regression model for the analysis of the energy demand.

There are three different approaches for estimating energy consumption of a building: the so called "white-box" approach (engineering), the "grey-box" approach (hybrid) or the "black-box" approach (data-driven and AI-based). In this section of the thesis, it will be presented these three types of methods, with particular attention to the "black-box" models, because all tools applied in this work are belonging to this category.

3.1.1 "White box" models

This is also called, as said before, "engineering approach." The engineering approach uses thermodynamic equations to predict energy consumption based on the physical behaviour of building components and their interactions with the environment. This method is known as the "white box" approach since the inner logic is clear.

This method considers the physics of the building, so there is the necessity of detailed building information to simulate the inner relations utilized to predict energy use or for analysing the demand. All models are based on the solving of equations describing the physical behaviour of the heat transfer. The fundamental law, common to all the studies, is the energy conservation law:

$$Q_{in} + Q_{source} = Q_{out} + Q_{stored} \quad [3.1]$$

Where Q_{in} is the flux entering in the system, Q_{source} the heat flux produced by a heat source, Q_{out} he heat flux leaving the system and Q_{stored} the heat flux stored inside the system. In this approach are studied in detail the phenomena of conduction through walls, the convection, the longwave and shortwave radiation, and the ventilation, for the incoming heat flux, but also losses and generation of heat.

All these equations are put together and solved with software (for example EnergyPlus or TRNSYS). There are many of them to aim this goal and there are many reviews and comparisons among them. Just to name a few, works of Wall M., McDowell Timothy et al. or Wolosyzn Monika et al. But, as discussed in the review of the last authors mentioned, some of them are not so good in considering moisture effect. It is important to choose the appropriate software, together with the appropriate equations.

This is only one of the critical aspects in the usage if withe-box models. The leitmotiv is that for developing this type of models to study the energy consumption of existing buildings can be time-consuming and requires expertise.

In general, the main disadvantage of the physical approach is that it requires a detailed understanding of the physical behaviour, which can be expensive and difficult to obtain for all the mechanisms occurring inside and outside the building geometry. This method (and grey-box has the same problems) may not be practical for these buildings where there are difficulties in obtaining information about building envelope specifications and mechanical systems.

Regarding white box approach there are models for every aspect of the building: space heating, natural ventilation, air conditioning system, passive solar, photovoltaic panel, hygrothermal effects, financial issue, occupants' behaviour, climate environment, etc.

Currently, there are three primary thermal building models in use: multizone, zonal, and Computational Fluid Dynamics (CFD) methods. It's difficult to determine which of these models is superior to the others as each has its own strengths and limitations, and the selection of a particular model depends on the specific problem being addressed. The detailed description of these approaches are beyond the scope of this thesis, but the analysis and review of Aurélie Foucquier et al. does a perfect analysis of these approaches. It is important to highlight the fact that all of them share the problems described before.

As a point to underline, all building simulation techniques require input parameters such as meteorological data, geometrical data, thermo-physical variables, and occupancy and equipment scenarios, among others. However, these parameters are subject to uncertainty, and in addition to that, there are uncertainties induced by the assumptions made to simplify the thermal mechanisms occurring in buildings. Therefore, evaluating the accuracy of building models is challenging due to these uncertainties, making it difficult to obtain a comprehensive overview of heat transfer in buildings without accumulating too much uncertainty.

The conclusion is that this approach is not always feasible, in contrast to statistical tools, which can generate models based solely on measurements, as shown in a few lines.

3.1.2 "Black box" models

In contrast to the with-box model, the AI-based approach, or the data-driven approach (also called "statistical" approach) is considered a "black-box" method since it estimates energy consumption without knowledge of the building's internal structure. They do not require any physical information such as heat transfer equations or thermal parameters.

Instead, they rely on a function based on training data that describes the behaviour of the system. These methods are useful when physical features of a building are unknown. Statistical tools, including learning methods, can build prediction models without much knowledge of building geometry or physical phenomena. However, they rely heavily on data, and collecting data can be difficult in certain cases. Another problem is that they depend also on the quality of data: having an incomplete dataset, with many holes and values not saved could be a large problem for the goodness of the model.

The statistical techniques most used for building energy forecasting, according to the review analysis of Tanveer Ahmad et al. and the work of the already mentioned Aurélie Foucquier, are linear multiple regression (CDA), artificial neural network (ANN), genetic algorithm (GA), and support vector machine (SVM). These techniques are part of the field of artificial intelligence.

As already said for white-box models, the detailed description of every single method it was mentioned in this paragraph is useless for the aim of this work. It was provided a detailed description of every method and equation it was used in the analysis of thermal energy demand of the building in the chapter 4.

It must be specified more in detail only the approach of the multiple linear regression or conditional demand analysis (CDA) because the Energy Signature Curve method used for the analysis of the demand belongs to this category.

The conditional demand analysis (CDA) is a technique used in building forecasting that applies a linear multivariate regression approach. The use of linear regression in energy consumption prediction for buildings was first introduced by Galton in 1886. In 1980, Parti and Parti proposed a new method that used linear regression, called the conditional demand analysis, for energy consumption prediction in buildings. The approach involved deducing energy demand by summing up several end-use consumptions with a noise term. By using this method, they were able to predict residential end-use consumption on a monthly and yearly basis using household invoices in San Diego.

The linear multivariate regression technique predicts a variable Y (in the examined case Y is the thermal energy demand of the building) by using a linear combination of input variables (X_1, X_2, . . ., X_p) plus an error term (E_i).

$$Y_i = \alpha_0 + \alpha_1 * x_{i1} + \alpha_2 * x_{i2} + \cdots + \alpha_P * x_{iP} + E_i, \;\; i \in [1, n] \;\; [3.2]$$

The other terms that appear in the general law are: α, the regression coefficients, p is the number of variables used in the model and n is the number of elements in the sample data.

The CDA technique is a versatile tool that can be applied for both predicting and mining data. One of its significant benefits is its user-friendliness, making it ideal for beginners who do not require any specialized knowledge or parameter tuning. It has, also, the advantages of every black-box model: it doesn't require physical information about the building.

However, the traditional method of multiple linear regression has a significant limitation in dealing with non-linear problems, which restricts its forecasting flexibility and creates difficulties in managing correlations between variables. One way to overcome these challenges is to use a feature selection approach as a preliminary step.

Regarding its application field, multivariable regression is a common technique used in the building sector to predict energy consumption and to compare energy demand between different time periods (Markel Eguizabal, Roberto Garay-Martinez et al.). It is also utilized for forecasting indoor air conditions (Lam Joseph C. et al.), controlling HVAC equipment, managing system reliability, and overseeing system management (Freire Roberto Z. et al.).

By not requiring physical information about the structure in analysis, it is possible to apply a CDA at any scale, from single building to regional level (work of Lafrance and Perron on residential electricity demand in the region of Quebec, in Canada or the research of Tiedermann about the annual end-use consumption in the region of British Columbia, in Canada) and even to national level, as shown by Aydinalp-Koksal and Ugursa that used CDA to estimate residential consumption in the entire nation of Canada.

Other examples of the application of the regression-based method for analysis and prediction of heat load con be found in the works by Catalina T. et al. for example, where they investigated the modelling of monthly heating demand for residential buildings.

To mention other studies, a Comparisons between the energy signature method and the Eta method based on statistical regression models can be found in the article published by Bauer M. et al., which found a high degree of predictability for both heating and cooling loads when treated simultaneously.

The flexibility in terms of application of the regression model offers the possibility to use in even in other sectors, aside from residential one. The banking sector, for example, was studied by Aranda et al., who applied regression models to predict annual energy use.

Overall, it is clear that the regression algorithm is widely utilized due to its ease of use and accuracy.

The primary challenge associated with this method is the requirement for a large amount of data to ensure accurate predictions. Additionally, non-collinearity between variables is essential for optimal results. These facts are highlighted is the research work of Aydinalp-Koksal Merih et al. where a comparison between AI and engineering method for modelling energy consumption is discussed.

For the analysis of the demand in the cases study analysed in this study the CDA is a perfect tool. It was entered more in detail in the explanation in chapter 4.

3.1.3 "Grey-box" models

The hybrid approach, known as the "grey-box" method, combines the strengths of both the white-box and black-box methods to overcome their limitations. In particular, as described in the previous paragraphs, white-box methods assume that all building characteristics, including thermal and geometric properties, are well-defined. While this may be true for newly designed buildings, collecting such extensive information for existing structures can be challenging. Nonetheless, this data is critical for establishing effective monitoring strategies.

Additionally, these approaches require an elevated level of accuracy in describing all physical mechanisms. Although most thermal phenomena are well-understood, some are based on assumptions and are difficult to model accurately, such as natural ventilation, which is often described using empirical equations.

Black-box methods, on the other hand, are primarily limited by their dependence on copious amounts of data. Additionally, it can be difficult to interpret statistical results in physical terms, and data mining techniques are specific to individual buildings, requiring new modelling for each new structure. In contrast, white box methods, which utilize general heat transfer equations, can be applied more generally.

One way to overcome the limitations of each technique is by combining them through a hybrid approach, known as "grey box ", methods.

By retaining some physical meaning, interpretability of the problem is maintained, and building characteristics can be determined through optimization techniques like AI genetic algorithms. This removes the need for all physical and geometric input parameters, resulting in a more flexible and adaptable approach.

Overall, the advantages of one method can offset the drawbacks of the other, resulting in a more comprehensive and effective modelling strategy.

According to the work of Aurélie Foucquier et al., there are several strategies to implement this type of method. As it happened in the case of white-box and black-box, it will not be described in detail all these strategies, because it is not useful in this thesis. It will be synthetized only the general characteristic of every strategy.

The first approach involves using machine learning to estimate physical parameters, often by coupling a nodal model with genetic algorithms (the case of the model for the consumption of a building in Tunisia made by Znouda et al.). Another approach involves using statistics to create a learning model that describes the building's behaviour. This learning model is based on a physical approach and can be used to make predictions.

A third approach involves using statistical methods in areas where physical models are not effective or accurate enough. For example, end-uses are difficult to account for in physical models but can be better considered using statistical techniques. In this case, physical and statistical methods can be combined to create a complete system (as an example, the work of Siddarth et al., which used a genetic algorithm to create a database for the implementation of a regression model).

The hybrid method offers several advantages, such as requiring a limited amount of data and not needing fixed input parameters at the beginning of the simulation, only bounds on physical parameters are required. Additionally, it still provides a physical interpretation.

However, there are still some drawbacks to consider, such as the need for free parameters for statistical tools and the computation time needed for both physical and statistical codes. Moreover, since it combines two different scientific domains, the grey box method may bring some difficulties for users to understand, but it should be of great scientific interest.

3.2 Tools for buildings energy consumption analysis

After the general introduction to the diverse types of approach in the modelling of energy consumption in buildings, now it enters more in details explaining the literature background of tools used in the analysis. Even in this case, the detailed description of the tools is not provided, because it is presented in the chapter 4. The intention of this paragraph is to propose a literature review and provide some references to the method used in the analysis of thermal energy demand od the building.

3.2.1 Energy Signature Curve

This is a powerful instrument to evaluate the energy demand of a building. It belongs to the statistical tools, so there is no necessity of physical parameters of the building or any specific information about end of use, users' behaviour, schedules, or any other information.

Energy Signature Curve (ESC) is a linear regression method, developed for estimating ad forecasting the demand of a building, with particular attention to the space heating demand in winter period. The first time this method was proposed was in 1986 by J. D. Balcomb and

it was firstly developed as a passive solar design tool. After that it was adapted for the analysis of energy demand in buildings.

The idea at the base of this method is that it is a well-established fact that outdoor temperature is the most significant factor affecting SH heat use (this is shown in many articles, just to mention a few works: Nicol JF. et al. about the thermal comfort of people inside a building or Schneider S. et al. talking about a demand load curve model for the Swiss national territory).

There is, also, a seasonal variation in DHW use that is also related to outdoor temperature (it is evident in the study of Gerin O. et al. that there is a seasonal variation in DHW energy consumption, in particular in winter consumption is 12% higher than the average and in summer it is 13% lower).

However, an hourly analysis has shown that there is no significant correlation between DHW use and outdoor temperature. This fact is shown in the article of Ivanko D. et al., where they built a model for the prediction of DHW consumption in hotels in Norway and shown that outdoor air temperature has almost no influence on DHW energy demand, the main variable in this case is the number of guests in the hotel.

Therefore, the regression model between total heat use and outdoor temperature is mainly influenced by SH, and the residual of this model accounts for DHW heat use.

The Energy Signature Curve (ESC) is a useful tool for analysing heat use in buildings and its relationship with outdoor temperature (the work presented in the doctoral thesis of Linda Petersen is noticeably clear in this field and this is one of the bases for the construction of the ESC in the method used in this study).

The ESC typically consists of two parts for buildings with heating seasons and no cooling, separated by a critical outdoor temperature known as the Change Point Temperature (CPT).

Figure 1 An example of ESC

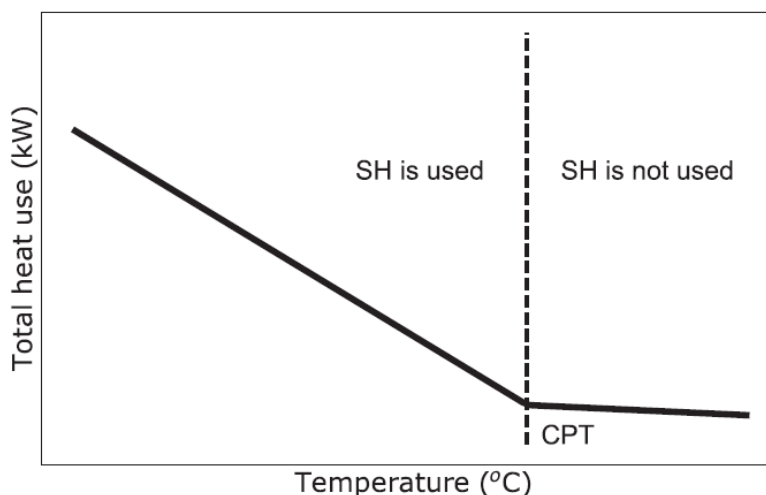In the figure 1 there is an example of Energy Signature Curve. The CPT (Critical Point Temperature) is a pivotal outdoor temperature that establishes the threshold between the onset and conclusion of the heating period.

Once the CPT is reached, the use of space heating (SH) within the building is restricted. The portion of the graph preceding the CPT represents the SH season, during which the SH

system is responsible for a significantly greater amount of heat use than the domestic hot water (DHW) system. This is also called the Temperature dependent part of the ESC because energy consumption depends on outdoor temperature.

The function after the CPT illustrates the warm season, during which SH is not necessary, and the primary source of heat use is the DHW system. This il also called the temperature independent part of the plot due to the fact that, being DHW the highest energy consumption of the period, there is almost no dependence on outdoor temperature. However, depending on the type of system in use, a small amount of heat may still be required to operate the SH system during this period.

Usually, when examining an energy signature diagram, simple regression models are employed to comprehend the energy usage patterns within a building. Since the energy signature diagram is influenced by the beginning and ending of the heating season, as specified before, it is often divided into temperature dependent and temperature independent sections, which are frequently described by the equations presented in various publications, among others the already mentioned ones. The equations are similar trough all studies of ESC, and their general form is:

$$Y = \alpha_0 + \alpha_1 * x_1 + \alpha_2 * x_2 + \cdots + \alpha_P * x_P + E \quad if \ T_t < CPT \quad [3.3]$$

$$Y = \alpha_0 + E \quad if \geq CPT \quad [3.4]$$

Where E is a random error, $\alpha_0, \alpha_1, \alpha_2, \ldots$ are coefficient of regression and they described the expected change in the variable Y when there is a unitary change in input variables x_i and the rest of predictors remain constant. From the expression of the equation, it is clear the fact that ESC is a statistical model, in particular a CDA model.

It is important to underline the fact that CPT depends strongly on the building, especially on the users' behaviour, the climate, and the location. In some particular cold climate, it is possible that SH is necessary even in summer, so the biggest part of the ESC is a temperature dependent part.

The individuation of the CPT is a crucial part for a good ESC calculation. For certain buildings, either the last day of the heating season or the CPT is already known. In cases where the CPT is known, an Energy Signature Curve can be constructed using the least square method for two parts of the model, as depicted in Figure 1.

Alternatively, if the CPT is not known, the piecewise regression method is one of the methods that can be employed to identify it. This is not, by the way, the unique possibility to find CPT.

Furthermore, several publications have indicated that the selection of the CPT is often based on the intuition and experience of researchers, but this will introduce inaccuracies and this method becomes too much related to the experience and the intuition of the user.

Another method was proposed by Linda Petersen in her doctoral thesis. Starting from the work of Aronsson, she proposes to find CPT checking where there is a significant the slope of interpolation line of experimental points, starting from temperature independent zone. It will be described better this approach in the chapter 4, because it was implemented this method in this analysis.

It is important to highlight the fact that in some publications the separation between temperature dependent zone and temperature independent is found not in temperature but in the demand. This is the case of the work by Lumbreras M, Garay-Martinez R. et al. about data-driven model for heat load prediction. In their study it was necessary the creation of a model to estimate the heat demand of the building in analysis. To calibrate the model, a reference heat load called Q_{REF} is used to split the data into two parts. It is shown an example of separation of heat consumption with a demand change point in the figure 2.

The data below Q_{REF} is considered to be independent of weather conditions (th4e correspondent Temperature independent part), while the data above this point is assumed to have a linear correlation with the climatic variables (it corresponds to the temperature
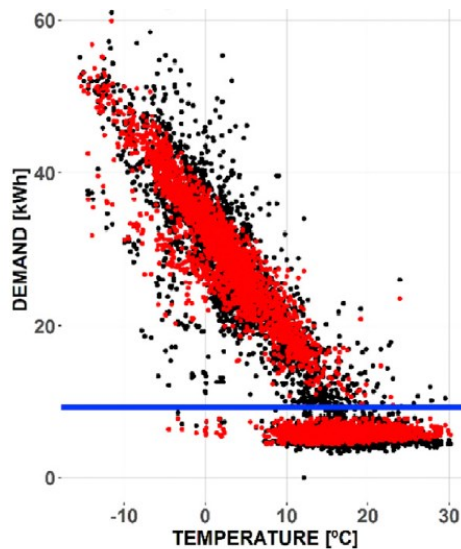
Figure 2 Experimental data (black) vs model estimation (red) in case there is a change point in heat load (blue line)

dependent part). To determine Q_{REF}, an iterative process is followed using different heat load thresholds, ranging from 0 to 0.5 *Q_{MAX}, where QMAX is the maximum heat load.

It is shown that this substation of the change point temperature with a change point value of heat load produces impressive results when multiple variables are taken as input parameters. In the studied mentioned the regression model to estimate building consumption was built using outdoor temperature, but also other weather parameters such as the global solar irradiance on the horizontal plane, the wind speed, and the wind direction. In models that use only outdoor temperature ad an input variable the CPT is ok for subdividing ESC in the two zones. In this work it was used CPT and not the change point in demand.

3.2.2 DHW + SH splitting.

This is one of the crucial points of the work presented in this thesis. Few studies propose a complete procedure to split DHW and SH consumption. The method proposed usually are made for specific building categories or are not studied with a large database of building.

This part of the analysis is particularly important because according to the European Directive 2018/844, it is necessary to analyse the energy performance of buildings based on calculated or measured energy use. This analysis should take into account the typical energy use for space heating (SH), domestic hot water (DHW), and other technical systems in a building. These guidelines are specified in the document mentioned before. This approach is crucial for developing energy-saving solutions in all technical components of the building.

The importance of splitting DHW and SH consumption from the global thermal energy demand is evident since smart heat meters are not typically available in buildings (fact explained in the work of Volkova A. et al.), and many buildings use only one heat meter for the total heat use, which cannot measure SH and DHW heat use separately. It is important to highlight the fact that according to the European Commission rules about smart grids promotion strategy, by 2024 it is expected that almost 77% of European consumers will have a smart meter for electricity and about 44% will have one for gas, but now we are far from these results.

Since SH and DHW systems have different operating regimes and influencing factors, it is necessary to analyse their heat use independently. It is crucial to have distinct statistical data for the use of domestic hot water (DHW) and space heating (SH) to address several issues, such as accurately sizing SH and DHW systems, designing efficient energy management and control systems, improving existing standards, predicting models, and energy use profiles, have information about building's behaviour and have better data for studying an eventual retrofit intervention. Therefore, dividing the total heat demand into SH and DHW components is a vital task.

Regarding the method for doing this operation, a number of research teams are investigating the issue of how to extract data on SH (space heating) and DWH (domestic hot water) usage from measurements of total heat consumption. However, this is a complex problem, and different researchers have different criteria for what constitutes a satisfactory outcome, meaning that there is no single standard method for conducting such data analysis. In the next lines it will be presented the most relevant method founded in the literature analysis.

One method that has been suggested for separating the total heat demand in a building into SH and DWH usage is presented in the work by Bacher P. et al. This research uses data collected at 10-minute intervals from a single-family house in Denmark. The method assumes that DWH usage generates short-lived spikes in the time series, while SH usage changes more slowly over the course of the day in response to factors such as climate and user behaviour.

To estimate SH usage, the authors propose using a non-parametric kernel smoother. Any values that are significantly higher than the kernel smoother are then considered to be spikes in DWH usage. This is an interesting idea and the intuition to use sudden spikes as an indicator of the presence of DHW consumption is used in other works, for example in the second part of the analysis of Dmytro I. about the splitting of SH and DHW consumption in a hotel in Norway. However, this method has not been validated using data collected separately for SH and DWH usage. As a result, it is difficult to assess its accuracy and reliability.

The work of Burzynski et al. explores the issue of separating weekly heat use into DHW and SH. The authors of this paper assume that the period during which the outdoor temperature exceeds the base temperature corresponds only to DHW heat use. This is an assumption used in many articles that talk about ESC analysis, as discussed in the previous paragraph.

Using this assumption, they were able to identify DHW heat use for several warm weeks throughout the year. They then proposed using monthly variation factors for DHW heat use to extrapolate from warm months to other months of the year. These factors are available for UK dwellings in "The government's standard assessment procedure for energy rating of dwellings", but the problem is that this document is absent in some nation.

There is, also, the issue that they are standard value, fact that produces inaccuracies in the analysis of a single specific building. In addition to this, these factors are not available in the literature for other types of buildings aside from dwellings. It is worth noting that in some buildings, SH heat use can occur even during warm months.

For all these reasons, they are methods usable in few situations and not in every type of building.

In the work of Marszal A.J. et al., it is presented a method for estimating the hourly space heating and daily DHW heat use profile. This study uses hourly values of the total heat demand in the building and involves several steps.

Firstly, the daily total heat use profile for an average summer day is calculated because there is the hypothesis that in summer SH consumption is extremely low or totally absent.
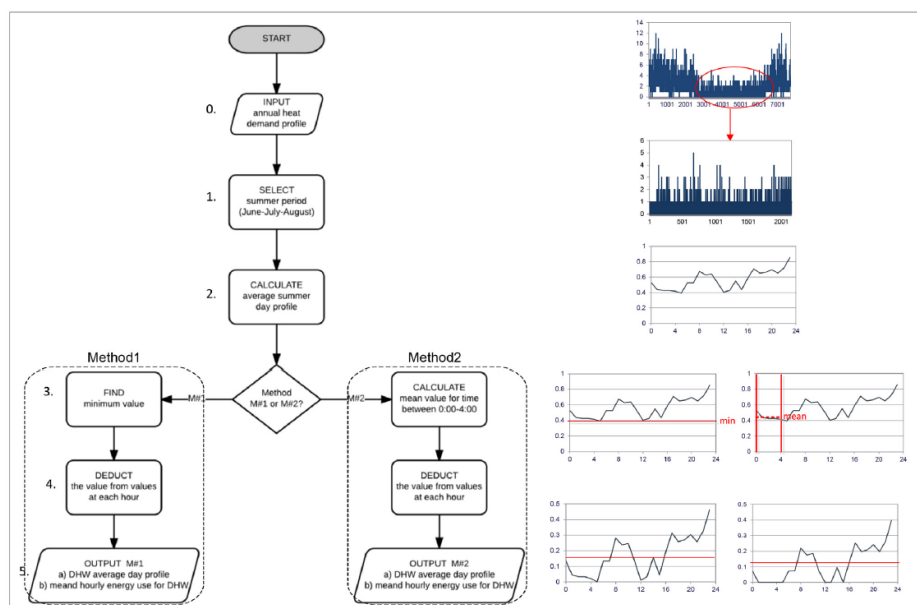
Figure 3 Flow chart of the methodology by Marszal A.P. et al.

Secondly, non-DHW use is determined as the minimum of the total heat use profile for an average summer day or the average for hours from 0:00 to 4:00. This step, together with the first one, is very good for the estimation of losses in the DHW circuit and it was used this procedure in the analysis described in chapter 4.

Finally, the DHW profiles are calculated by subtracting the non-DHW heat use from the value of the heat use at each hour of the day. This is the most problematic point: in other seasons

we aren't sure that the load founded in the step 2 is the only non-DHW load, because there is a strong relevance of SH heat load.

Figure 2 presents the flow chart of the methodology. It will be a reconnection to this article when there will be a discussion about DHW circuit losses estimation.

The results of this study demonstrate that the method produces satisfactory outcomes when DHW use during the summer is at least at the same level as space heating. However, this method does not account for DHW heat use during other periods, aside from the warm season.

There are some alternative approaches for identifying SH and DHW heat use, which involve the use of building simulation tools. For instance, one such methodology is discussed in the work of Neu. O. et al. which employs an occupant-focused approach and a time-of-use survey (TUS). The authors apply Markov Chain Monte Carlo techniques to TUS activity data to develop activity-specific profiles for occupancy and domestic equipment use. They assume that heat demand is influenced by household size, the type of day, and the season. The DHW heat use profiles combine probability distributions for specific TUS activities with the average daily DHW heat demand. This is an interesting method, but it requires additional information about the use of the building, the number of occupants, the type of occupants… to create an accurate model of the specific building in analysis. In most of the cases, especially for privacy problem, this type of information is not available, and the application of the method becomes very hard.

Since this moment it was presented only statistical approaches, but it is possible to split DHW and SH even with hybrid methods. This is the case, for example, of the work done by Verhaert I. et al. which proposed a method to estimate DHW consumption based on tap patterns. They used building performance simulation archetype models to estimate the SH and DHW heat use in a building.

They measured volumetric flow rates and water temperatures in 20 different buildings to characterize DHW use. Stochastic simulations were then executed on the measured data to obtain representative DHW use profiles, which can be used as input for simulation tools like IDA ICE, EnergyPlus, and TRNSYS.

However, developing a complex model for all the components in a building is usually required for such simulation tools, and these models are often only suitable for a particular building. Furthermore, these models have been shown to be less accurate than data-based methods as shown in the comparison analysis done by Tronchin L. et al., where it is said that these calculation codes do not give a comparable results of energy consumption in buildings.

Other authors suggest using statistical models and heat use profiles of typical buildings obtained from building stock databases to estimate the SH and DHW heat use. For example, a Neural Networks model is proposed in the research of Aydinalp. M. et al. to estimate the SH and DHW heat use in typical Canadian households, which uses data from the 1993 Survey of Household Energy Use (SHEU) database.

Such models can be used to separate the SH and DHW heat use in typical buildings, but their accuracy for individual buildings may be uncertain. This method doesn't consider the specific operation of the particular building in analysis, and this may conduce to not

consistent results. The availability of appropriate databases is also required to develop such models and this type of data is rarely available.

Talking about linear regression models, it was already said that they have been used in several studies to predict heat demand in buildings. A good example of that can be found in the already mentioned work of Linda Petersen or in the analysis of energy management in neighbourhoods done by Sorensen A.L. et al. They use linear regression models to separate DHW heat delivery from the total heat delivery.

In Sorensen's article, a linear regression model for total heat delivery was developed by taking various factors such as outdoor temperature, hour of the day, weekdays, and holidays into account. When estimating DHW heat delivery, the outdoor temperature was set to the approximate break-point temperature of the model, resulting in a DHW daily load profile with hourly mean values.

In their work, Riachi Y. et al., investigate the hourly energy loads of space heating (SH) and domestic hot water (DHW) in buildings. They use various factors to estimate the hourly DHW heat use, including water volume usage, building activity, and type of DHW system. The study also models hourly SH loads by considering outdoor temperatures, building setpoint temperatures, night setbacks, and weekends. So, it is still present the problem of knowing detailed information especially about users' behaviour but, as already discussed, they are hard to find.

The review of literature reveals that the issue of separating the total heat consumption of a building into SH and DHW components remains unresolved, especially for larger buildings with limited knowledge of user behaviour. Most of the existing methods are simplified and focus only on obtaining average daily profiles for a given year. Some methods can provide general models of SH and DHW heat use for a particular category of buildings, but not for an individual building. Other methods only solve the problem for several warm months, assuming that the SH is not operating during the summer. Several approaches require extensive knowledge of user behaviour, building physical properties, and system parameters, which limits their application. Additionally, most of the research work analysed heat use in residential buildings, whereas non-residential buildings, including hotels, have received less attention.

For solving these problems, Dmytro I. et al. wrote an article that introduces a technique for separating the total heat consumption measured hourly into two parts: space heating (SH) and domestic hot water (DHW) heating. The first step of this approach involves creating a SH heating model using total heat use data, which relies on an energy signature curve (ESC) and singular spectrum analysis (SSA). The idea at the basis of the SSA analysis is the same as the intuition had by Bacher P. et al.: DHW demand produces a peak in the global heat load demand and by performing a SSA analysis it is possible to find these peaks. The second step involves extracting a DHW heating model from the difference between the SH heating model and the total heat use data.

To test this method, the researchers applied it to one year of hourly data collected from a hotel in Eastern Norway, where they also measured the SH and DHW heat use separately. By comparing the results of the total heat use splitting with the actual measurements, they found that this method is helpful for obtaining valuable information about the SH and DHW heat use in a building with only one heat meter.

The starting point of this work is what proposed in this article, because it offers the solution to the main issues of the procedure for splitting SH and DHW and because it was searching for a method that can work with zero knowledge about building or occupants' habits, exactly the case of the dataset of Tartu and Verona, analysed in the case studies section, in chapter 6.

3.2.3 Detailed DHW consumption analysis

After the operation of splitting the global heat demand in the space heating and domestic hot water heat load consumption, it emerged the necessity to analyse the DHW profile more in depth. To anticipate some arguments discussed in chapter 5, one of the main problems to solve after the splitting of the two parts of heat consumption is the accuracy of the DHW model. It will be discussed better this point in the following chapters, for now it is important to know the fact that it became necessary to have a tool to perform a detailed DHW consumption analysis.

On the other hand, even if the starting point is a database with the subdivision between SH and DHW done by smart meter, an analysis of DHW consumption profile is extremely significant. In many European countries a lot of focus has been placed on investigating the performance of space heating systems, while domestic hot water (DHW) heat use has been considered a minor aspect of energy needs for heating, especially in countries with cold climates.

However, with the introduction of passive house technologies and improvements in building envelopes, space heating needs in buildings are decreasing, while DHW heat use remains relatively unchanged. For example, a study in Denmark conducted by T. Peters et al. focused on achieving low heat use in passive buildings through highly insulated, resource-efficient, and airtight designs, but did not consider DHW use. However, the study found that energy demand for DHW use was almost twice as high as for space heating, underlining the importance to not neglect this type of energy load.

In another study done by J. Hirvonen et al. about buildings in Finland, DHW heat use was assessed by using profiles obtained from measured DHW demand as input in simulation software. The results showed that in modern buildings, DHW heat use was the most significant component of heat use (it is shown that in two well-insulated buildings taken into consideration the DHW produces the 52% and the 63% of the total energy consumption), while in older buildings, it contributed 24% to 30% of total energy use. Unambiguous evidence that the DHW consumption plays a significant role in every building, and its importance becomes higher and higher with the improvement of building insulation and a consequently energy demand for space heating reduction. There is the evidence of this fact in many articles, it can be mentioned the work of S. Attia et al. which provides an interesting overview of the phenomenon of the nZEB buildings in Europe, raising interesting common goals and design suggestions.

A tool that can provide a detailed analysis of DHW consumption is fundamental to promote the shifting of the focus of energy-saving measures in buildings from improving space heating to improving DHW systems. To fully realize the potential for energy savings in DHW

systems, research, and innovation in the field of DHW energy performance are becoming more relevant and valuable.

It is important to acknowledge that DHW systems are associated with issues related to sanitation and health safety, in different modality than space heating. When we talk about DHW systems, o0ne of the most serious question to deal with is he appearance of the Legionella bacterium, which can cause different forms of pneumonia and even death. The conditions that promote Legionella growth include water temperatures ranging from 25°C to 42°C, stagnant water, and nutrients. Therefore, many countries, including Italy, have developed regulations aimed at minimizing the risk of Legionella disease. All the rules for preventing the birth and development of the bacteria are collected in the document "Linee guida per la prevenzione ed il controllo della legionellosi" provided by the Italian government. Just a few data: in the storage tank water temperature must be always higher than 60°C and higher than 50°C at the basis of each recirculation column. Thus, ensuring the safety of energy-efficient solutions is a critical factor in DHW systems.

To simplify calculations, some methods suggest assuming a constant DHW use. This is the case, for example, of many standard design methods for building. In the literature can be found specific values of DHW consumption for every type of building (residential single house, condominium, hotel, hospital, …) However, practical experience shows that these standards do not reflect the real DHW use in buildings. This concept is the conclusion of the research work by T.A. Koiv et al. In fact, they underline the necessity of new calculation formulas for determining the DHW design flow rates for instantaneous heat exchangers and the fact that there is a significant difference between design flow rates obtained from standard method EVS 835 and their empirical law: with their own formula flows are two times smaller.

This is just an example of the problems caused by oversimplification of DHW consumption: oversizing of DHW system components, leading to extra financial and energy losses.

To solve this problem the solution is a detailed analysis of DHW consumption, and the DHW heat use profiles are a crucial tool for predicting the amount of DHW heat use in buildings. By analysing these profiles, it is possible to observe the variations in energy use over different time periods. DHW heat use profiles provide valuable information such as peak energy loads and other energy load characteristics of the building.

The analysis of performance of DHW systems is a complex issue that involves various areas such as economic, sanitary, behavioural, and technical aspects. DHW heat use profiles are helpful in identifying energy-efficient solutions in all these areas. For instance, economic analysis can be performed to evaluate DHW pricing and its correlation with income and water prices. An example of this type of analysis is the work done by Reynaud A. et al., where they talk about the impact of individual metering for DHW consumption, finding that this method of measuring demand, instead of collective metering, has a high impact on the price elasticity and produces different values.

Technical solutions can be designed to address sanitary problems, but they require knowledge of when and how much DHW water is used. Additionally, distinct types of DHW heating systems can be investigated to reduce DHW energy use by combining traditional and renewable energy solutions. In the article written by H. Ahmad there is a good example of this type of discussion. The scope of his research is to present different technologies for

the domestic hot water generation and to present the difference in performances among them in terms of efficiency, CO2 emissions and greenhouse gases emissions.

Building simulation software tools such as EnergyPlus, TRNSYS, TRANSOL, etc. also require DHW profiles to simulate DHW systems' performance accurately. Accurate DHW profiles are essential for effective demand-side management, energy conservation measures, and legislation improvement. For all these reasons is important to perform a detailed DHW consumption analysis even if no splitting operation was done before.

However, due to differences in building characteristics, data quality, and calculation requirements, there is no universal method for performing appropriate analysis. Numerous scientific works have been dedicated to developing and analysing DHW energy profiles.

For example, one study measured volumetric flow rates, cold and supply temperatures to characterize DHW use in variously sized buildings and used stochastic simulations to obtain representative DHW use profiles for end-users. This is the case of the already mentioned work by Verhaert I. et al. Some methods for DHW profile development are based on operating schedules for primary DHW energy users and occupant activities (time of shower, cooking, cloth washer, dish washer). There are many standards that offer schedules and typical profiles useful for analysing DHW use in new and existing apartment buildings. Just to mention an example, the document "Building America House Simulation Protocols" a document from 2014 written by national renewable energy laboratory (NREL) and provides, aside from profiles for the DHW study, also instruments for the analysis of building envelope, space conditioning equipment, distribution system, appliances, lightning for both new constructions and existing homes.

There are other studies that uses the occupant activities to perform a detailed analysis and a precise simulation of the DHW load consumption. This is the case of the studies by N. Good et al. and by D. Fischer et al. In the first mentioned article they can build high resolution profiles for DHW and other loads by taking into consideration the number of activities and appliances that use this type of load. To create a such precise profile, it is necessary to know the exact number of appliances in the building, the number of occupants and their habits in terms of time of use. In the Fisher et al. work the reference loads for DHW are taken from standard VDI 4655, with the cons, highlighted by the autos themselves, that there is no possibility of capturing the random nature and the effect of removing irregularities that may arise when multiple households are simulated simultaneously.

There is no shortage of examples of use AI-based method. They are used to generate in an automatic way DHW profiles, as it happens in the simulations done by L. G. Swan et al., where a neural network technique was applied.

The majority of the research reviewed focuses on apartments and households, where it is easier to obtain the necessary parameters. However, in non-residential buildings, acquiring detailed information on occupant activities and equipment operation can be a time-consuming and costly task. This limited availability of input data makes it challenging to apply these methods in practice. This difficulty in finding data makes difficult even the operation of validation of these models for this type of building, another problem that affects the applicability in frequent practice, as emerged in the analysis done by Ivanko D. et al.

Researchers have also acknowledged the challenges of comparing actual DHW energy use profiles with standards and verifying them. For instance, in one study, the actual DHW profiles in apartments were compared to the profile proposed by the American Society of Heating, Refrigeration, and Air Conditioning Engineers (ASHRAE). The study conducted by J. Wiehagen et al. found that the primary difference between the actual and ASHRAE-derived data is that water use is less evenly distributed in the actual data, with higher peaks and lower troughs, and much less use in the early morning hours. This is a typical problem when there is the necessity to deal with DHW data: standard profiles differ too much from the actual consumption and the differences between one building and another one is significant. There are other examples of studies that proves the thesis mentioned in the lines above. The authors Vine E. at al. conclude their work of comparison between standard profiles for DHW consumption and experimental demand suggesting using real profiles obtained by experimental points in the analysis of buildings.

Starting from the problems emerged in the previous discussions, Ivanko D. et al. proposed a new method for the analysis of DHW consumption profiles of nursing homes in Norway, with the objective to improve upon the current methods used for analysing DHW heat usage. They want, also, to increase the knowledge of non-residential structures such as nursing homes, hospitals, hostels, and schools that, as established in the previous lines, have not received as much attention from researchers.

In this analysis it will be followed the method proposed in this article, expanding the applicability field, and testing it into new databases and with different profiles. With this method the authors were able to evaluate the comparability of DHW profiles between seasons and days of the week and determine the timing of peak DHW heat usage. To test these methods, they utilized one-year hourly measurements collected from three nursing homes located in Eastern Norway. In this work, as it will be explained later, it was considered a lor of several types of building.

Through this process, they were able to identify unified DHW profiles with similar usage characteristics for each month and day of the week. Using these profiles, they estimated the timing of peak, average, and low DHW heat usage, so do I. after this type of analysis they proposed, also, a comparison between DHW profile obtained from data deep processing and standard profiles from the national guidelines of Norway: SN/TS 3031:2016 and the international standard NS-EN 12831-3:2017. As emerged in previous study, standard profiles do not give an accurate representation of DHW profile in real buildings' operation. In particular, it emerged the fact that the European standard, NS-EN 12831-3, overestimated the daily DHW heat use by 1.65 times, and the Norwegian standard, SN/TS 3031, overestimated it by 3.5 times. The magnitude and timing of the peak heat use in the buildings was also different from the standards.

This is a clear evidence that a detailed DHW consumption profiles analysis based on real experimental data is necessary to understand the real behaviour of the building. It is impossible to blindly trust standards because they could give some important inaccuracies. It retained, for these reasons, an especially important point to focus the attention in this study on the DHW profile analysis and to use the information obtained not only to improve the model of DHW and SH splitting, but also to use to understand in a clearer way the behaviour of users and of the building.

3.2.4 DHW boxplot analysis

The next tool proposed in this thesis is an instrument to check the DHW profile consumption during summer. Being a tool focused on the analysis of DHW profile, the literature review for this point is strongly connected to the previous argument, when it was talked about the detailed DHW consumption analysis in the point 3.2.3 of this chapter.

It is behind the scope of this thesis to discuss the born and the evolution of the boxplot itself, it is not useful and a uselessly high time-consuming operation. In this section, so, it will be shown an example of the instrument it was used, and it was mentioned just some articles where there is an application of this boxplot. It is important to underline the fact that, due to his extremely easy implementation and the extreme general field of possible uses, the boxplot is commonly used in a multitude of articles and studies, from the most different arguments of research.

Boxplot belongs to the statistical tools. For this reason, it is commonly used to understand experimental profiles and to obtain some interesting information from its analysis. An interesting use of the analysis of boxplot is presented in the work of Lomet A. et al., where they did a statistical modelling for real domestic hot water consumption forecasting. To make these forecasts they studied the consumption profiles of various individual residences using boxplots of volume consumption for each month (an example of average consumption profile and boxplot con be found in the figure 4 and 5).
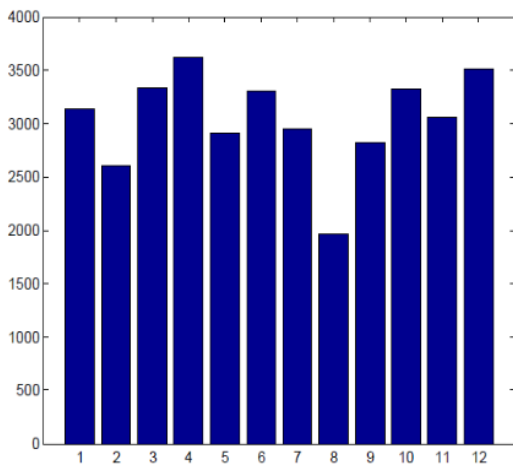
Figure 4 An example of average DHW consumption profile used in article by Lomet A et al.
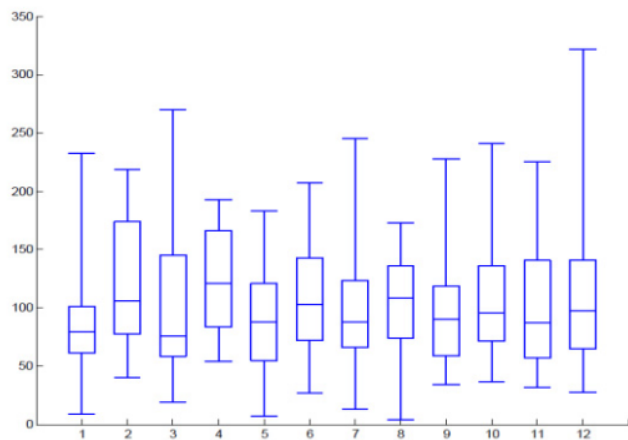
Figure 5 an example of boxplot used for DHW consumption studies in the article of Lomet A.

This model was developed to solve the main problem of other methods, such as neural networks, grey box modelling approaches or regression models: strong computational time and detailed information about the residents like their number, age or social profile that is often not available. It must be underlined an important question emerged in this article: a boxplot analysis can give essential information about weekly periodicity, random fluctuations and the different profiles of consumption following the residence, the season, and the day of the week.

Another use of this analysis, boxplot, and average profiles of DHW consumption, is presented in the already mentioned article by Anna Marszal-Pomianowska et al. Here there is the suggestion to study DHW profile consumption using an average profile created in summer months.

Just to mention the potential of the boxplot analysis: it can be used both for studying water demand in a single building, but also for an entire city, providing in both cases useful and consistent information. In the study conducted by J. C. Evarts, they used a boxplot representation for observing the water consumption of the Halifax Regional Municipality.

Another important application of this type of method can be found in the work by R. Hendron and J. Burch, where they use, among others, boxplot for putting together DHW profile consumption from diverse types of building and time of the year for creating a Standardized Domestic Hot Water Event Schedules for Residential Buildings.

For the interest of the analysis presented in this thesis it will not proceed over with the presentation of possible using of boxplot analysis, because it would be impossible to cover all the infinite possibility this instrument offers. For this analysis, it was used boxplot and average profiles analysis of DHW consumption just to identify summer season, where consumption is lower and more similar among different moths, and to notice if there is the presence of some outliers in the experimental database. As it happened for the previous paragraphs, it will be explained in detail and it will be provided examples and results of this type of analysis in the next chapters, in particular in chapter 4 and 7.

3.2.5 Boxplot of energy consumption

Even in this case, the same considerations did for the boxplot analysis of DHW consumption profiles are true. It will not go further with the repetition of concept already expressed. It will be provided, by the way, an example of the use of boxplot analysis for the global energy consumption.

The object of interest is the work conducted by Tereshchenko T. et al. about the analysis of energy signatures and planning of heating and domestic hot water energy use in buildings in Norway. Their goal was to present a more advanced approach to plan and analyse energy consumption for domestic hot water and heating in LEBs, utilizing the energy signature method. The relevant point is that they used the boxplot analysis for dealing with outliers.
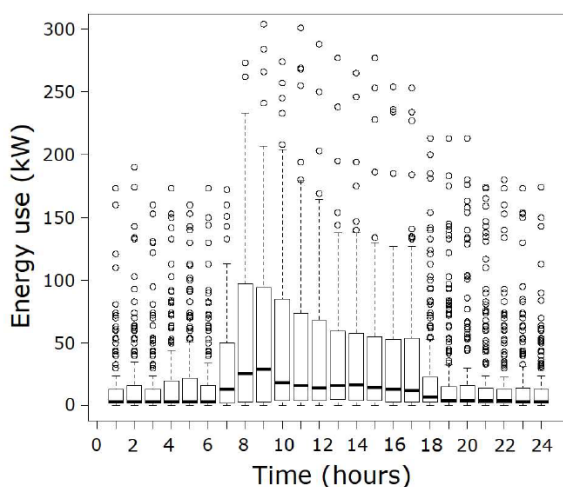
Figure 6 Boxplot Energy demand vs Time of the day taken from Tereshchenko et al. study

In fact, it is crucial to determine the factors behind the common and atypical energy usage patterns in every data analysis. So, they did an investigation into the tailed data presented in the experimental points and to gain a better understanding of the energy usage patterns, a boxplot was utilized. It is reported below, in the figure 6, one of the boxplots taken from

the article, to highlight its extreme power into the identification of mean values and, even more important, the outliers (represented as dots in the upper part of the plot in figure 6). In addition to that it can show the maximum, median and minimum energy use, and density of data points.

It will be provided an analysis of global energy consumption with this tool even in the method it was used for the analysis of case studies. As it happened for the previous tools, a detailed description of method and results is given in chapter 4 and 7.

3.2.6 Analysis of temperature lag

This tool was introduced to consider the inertia of the system. It can be noticed that the previous tools are belonging to the area of statistical approaches. They suffer of a problem of low physical interpretation of the system. It could be hard to use these statistical data-based model in some physical methods because it would be hard to observe some parameters of interest.

It can be found an interesting work in these terms in the article by Karen Byskov Lindberg et al., where the authors discuss the hourly load modelling of non-residential building stock. They improve the single regression model for building consumption estimation developed by Petersen (already mentioned in the previous paragraphs) by adding other variables as input parameters, together with outdoor temperature. They also use wind speed and solar radiation and expand the sample of investigation. Using the conclusions of the work done by Stokke G. et al. about demand in residential sector, they consider a non-linear effect of temperature in the part of the ESC (see paragraph 3.2.1) that exceeds the changing point temperature (no need for more heat) or reaches very low temperatures (the technical heat system is not able to deliver more heat), creating an S-shaped curve. To consider this phenomenon they added a quadratic term of temperature to the regression formula.

But the most important thing highlighted in this work is the implementation of a moving average of outdoor temperature. In fact, due to the presence of thermal mass in the building, which causes a delay in temperature response, a 24-hour moving average of the outdoor temperature was incorporated into the model. The model with TMA shows an increment by 0.6% to the index R^2, passing from 75.4% of the single regression model, with only outdoor temperature as an input variable, to a value of 76% with both outdoor temperature and TMA.

Seen in this term the improvement may be not so interesting. But if we consider the work of the already mentioned Tereshchenko et al., we found a peculiar implementation of TMA. Here The purpose of the TMA analysis was to investigate the impact of outdoor temperature lag, in terms of the hour of the day, on building energy consumption. Building structures have varying levels of thermal inertia, with some being capable of storing heat for later use in improving indoor thermal conditions. To determine the extent of the effect of temperature lag on energy consumption, the outdoor temperature was shifted every hour for a period of 48 hours, and correlations between the variables were examined.

The analysis revealed that the common practice of using 24 or 48-hour lag periods to define and describe building inertia is misleading. The results indicated that the appropriate lag period is dependent on the building's construction type and time constant parameter.

Starting from these conclusions, it was implemented a similar analysis in these case studies, trying to understand how every single building react to the thermal inertia. As usual, it was presented the procedure followed in the chapter 4 of this thesis.

3.2.7 Parameters for testing accuracy of models

Regarding parameters and indicators for evaluating the accuracy of the tools proposed the choice was easy. In all articles mentioned before that used statistical methods in their analysis there is the presence of the factor R^2 to evaluate the goodness of the model. The most accurate and complete expression and explanation of this parameter can be found in the article by Lumbreras M, Garay-Martinez R. et al. regarding data-driven model for heat load prediction. In this article there is an entire paragraph dedicated to the explanation of the tools for the analysis of results.

To assess the precision and effectiveness of the model, it was evaluated numerically using the coefficient of determination, R^2. This value quantifies the percentage of the variance in the data that can be predicted by the model's predictors. The R^2 variable is computed using the following formula:

$$R^2 = 1 - \frac{SSE}{SSYY} \quad [3.5]$$

$$SSE = \sum_{i=1}^{N}(X_i - Y_i)^2 \quad [3.6]$$

$$SSYY = \sum_{i=1}^{N}\left(X_i - mean(X)\right)^2 \quad [3.7]$$

where Y_i is the value obtained from the prediction model, X_i represents the measured experimental data and N is the number of observations.

In the same article, they propose another indicator. This is necessary because their model's practicality lies in its ability to forecast heating demand and optimize the heat generation process. The DH operator is responsible for managing the heat production process across the entire DH network, so it is essential to evaluate the model's accuracy in energy terms. Solely relying on the R^2 value can result in an overfitted or biased model. For the current application, fluctuations caused by the high thermal inertia of the DH network need to be considered, and the analysis focuses on overall energy outcomes, as said by authors.

So, they introduce another parameter: the yearly energy consumption deviation (YEC). The expression is reported in the equation [3.8]. A 0% indicates a perfect match between measurement and prediction.

$$YEC = 100 * \frac{\left|\sum_{i=1}^{N}X_i - \sum_{i=1}^{N}Y_i\right|}{\sum_{i=1}^{N}X_i} \quad [3.8]$$

This parameter is a sort of MAPE (mean absolute percentage error). It is used to estimate the error present when there is the estimation of the entire year consumption. It takes into account the general deviation of modelled points from the experimental data. This is an

extremely useful instrument when we talk about model for forecasting or for modelling building consumption.

In this application, instead, it was decided to implement only R^2 because it would be meaningless apply YEC when it was built the model for DHW and SH splitting. The reason why this happens is explained in detail in the chapter 4.

With the literature analysis and the insertion of references for this last tool it is concluded the literature review. It must be emphasized the fact that in this site there are proposed some articles per argument and there's the possibility to include a lot of other articles. Nevertheless, in this review they are presented only the most relevant articles and the ones that contain information about the law and the principles that have been implemented in the analysis. In each paragraph it is highlighted what research are used in the effective analysis, by inserting a reference to the chapter 4 or the chapter 7.

Chapter 4: Description of method with a complete database

after a chapter dedicated to the literature review, in this site it will be discussed the application and the method followed for the analysis of energy consumption. In this chapter It will be analysed buildings with a complete database. By saying that it is meant that there must be data from meters collected with hourly resolution for an entire year. The only data used in these tools is the hourly global energy consumption. It will be described better the characteristic of the case study in the chapter 6, but for now it must be said that the only component to the demand of thermal energy is the Domestic Hot Water (DHW) and the Space Heating (SH). In addition to that, every building taken in analysis is part of a district heating network.

It will be analysed in the next chapter, the number 5, the case of an incomplete database.

The paragraphs presented in this chapter correspond to the different tools implemented for the analysis. In every section it will be described step by step the formula and the reasonings behind every passage and it will be underlined what is taken from literature and what are the contributes from the author of this work. It was not possible to make a detailed list at the beginning of the thesis because modifications and improvements are done in distinct parts of the whole analysis and in different tools.

The steps are: Data pre-processing; Hourly data splitting; Energy Signature Curve reconstruction, including discussion about the choice of every parameters useful for the construction; Boxplot of energy consumption; Analysis of temperature lag; Implementation of Confidence Intervals; DHW boxplot analysis; Splitting of DHW and SH energy consumption; Detailed DHW consumption analysis, with all the sub-steps present in this part of the examination.

4.1 Data pre-processing

This is a crucial phase of the entire analysis. It is impossible to obtain satisfactory results if the starting dataset contains errors, missing values, or inaccuracies. The pre-processing phase includes all the operations that must be done in the initial datasheet. For this phase it was used the program Excel, but any software that gives the possibility to deal with tables Is ok.

It will be described in the chapter 6 the characteristics of the case study of Tartu, the one analysed that present a complete dataset for every building. In this site it was wanted just to show all phases and operation of the work, so it was decided to take a building as an example. It is considered building named "10922" just for take some images and table, useful to better understand the steps of the analysis.

To understand what operations, must be done in this step it is wanted to show you an example of datasheet, in the figure 7. This is taken from the building "10922", as said before. This thing will not be specified in the following part of the explanation.

METERID	ENERGY	VOLUME	HOURS	POWER1	FLOW	FLOW_TEMP	RETURN_TEMP	HOT_WATER	SEC_FLOW	SEC_RETURN	READ_DATE
10922	31.105	858.43	4550	82.7	2313	75.88	44.81	46.06	52.61	43.27	martedì 1 gennaio 2019
10922	31.188	860.75	4551	82.8	2312	75.97	44.82	46.74	52.61	43.34	martedì 1 gennaio 2019
10922	31.273	863.1	4552	87.5	2415	76.11	44.49	45.58	52.48	43.34	martedì 1 gennaio 2019
10922	31.359	865.48	4553	86	2416	75.7	44.81	46.75	52.67	43.34	martedì 1 gennaio 2019
10922	31.446	867.93	4554	83.8	2381	75.54	44.95	46.05	52.79	43.4	martedì 1 gennaio 2019
10922	31.531	870.32	4555	82.4	2349	75.51	45.02	46.43	52.86	43.53	martedì 1 gennaio 2019
10922	31.615	872.73	4556	82.5	2364	75.43	45.08	46.33	52.86	43.53	martedì 1 gennaio 2019
10922	31.697	875.1	4557	82.5	2417	74.71	45.02	46.81	52.67	43.46	martedì 1 gennaio 2019
10922	31.779	877.5	4558	83.5	2409	75.16	45.02	45.99	52.86	43.46	martedì 1 gennaio 2019
10922	31.863	879.9	4559	87.6	2475	75.68	45.07	46.35	52.98	43.65	martedì 1 gennaio 2019

Figure 7 Example of a complete database (building 10922)

In this file we can notice some relevant and some useless information. Starting from the left there is the name of the meter. Every building in this case study had its own meter. This is also used to identify the building because for privacy reasons it was impossible to know addresses, location, or any other sensitive information.

On the second column there is the voice "ENERGY." Here we can find, expressed in MWh the cumulative thermal energy consumption of the building. This is a particularly important data. It could be missing in some dataset (it is evident, for example, in the analysis of the chapter 5, the section about incomplete datasheet). This is not a problem, because it can be calculated with the following law:

$$E = m * cp * \left(T_{flow} - T_{return}\right) \quad [4.1]$$

Where E is the energy, m is the mass of the fluid, T_{flow} is the temperature of the flow and T_{reutrn} is the return temperature of the fluid, after the heat exchange. Particular attention must be done to the units of measure. In the analysis it was kept kWh as a measure of energy consumption, to have a clearer vision of the results and because it is the common unit used when we talk about building energy consumption.

If data is taken from this datasheet, it must be reminded the fact that the mass is absent, but in the third column is present the "VOLUME." Here we can find the cumulative water consumption expressed in m^3. It is possible to obtain mass from this data by applying the definition of density of water, expressed in the law [4.2]. Because we are talking about temperature of water around between 40°C and 80°C, the value considered for density is 980 kg/ m^3.

$$m = V * \rho \quad [4.2]$$

For double checking the fairness of data it was compared the hourly energy obtained by subtracting two following hours taken from the cumulative energy consumption column ("ENERGY") with the result of the application of the equation [4.1] hour per hour. The two values are the same.

Proceeding to the right there is the column "HOURS." This is a progressive number which increases by one every hour. It is a relatively useful information, because the knowledge about day and hour of the day comes from the last column to the right named "READ_DATE." This is an especially essential information and allowed me to select any days with missing values or with double identical values.

There was, in fact, some days that presented some errors in the reading of meters. It was found days with 25 hours, because there were two close lines identical. So, it was cancelled duplicate. They are extremely dangerous because, when it was calculated the hourly energy

consumption as a difference between two following values of the line "ENERGY" there is the creation of 0 in case the difference is done between two following identical values. These 0 create big inaccuracies in the demand profile with consequent problems in the following steps of the analysis.

Proceeding to the right there is the column "POWER1". These data represent the mean hourly thermal power of the plant expressed in kW. It is evident that this data, in numerical terms, is equal to the energy per hour expressed in kWh, previously mentioned.

On its right there are the columns called "FLOW," where the is the information about the hourly water consumption expressed in l, "FLOW_TEMP" and "RETURN_TEMP" which represent the temperature of water, in °C, before and after the thermal exchange with the building. These values are used for the evaluation of the thermal energy per hour.

The other columns, "HOT_WATER," "SEC_FLOW" and "SEC_RETURN" are representing temperatures in different zones of the water circuit, but they are not used in this analysis. In the figure 8 it is possible to find the scheme of the plant, with the position of all meters that collected data mentioned in this paragraph.
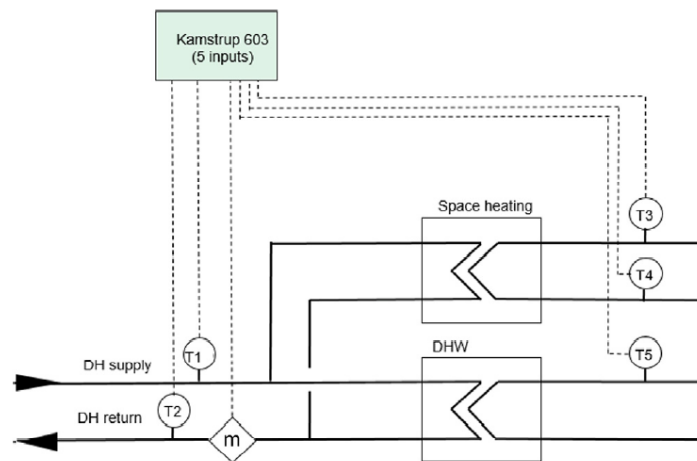
Figure 8 Location and lay out of the smart energy meters in the District Heating in Tartu.

Clearly this scheme has as reference Tartu's building. In other types of datasets there could be different measurements. This figure (8) is shown for clearness of exposition. The fundamental input for the analysis, by the way in the hourly global energy consumption. If this data is not available, it can be calculated by using flow consumption per hour and difference between inlet temperature and outlet temperature.

On the other hand, the other input is the outdoor temperature. The only request is that there must be a database with hourly value of this variable. It can be easily found from meteorological station close to the building in analysis or it can be measured with a specific meter near the building itself. If these data are missing, it is possible to use daily mean temperature values, but the analysis will be less accurate.

4.2 Hourly data splitting

This is the first operation of the analysis. This step and the followers were implemented in MATLAB R2020b. All the functions used and the general scripts for this step and the following ones can be found in the appendix of the thesis. As it was said this thing happens for every step, so it will not be repeated this concept for every paragraph.

It is relevant to underline the fact that from this point of the analysis, the implemented script in MATLAB is completely automatic, so there is no necessity of intervention after the pre-processing operation. This means that this algorithm can be applied in a fast and uncomplicated way. It is, also, general, so the tools implemented can be applied to every building even outside from the considered dataset. In other words, it can be used to study buildings in every climate zone, as long as a completed year dataset and weather data are available.

The concept at the base of the hourly data splitting operation is to subdivide the global energy demand into smaller vector for a better analysis of the consumption. It was created 24 vectors, one for each hour of the day. Each vector contains 365 values. Each value represents the global energy consumption of the hour indicated by the name of the vector per each day of the year. For example, the vector called "ENERGY_0AM" is composed by the global thermal energy consumption of the hour that goes from 0.00am to 1.00 am of the 1$^{st\,of}$ January, then the same hour of the 2nd of January and so on. In the equation below there is a clearer representation.

$$ENERGY_{0AM} = \left[DEMAND\ 0AM_{1JAN}, DEMAND\ 0AM_{2JAN}, \dots DEMAND\ 0AM_{31DEC} \right]\ [4.3]$$

It was done this operation for every hour of the day. These vectors are the starting point for the operation of DHW and SH heat lead consumption splitting. This procedure, in fact, is done both for hourly vectors and for the global vector containing the 8760 values of demand for the year.

The scope of the hourly data splitting is to identify hours with an eventual particular consumption. It was developed this type of analysis starting from the considerations done regarding the same dataset in the article of Lumbreras M., Garay-Martinez R. et al. From their work, in fact, it emerged that some buildings could present a night setback or a reduction in the demand in specific hours of the day. The operation of hourly data splitting offers the possibility to find if particular heat loads happen always in the same hours of the day.

Regarding this argument, it is possible to find, for example, moments of the day of typical maximum consumption or if there is an operation of the plant different from the usual. This is the case of the partial dataset analysis, where thanks to this tool it was possible to find hours where the space heating plant was on operation and when it was turned off. For seeing the results and the comments of this fact, check the paragraph 7.3 in the chapter 7.

Another important possibility given by this tool is that it represents the initial operation for the evaluation of the thermal losses in the hot water circuit. This fact is explained in the sub-paragraph 4.9.1.

4.3 Energy Signature Curve reconstruction

It was discussed about the meaning of Energy Signature Curve and its literature background in the literature review in paragraph 3.2.2. In this site it will be explained all mathematical passages and algorithms used for the construction of the curve. As it was said in the previous paragraph, this operation is done both for hourly energy vector and for the global energy vector, with all consumption values of the year.

The scope of building the energy signature line is obtaining a model for the estimation of the space heating component of the demand. From this curve it is possible to obtain this part of the thermal demand, as happened in the study conducted by Ivanko D. et al., already mentioned in the literature review. The focus on this passage can be found in the paragraph 4.9 of the current chapter.

The first operation is the calculation of the Daily Mean Temperature (DMT). This is an optional operation. The analysis, in fact, can be done also with hourly values of outdoor temperature. It was decided in this site to use DMT for operating with hourly energy vector, and to use hourly values of outdoor temperature for the global curve. Some authors use DMT, for example the already mentioned Petersen L., and others use hourly values.

DMT was calculated by considering the average value of 24-hourly outdoor temperature day per day, as showed in the equation 4.4. So, there is the creation of a 365-values vector named "DMT".

$$DMT_i = \frac{(T_{out\ 0am\ i} + T_{out\ 1am\ i} + T_{out\ 2am\ i} + \cdots + T_{out\ 23pm\ i})}{24} \quad [4.4]$$

At this point there is the first phase of the ESC reconstruction, which is the individuation of the CPT.

4.3.1 Determination of CPT

Determination of the Change Point Temperature is a crucial point in the ESC analysis. As shown in literature review, in many applications this point is found by experience of the authors or by previous data regarding the building or the location in analysis. All these methods are effective, but with the problem of the necessity of an expert or availability of data regarding the begin and the end of the heating season. These data are usually hard to find, especially if we are talking about buildings where information about users' behaviour is absent for privacy reasons. In addition to that, the length of the heating season is dependent on the climate as well as the type of building, isolation thickness, control system, consumers, and more.

For solving these problems, in this analysis it was used a mathematical approach for the determination of the CPT. With this method no extra information are needed, except for the already mentioned as input parameters. Guidelines and general steps for the method are taken from the work of Petersen L. about a method for load modelling of heat and electricity demand. This is a compete statistical approach.

Role of the change point temperature is to find the point of switch between two different regimes in the behaviour of the heating plant of the building. The temperature independent part (on the right of the figure 10) and the temperature dependent one (on the left of the figure 10). In the first part there is the presence of only DHW, in the other side there is the sum of DHW and SH.
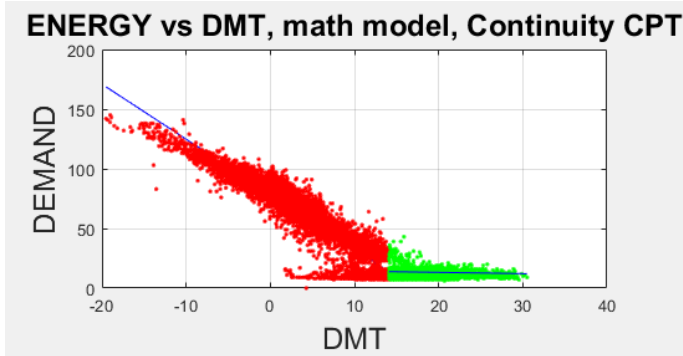

Figure 10 An example of ESC curve. In green the T-independent part, in red the T-dependent part
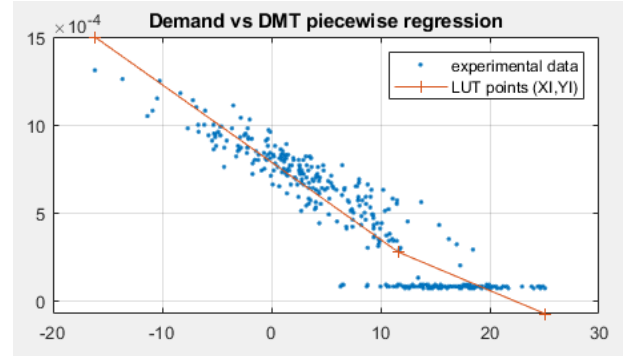

Figure 9 An example of test for the piecewise regression with the yearly global demand between 0.00 am and 1.00 am.

To perform an accurate regression analysis, it is crucial to distinguish between temperature dependent and temperature independent consumption. It can be done even with other methods, such as a piecewise regression, maybe by implementing an automatic for the operation. There were some attempts to create the ESC without the separation in two zones made with CPT using various method: BEAST algorithm developed by Zhao et al. or the tool for creating a continue interpolation curve by optimization toolbox in MATLAB, but none of them give satisfactory results.

In figure 9 there is the example of a test done with the algorithm of automatic piecewise regression, and it is evident that even if for the temperature dependent part, the interpolation is good, for the temperature independent part it is completely wrong.

The only way to perform a good ESC rebuilding is to identify CPT, better if this operation is done in a mathematical automatic way. It was report here the general expression of the regression equation [4.5] for the clearness of exposition.

$$E_i = \alpha_i + \beta_i * DMT_i + e_i \quad [4.5]$$

Where E_i is the global energy consumption of the considered hour (for example between 0.00 am and 1.00 am) at the i-day measured in kWh, α_i is a coefficient obtained from linear regression of experimental data expressed in kWh, β_i is a coefficient obtained from linear regression of experimental data expressed in kWh/°C, DMT_i is the daily mean temperature of the considered i-day and e_i is residual (the error in the fit).

Researchers, as established in Petersen's work, have developed a mathematical methodology for determining the division between temperature dependent and non-dependent seasons, i.e., for determining CPT. The slope of the regression equation in equation [4.5], represented by the β value, indicates the degree to which the heat load

decreases as the mean daily temperature increases. The α and β values are determined via the least squares method.

To identify the temperature dependent season, it was calculated β values for temperatures decreasing from the maximum temperature registered in the dataset (for Tartu's analysis in it around 23°C, but this value depends on the location in analysis) to the minimum temperature registered in the dataset (even this value depends on the location, for Tartu's analysis it is around -18°C) in increments of 0.1°C. The objective is to identify the temperature range where the β value remains relatively constant, fluctuating only within a few percentage points. The CPT, so it means the beginning of the temperature dependent season, is located in the region where β passes from relatively constant value to a rapid decreasing.

β, in fact, is almost constant when there is no correlation (or a low correlation) between thermal consumption and outdoor temperature but begins to be lower and lower (in relative terms, in absolute terms it becomes higher) when the correlation between demand and temperature becomes stronger (it means we are in the temperature dependent region). In figure 11 it can be easily seen that CPT is around 11.5°C.

The first part of the curve on the right does not count as variation: in this zone a low number of experimental points are considered, and this fact produces the instability in β value that can be notice in this plot.
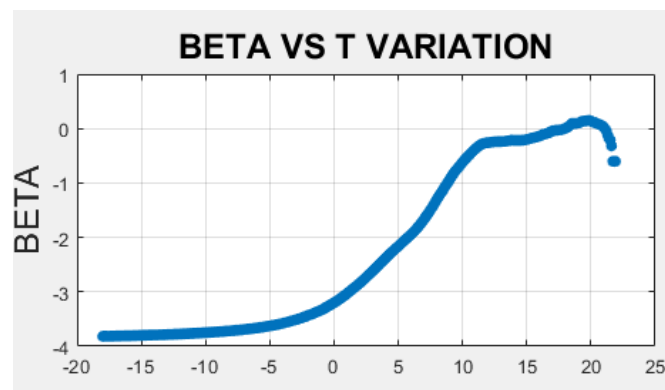
Figure 11 example of β variation with increasing temperature intervals

This operation is performed for demand of every hour of the day and for the global demand. Once all CPTs are identified, their values must be inserted in the special place of the general script (it is the section "CPT definition" in the general script "ESC_10922" that can be found in the appendix of the thesis).

Even if the operation is done for every hour, this is not necessary. It implemented this function just for controlling the similarity between CPT of different hours. It is confirmed, after various simulations using different buildings, that CPT value is similar for every hour. This fact is correct: CPT represents the temperature at which space heating plant is turned off. Because control of the space heating plant depends on outdoor temperature, it has perfectly sense that there is the same behaviour during different hour of the day.

4.3.2 Splitting of two zones in ESC

At this point, change point temperature value is known. This means that experimental points can be easily divided into the temperature dependent part and the temperature independent part.

This operation is done by the function "ESC_DIVISION" that can be found in the appendix. The result is the individuation of the red zone and the green zone showed in figure 10.

The other operation done in this step is the linear interpolation of the two zones. Each interpolation is performed independently from the other. In this way, of course, there is no continuity in CPT, as you can see from the figure 11. This problem is addressed in the next paragraph.
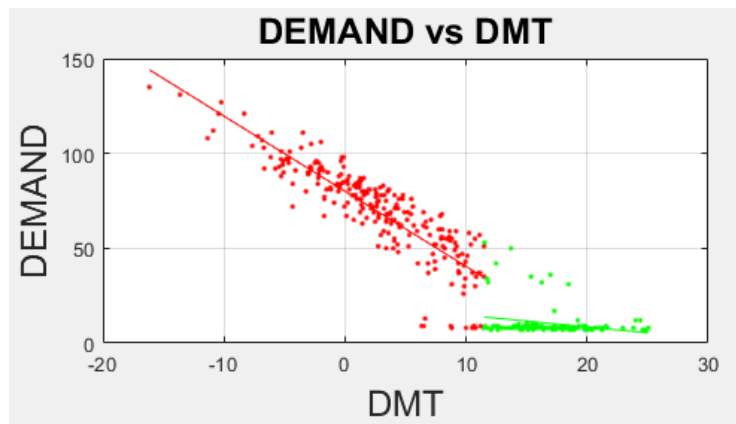

Figure 12 Example of ESC without continuity in CPT. It is evident from the behaviour of red line and green line.

4.3.3 Reconstruction of ESC

At this point, the next goal is to build an ESC continue in CPT. The red line and the green line in figure 12 are good interpolation of their respectively experimental points. For obtaining the continuity in CPT, the first step is writing the equation of the linear regression model in the temperature dependent part and the temperature independent part, as provided by the work of Tereshchenko T. et al. (equations [4.6])

$$E = \alpha_1 + \beta_1 * x \quad if \ x \leq CPT$$

$$E = \alpha_2 + \beta_2 * x \quad if \ x \geq CPT \quad [4.6]$$

Where E in the global energy demand, α_1 is a coefficient obtained from linear regression of experimental data expressed in kWh of the temperature independent part (α_2 has the same meaning but for the temperature dependent part), β_1 is a coefficient obtained from linear regression of experimental data expressed in kWh/°C, of the temperature independent part (β_2 has the same meaning but for the temperature dependent part) and x is the outdoor temperature.

By imposing continuity in x=CPT, E must be equal in this point, so:

$$\alpha_1 + \beta_1 * CPT = \alpha_2 + \beta_2 * CPT \quad [4.7]$$

Obtaining one parameter from this equation:

$$\alpha_2 = \alpha_1 + CPT * (\beta_1 - \beta_2) \quad [4.8]$$

By replacing α_2 in the equations [4.6] it is possible to obtain a mathematical expression of the ESC before CPT and after CPT. It assumes the shape of a piecemeal function. [4.9]

$$E = \alpha_1 + \beta_1 * x \quad if \ x \leq CPT$$

$$E = \alpha_1 + CPT * (\beta_1 - \beta_2) + \beta_2 * x \quad if \ x \geq CPT \quad [4.9]$$

By implementing these equations, it was obtained a valid representation of ESC, continue in CPT as it can be seen in figure 13. It is important to underline the fact that with the application of these equations, there is no changes in regression coefficients. This means the interpolation stays good. It is only an operation of shifting up for the temperature independent part. The function for doing this operation is called "ESC_CONT_CPT" and can be found in the appendix of the thesis.
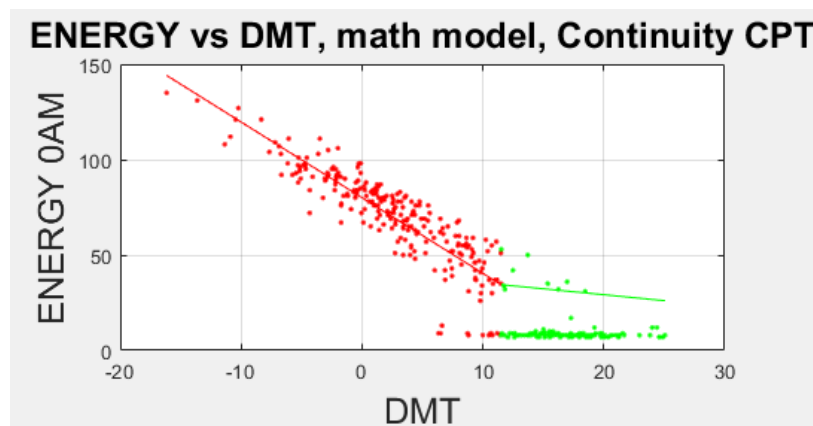
Figure 13 An example of ESC with continuity in CPT

4.3.4 Accuracy parameters

In this last section of the ESC construction, it was faced the problem of accuracy of the model. As introduced in literature review, for checking the goodness of the model it was used the parameter R^2, defined by equations [3.5, 3.6, 3.7]. It is calculated for the temperature independent part and the temperature dependent part independently one from each other.

This parameter is and indicator of how close the estimation is, is done by linear regression model to the real experimental data. In the results section (Chapter 7) it will be presented detailed values for this parameter. In this paragraph it is only mentioned that, in general, this model of ESC gives particularly reliable results in terms of approximation to real data. Values for R^2 are between 0.75 and 0.85 for Tartu's simulations.

4.4 Boxplot of energy consumption

An analysis of global energy consumption with the boxplot tool is an interesting choice for many reasons. First of all, it provides clear values of the maximum, median and minimum energy use, and density of data points. Then it is a useful instrument for the detection of outliers. Last, but not least, if performed by subdividing demand into weekdays, Saturdays, and Sundays it can provide some useful information about different patterns that occur in different days of the week. Because of this fact, it is possible to obtain information about users' behaviour and, also, some general clues about the use of the building, information not always known.

4.4.1 Separation of weekends and weekdays

The first step of this analysis consists in the subdivision of the global demand into weekdays, Saturdays, and Sundays. This is done by checking what is the first day of the year (in the analysed dataset all measurements came from 2019, so it was a Tuesday) and implementing the operation shown in the section "boxplot analysis" of the general script "ESC_10922" that can be found in the appendix.

4.4.2 Creation of the boxplot

After the subdivision, the boxplots are created using the function "boxplot" from the MATLAB's database. In figure 14 it is possible to notice an example of boxplot. For every building four different boxplots are crated: the first for Saturdays, then Sundays, then rest of the weekdays and the last one is for the global demand analysis.
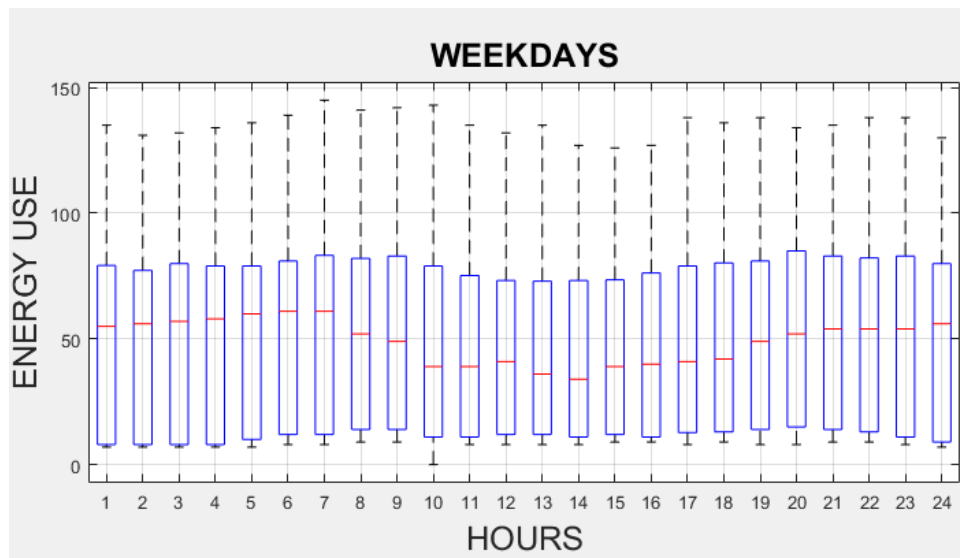
Figure 14 Example of boxplot. In this case it represents thermal energy consumption during weekdays

Indicated with red lines there are mean values for each hour of the day. For example, in correspondence of number 13 in the "HOURS" axis, there is the median value for global energy demand of all weekdays (if we consider the example on the figure 14) for the hour between 13.00 pm and 14.00 pm. Upper and lower values of the blue boxes are,

respectively, the 75% and the 25% percentiles. the whiskers extend to the most extreme datapoints the algorithm considers to be not outliers, and the outliers are plotted individually with a "+" if they are present.

4.5 Analysis of temperature lag

This part of the method consists in the implementation of an analysis of energy consumption versus a temperature moving average. The moving average is an operation that, starting from a vector of elements, produces a new vector with the same number of elements of the initial one, but where numbers are mean values of a window of elements of the first one. In simple terms, it consists in substituting every element with the average value calculated among n values around the considered element. For doing this type of analysis with outdoor temperature values, it was followed indication provided in the article by Tereshchenko T. et al., as explained in the literature review paragraph about this topic (3.2.6).

The MATLAB implementation is explained in the following lines. First, it was calculated temperature moving average vector, shifting the window for mean value calculation of 1 hour per each step, for 48 hours. In this way it was created 48 temperature moving average vectors. The first one is identical to the original outdoor temperature vector. The second one, in each i-element it has the mean between i and i-1 element. The third one has the mean among i, i-1 and i-3 elements of the original outdoor temperature vector and so on. It was used function "movmean" already present in MATLAB's database for the implementation.

After that it was performed a linear interpolation of measured demand points using the created TMA vectors as temperature values. Thanks to this operation it was obtain 48 β values (see equation 4.9) for the regression coefficients. As demand values for this type of demand it was considered only points before CPT, to be sure of being in the temperature dependent part of the heat consumption.

By plotting these 48 β values versus hour lag (a vector that goes from 1 to 48 and indicates how many elements are considered for the TMA calculation, i.e., the number of previously outdoor temperatures considered) it is possible to obtain for each building a plot. The maximum value of β (in absolute term because they are always negative) indicates the better fit between TMA and heating energy use. This is a good indication for accounting the thermal inertia of the building, because better correlation is obtained when previous values of outdoor temperatures are considered.
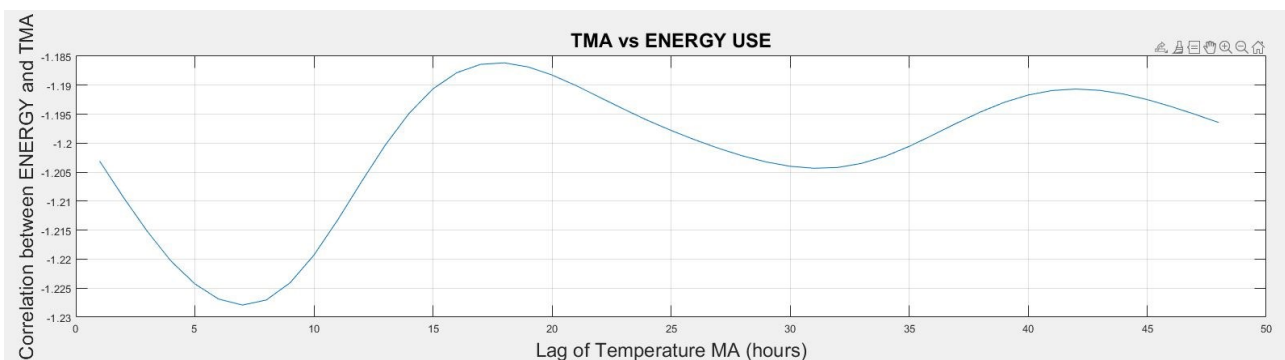

Figure 93 An example of plot showing the effect of temperature lag on energy consumption of the building.

It was reported an example of this type of plot in figure 93, highlighting the point of largest correlation. For the building in examine it is possible to say that the thermal inertia acts with an effect of 7 hours. This is a useful information in case a more detailed model for heat consumption is implemented. Usually in addition to the variable Tout, also in the regression model is inserted a term that depends on the TMA of 24 previous hours. Thanks to this analysis it is possible to improve this type of model, saying that it is better to consider 7-hours TMA instead of 24-hours TMA for this specific building in analysis (but this value depends on the building's construction type and time constant parameter typical of the structure itself).

4.6 Implementation of Confidence Intervals

Confidence intervals are especially useful for outliers' detection and as a filter in case data measured present a high dispersion. This is particularly useful not only for excluding points that could lead to possible errors and problematic results, but also with the goal to make easier any future work on these data. From many studies, in fact, it emerged that clusterization algorithms (programs that automatically create standard profiles from a large set of data) do not work well if data present high dispersion. This is evident in the analysis of Tereshchenko et al., just to mention one article. Their role as outliers' filters will play a vital role in the Verona's datasets analysis.

Their implementation follows guidelines given by Tereshchenko et al. in their already mentioned article about the analysis of energy signatures in buildings in Norway. Another article that provides useful formula and information for using in an appropriate and effective way confidence intervals is the work by Altman D.G. et al., that talks about the use of confidence intervals for data analysis in medical applications. However, being a statistical instrument, it can be used to study any type of data, in particular in cases when there is the necessity to examine one or two groups of individuals with respect to a single variable.

The goal of the application of this tool is the construction of two parallel lines, one above and the other below the regression line, which represents the interpolation of experimental data. In this way there is the creation of a region where measured points are expected to be. Now it will be discussed about how it is possible to determine the equation of the two lines of confidence interval's boundaries.

The general method for the construction of confidence intervals is quite simple. Each confidence interval is calculated by adding and subtracting a multiple of the estimated statistic's standard error (SE) or a transformed version of it.

The multiple is determined by the theoretical distribution of the statistic, such as the t-distribution for regression or the Normal distribution for correlation. To obtain the multiple, we need to identify the value that corresponds to including the central 100(1-a) % of the theoretical distribution, where "a" represents the level of significance. For instance, a 95% confidence interval is obtained by finding the value that cuts off 2.5% from each tail of the distribution. In this case we are using a level of significance equal to 5%.

There are many sites and tables in literature of the t and Normal distributions that can provide appropriate values. In the script implemented for the CI analysis, these values are obtained using MATLAB's database, in particular the function "tcdf" for obtaining values

regarding t-Student's distribution. Relevant variables are $t_{(1-a/2)}$ or $N_{(1-a/2)}$. For the t distribution the degrees of freedom, which depend on the sample size, must be known.

Starting from the degrees of freedom, its expression is given in equation [4.10]. in this expression, n is the sample size.

$$n_{free\ deg} = n - 2 \quad [4.10]$$

Another necessary parameter for the evaluation of confidence intervals lines is the residual standard deviation of y (in this case this letter represents experimental points of global energy demand) about the regression line. Its expression can be found in equation [4.11]:

$$S_{res} = \sqrt{\frac{(n-1) * \left(S_y^2 - b^2 * S_x^2\right)}{n_{free\ deg}}} \quad [4.11]$$

Where n is the sample size, S_y is the standard deviation of the global energy demand, S_x is the standard deviation of the outdoor temperature, b is the sloping coefficient of the linear regression line, obtained in the previous paragraphs and $n_{deg\ free}$ is the variables of degrees of freedom of the sample data.

As said before, value $t_{(1-a/2)}$ is obtained from the MATLAB's database, using the function "tcdf" (which implements the cumulative distribution function of a t-Student's distribution, using as input parameters the significance level a, and the number of degrees of freedom. So, the equation for this parameter is not an equation, but an expression for software implementation [4.12]:

$$t_{\left(1-\frac{a}{2}\right)} = \text{tcdf}\left(1 - \frac{a}{2}, n_{\deg free}\right) \quad [4.12]$$

As highlighted by Tereshchenko et al. in their study, The shape of many sample distributions when we are talking about buildings' energy demand can be approximated by a normal distribution. A convenient aspect of normal population distribution is that we can apply 95% confidence interval to describe desired population range. So, this is also the case of this analysis. It was applied the same reasoning in the presented analysis.

In the equation [4.13] there is the expression, point per point, of the higher (if the law is read with sign "+") or lower (if equation is read with sign "- ") confidence interval.

$$C.I. = Y_i \pm S_{res} * t_{\left(1-\frac{a}{2}\right)} * \sqrt{1 + \frac{1}{n} + \frac{(x_i - x_{mean})}{(n-1) * S_x^2}} \quad [4.13]$$

Where Y_i is the predicted value of energy use, obtained from the linear regression model, S_{res} residual standard deviation of y about the regression line, $t_{\left(1-\frac{a}{2}\right)}$ is the value from Student's criteria, n is the sample size, x_i are the actual values of the independent variable (in this case it represents the outdoor temperature), x_{mean} is the mean value of the independent variable (outdoor temperature), S_x is the standard deviation of the independent variable (outdoor temperature).

Thanks to equation [4.13] it is possible to draw confidence intervals lines for every experimental points distribution. It is represented in a plot where global energy demand is plotted versus outdoor temperature (figure 15). Any comments about results of the C.I. application will be done in chapter 7.
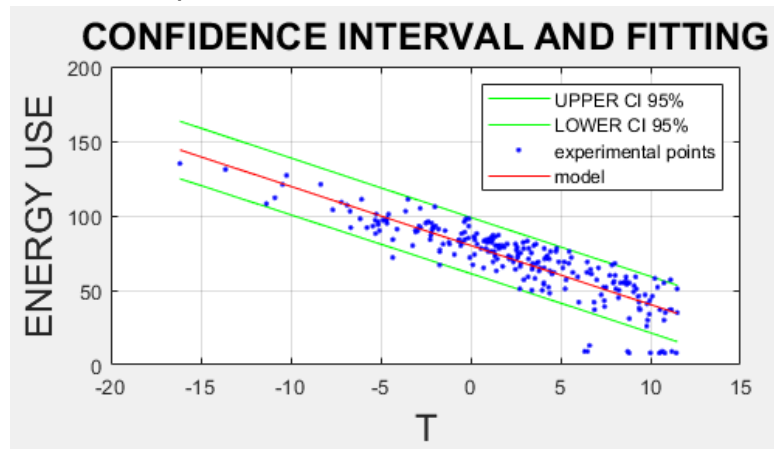

Figure 15 An example of plot with Confidence Intervals implementation

It is important to underline the fact that this type of analysis is done only in the temperature dependent part. This happens because in the temperature dependent part, points outside from confidence intervals may represent special point of use, particular days, or particular hours of the day. The application is, for this reason, much more interesting in the temperature dependent part, rather than the temperature independent part, where the demand is much more regular.

4.7 DHW mean profiles analysis.

This type of analysis is interesting not only for seeing the typical users' profile during certain month, but also for obtaining an estimation of losses into hot water circuit. The aim of this step is to have a confirm of what are the months where heat demand is temperature independent. As said in previous analysis, the temperature independent part of the demand happens when outdoor temperature is higher than CPT. This happens, of course, during summertime. This means that in summertime there is the presence of only DHW consumption.

With a representation of mean profiles for heat demand during summer, it is possible to note if there are differences between summer (months of June, July and august) and May and September. Months of summer are not the same for every building or for every location. This analysis helps also to identify which months belong to summertime. It could happen that in some cold climates also during June it is necessary space heating. This will produce a higher mean monthly profile than, for example, July.

By not knowing the intended use of the building in analysis, subdivide weeks into weekdays and weekend could give some clues about that. In residential buildings usually profiles in summer, due to the fact that they represent only DHW heat load demand, are similar for weekdays or weekends. If a building is used as office space or as commercial building, for example, it is possible to identify some differences in profiles of weekdays and weekend, just because during these last days there is no activities in the building.

So, the first step is to identify months of May, June, July, August, and September. It was done this operation in the section "DHW mean profile analysis" of the general script "ESC_10922" in the appendix. This procedure is very simple because all databases start from the 1st of January, so just select values of interest from the global energy vector by inserting the number of day of the year where each month starts and finishes. For May, for example, in every year (apart from in case of leap year) it starts from day 121 and ends with day 151.

The second step is to identify in each month weekend and weekdays. The operation is very similar to the one described few lines above. By knowing that the 1st of January of 2019 (year of the database) is a Tuesday, it is possible to implement an algorithm that can automatically select weekend and weekdays within each month.

After the selection of the month and the subdivision in weekend and weekday, a mean profile for each hour of the day is calculated. An example of the results of this operation can be seen below, in the figures 16 and 17. For example, in correspondence of the index "18" in the time axis in figure 16, mean value for all weekdays (that have also called workdays) is represented for May (above) and for June (below). Same concept but for weekends is represented in figure 17.

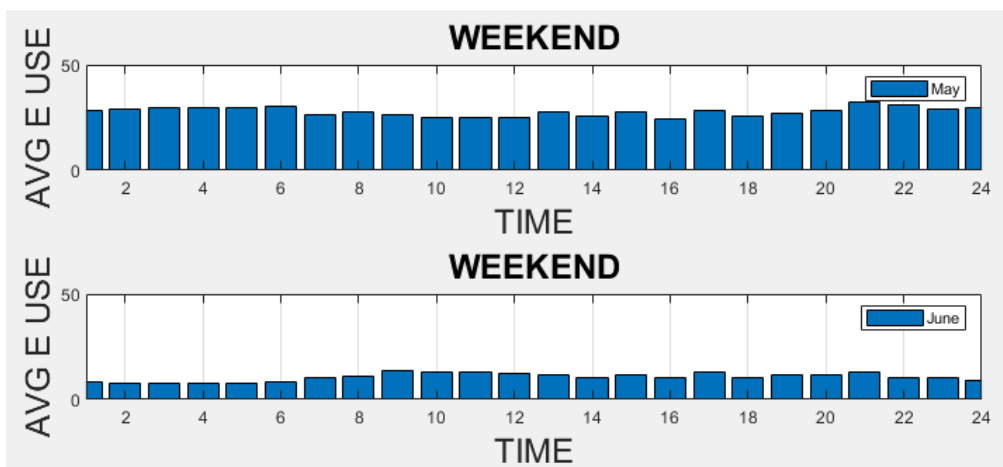*Figure 17 An example of DHW mean plot analysis of workdays of May and June*

Figure 16 An example of DHW mean plot analysis of weekends of May and June

It can be noticed a high difference between mean profiles in June and in May for the example proposed in figures 16 and 17. This is just an example and, as happened before, comments

45

on results and real case studies are done in chapter 7. Apart from that, this is a clear example of the relevant information that can be obtained from the application of this tool.

4.8 Splitting of DHW and SH energy consumption.

All tools presented until this point provide certainly interesting information, but they are also preparatory for this phase. In this paragraph it will be explained how it is possible to split the global energy demand obtained from measures into a component of DHW and the other of SH.

The starting point is the energy signature curve, continue in CPT, built in paragraph 4.3 of the current chapter. For the following passages it was used as references the work by Dmytro I. et al. about the DHW and SH splitting in a hotel in Norway. From their work, it emerged that ESC represent an incredibly good model for the space heating consumption. In fact, as it happens for the space heating, when outdoor temperature decreases, heat demand for heating up the indoor environment increase. When outdoor temperature becomes higher than a certain value (CPT), space heating is useless, and it is turned off.

However, as the model considers the total heat use, which includes domestic hot water (DHW), a shifting coefficient was introduced to adjust for this. The shifting coefficient was determined by analysing the behaviour of the space heating (SH) system during warm seasons, when the outdoor temperature exceeds the critical performance temperature (CPT). During these warm seasons, there were times when the SH heat use was zero, and the research of Tereshchenko et al. demonstrated that the minimum value of the energy signature curve (ESC) coincided with these periods.

Together with the study of Ivanko D. et al. conducted in a hotel, a previous study by the same authors proved same results for two other buildings. For this reason, it was used the minimum value of the ESC as the coefficient of shifting, and this enabled me to derive a reliable model of SH heat use by extracting this coefficient from the ESC. From these considerations, the equation for the model of SH is the [4.14]:

$$E_{SH} = f(x) - \min(f(x)) \quad [4.14]$$

Where $f(x)$ represents the point obtained by ESC model and x in the independent variable, i.e., the outdoor temperature.

The total heat use values, which exceed the modelled SH heat use, provide insights into the trend of DHW heat use. This fact is highlighted also in the work by Bacher P. et al. This work was already mentioned because their intuition to use heating spikes in the global energy demand for splitting domestic hot water and space heating. In this analysis, DHW consumption is obtained by a subtraction between the experimental point (global energy demand) measured from smart meters and E_{SH} obtained by ESC model shifted down.

However, in all buildings taken in analysis, there is a DHW system with continuous circulation where hot water is delivered continuously, because they are connected to a district heating network. For this reason, the system losses must be added to the DHW heat use obtained from the residuals. These losses can be calculated as illustrated in the following sub-paragraph.

4.8.1 Evaluation of DHW circuit losses

For the evaluation of DHW circuit losses it was followed the method proposed by Marszal A. P. et al. in the first part of their analysis regarding a simple methodology to estimate the mean hourly and the daily profiles of domestic hot water demand from hourly total heating readings. This article was mentioned in the literature review, and also, it is reported the general flowchart of their methodology in figure 3. In this site it will be explained better all passages that were followed.

Starting from the annual heat demand profile, it was selected summer period. This could be identified by using considerations from the paragraph 4.7 of the current chapter. It is important to select only summer period because it is fundamental to have only DHW consumption in the building. If there is presence of SH there is a great overestimation of the losses in the circuit.

For all buildings it was selected June, July, and August as month of summer period. By analysing the energy demand profile versus time, it is possible to note that even in some days of May and of September SH is still turned off. It was preferred to not consider these periods for avoiding the risk of considering some days with SH heat load consumption and, as a consequence, obtain some errors in the creation of mean summer profiles of consumption.

Days selected goes from day number 152 to day number 243 (from 1st of June to the 31 of August). In figure 18 it is evident that between these two days demand is much different than the rest of the year.
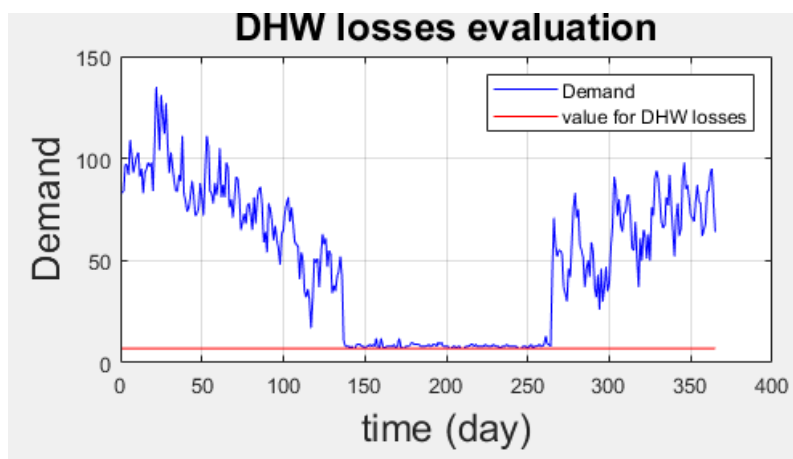
Figure 18 An example of yearly demand plotted vs day of the year and a representation of DHW losses value.

After summertime days selection, the next phase consists in creating one mean profile for summer consumption. Hour by hour, mean values of global heat consumption are obtained. At this point a mean profile for summertime period is available.

After this operation there are two possible methods that can be followed. The method 1 consists in finding the minimum value of demand of the profile created in the previous step. This minimum value is assumed to be the DHW circuit losses value. The method 2, on the other hand, calculates DHW circuit losses value using a mean between 0.00 am and 4.00

am of the demand of mean summertime period profile. It was implemented both method in the script and it is possible to choose what is the value for losses by modifying the special value in the function "SHDHW_split_function" that can be found in the appendix.

The idea at the base of method 1 is that in correspondence of the minimum value of standard summertime period profile people are not at home. Due to their absence, heat load consumption for domestic hot water must be zero. This fact means that if there is a value different from zero, it must represents losses. There is high probability that people are absent from home in a certain moment of the day based on various studies about Europeans' behaviour. One example comes from the author K. Winqvist, who investigated how Europeans spend their time everyday life. He found that on a daily basis, individuals typically spend approximately 17 hours in their homes. However, this number varies based on age, with those aged over 65 spending around 20 hours at home daily, while the youngest group aged between 15-24 spends around 14 hours at home daily.

The idea at the base of method 2, on the other hand, is that the other moment of the day when there is no DHW consumption is during the night, when people sleep. So, for that reason the period between 0.00 am and 4.00 am is considered. During this time window, in fact, more than 90% of Europeans sleep, as reported by the study of K. Winqvist, mentioned above.

According to the results of the study by Marszal A. P. et al. and based on simulations done on the Tartu's database, there are no substantial differences in using method 1 or method 2. With method 2 values obtained are slightly higher than method 1 values. This means there is the consideration of higher losses. The effect on the splitting between DHW and SH is not so relevant, due to the fact that the difference between them is small. It is important to underline the fact that in case of a mean summertime period demand profile with high differences in night hours these values could be very different. But it was never found this case in the analysis and because it is considering mean values of 3 months of demand, this case is almost impossible with common type of building. There could be this problem in building with high use of hot water and in specific hours of the day, for example, of some particular factories. However, with this type of building other type of considerations must be done.

After this sub paragraph dedicated to the evaluation of DHW circuit losses it is possible to go further with the splitting. Value obtained from the previous step is reported as a red line in the plot of figure 18.

4.8.2 Last operations for DHW and SH splitting.

After the evaluation of DHW circuit losses (called E_{loss}), the equations used for obtaining DHW energy consumption are the one presented in [4.15].

$$E_{DHW} = \begin{cases} E_{TH} - E_{SH} + E_{loss}, & E_{TH} > E_{SH} \\ E_{loss}, & E_{TH} \leq E_{SH} \end{cases} \quad [4.15]$$

Where E_{DHW} is the heat load for DHW, E_{TH} is the experimental point, the global thermal energy demand, E_{SH} is the heat load for SH, obtained in the previous lines, from equation [4.14], E_{loss} is the value for losses into the DHW circuit.
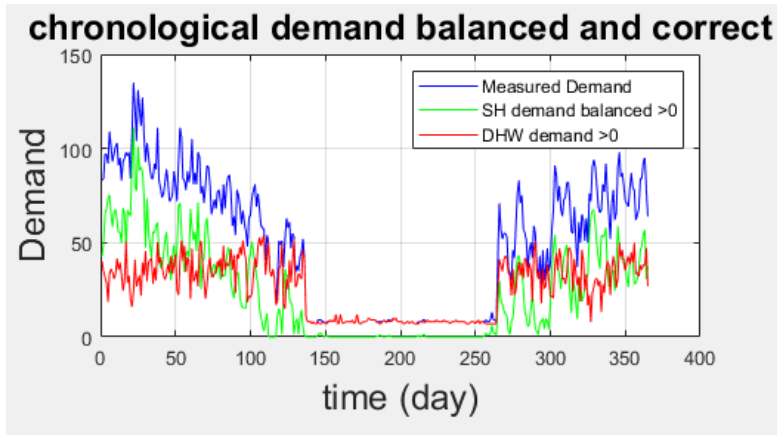
Figure 19 An example of plot where global demand is splitted into DHW and SH consumption.

After that there is the necessary operation of balancing E_{DHW} and E_{SH}. The balance of SH heat usage was adjusted based on the DHW heat usage model. The SH heat usage model was recalculated by subtracting the DHW heat usage obtained from equation [4.15] from the measured total heat usage (E_{TH}).
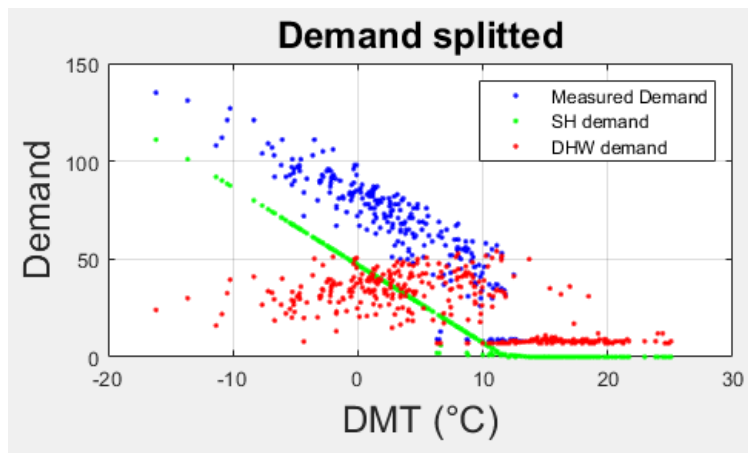
Figure 20 An example of global demand splitted into SH and DHW consumption. Plot is energy demand versus DMT.

To ensure that both DHW and SH heat usage values were positive, a condition was introduced. If one of the parameters (DHW or SH heat usage) became negative, the negative value was compensated for by the other parameter. The parameter that was negative, after this operation becomes zero. For instance, if the modelled DHW heat usage was negative for a particular point, it was compensated for by SH heat usage, and vice versa. This approach ensured that all restored DHW and SH heat usage values are positive, and their sum is balanced to equal the total heat usage.

While the proposed method could provide a reasonable estimate of the trend in SH heat usage, it has limitations. ESC relies on linear functions, which means that it cannot accurately capture sudden spikes or rapid fluctuations in SH heat usage. Additionally, due

to the fact that DHW heat load consumption is obtained from subtraction between experimental demand and model for SH, it suffers also from inaccuracies of SH's model.

Results at this point of analysis are presented in figures 19 and 20.

4.9 Detailed DHW consumption analysis

As explained is the literature review part about this topic, a detailed DHW consumption analysis is performed not only for the interesting results that this tool gives, but also to improve the model for DHW and SH splitting. In fact, one of the problems emerged after the operation of splitting is the higher number and intensity of oscillation in DHW demand in winter period. Due to the causes of this type of consumption, i.e., users' behaviour first of all, the demand pattern should be similar in summer and in winter season. An example of this fact can be found in figure 19. If red line is observed, it is clear that in winter variation among different days is much higher.

To improve the model, the necessity is to find a way to reduce these oscillations. The idea is to perform a detailed analysis on DHW profile consumption in summer. The hypothesis at the basis is that in summer there is only DHW consumption. This is a reasonable consideration, because in summertime period (from June to August) we are in the temperature independent part of the energy signature curve, so it means no space heating plant is in operation.

This type of analysis on DHW consumption in summer is useful for many purposes. First, it offers a useful method to find similarities between different days. In this way it is possible to group similar days into a single profile and treat them as they were the same day. Second, after this operation it offers a method to find hour of the day where there is minimum, medium, or peak consumption. Using these values, then, it is possible to build a standard profile for DHW consumption in summer for each group previously identified.

Thanks to this standard profile in summertime period, It was possible to translate this profile into winter period, by upscaling it by a factor obtained comparing mean value of DHW in summer and in winter. Thanks to this profile there is a standard value for DHW consumption hour per hour in winter period. By comparing this value with the one obtained from the previous operation of splitting it is possible to filter values that present a high difference with standard profile's values. With this method, in simple terms, it is possible to improve the previous model of splitting and offer the possibility to correct values for DHW consumption and, as a consequence, of SH consumption for a better adaptation to experimental values.

What have been presented in this introduction part is just the general passages of the method. For finding similar days and for identifying hours of minimum, medium and peak consumption it was followed the research work conducted by Ivanko D. et al. They did the same DHW hourly analysis, but the application was restricted to three nursing homes in Norway. It was applied in many buildings, from different datasets and case studies (also with different intended use) and then it was developed a method to use this information for improving DHW and SH splitting, as described before. The implementation of the general this part of analysis can be found in the script named "DHW_model_10922_v3" in the appendix. In the following sub paragraphs, it will be explained in detail every step and all the mathematical law it was used.

4.9.1 Step 1: Similarity in DHW heat use profiles in different days of the week

The goal of this first step is to identify with an analytical approach which days can be considered similar in terms of DHW heat load consumption. In previous analysis it was subdivided days of the week in weekend and weekdays or, as happened in paragraph 4.4, subdividing weekend also into Saturdays and Sundays. This was done because in general there is a different in consumption between weekdays, where it is more likely that occupants are away from home (or inside the building in case it is an office) and weekends, when it is, on the contrary, more likely to find people at home (or outside the building in case it is an office or a commercial activity).

However, this subdivision does not consider the intended use of the specific building in analysis. With this method it is possible to identify similar days in terms of DHW profile of consumption, for the single specific building in analysis. It is solved the problem of a general subdivision, based on probabilistic considerations, which could be not true when we talk about a single building.

The starting point consists in considering summertime global heat demand. As discussed before, in this part of the year there is only DHW consumption. This method requires profile of DHW consumption for working. It was decided to not use DHW profiles obtained from splitting because the scope of this method is to improve their accuracy and to filter some unusual (too high or too low) values.

After this consideration, there is an operation of selection of division of days of the week. Starting from the vector that contains experimental values for global thermal energy use, it was created 7 matrixes, one for each day of the week (Monday, Tuesday …) of the summertime period. Each column of each matric represent the corresponding day of the week of summertime period indicated by the number of columns. For example, the matrix called "DHW_saturday" is a 24x14 matrix. The column number 4 represents the 24 hourly value consumption of the Saturday of the 4th week of summertime period in 2019. It is important to note the fact that in case this script is implemented for other years, a correction on the first day of summer must be done. The period taken in exam, in fact, goes from the 1st of June to the 31st of August. In 2019 both these days are Saturday, but in 2020, for example, they are both Mondays. At the end of this operation there 7 matrixes 24xN, where N represents number of weeks in the summertime period that have the considered day. There are six 24x13 matrixes and one 24x14, the Saturday's one.

After that, it begins the operation of comparing days. This is a completely statistical operation, with all pros and cons of this type of approach. For doing that there is the implementation of the Student's t-test and Fisher criterion.

The Student's t-test was utilized to determine whether the mean values of DHW heat use from two different days of the week were equivalent. To accomplish this, experimental values contained in matrixes created before were used. The value used for the evaluation of similarity with Student's t-test is expressed in equation [4.16]:

$$T_{cal} = \frac{\bar{E}_{proof1} - \bar{E}_{proof2}}{\sqrt{\dfrac{S^2_{proof1}}{n_{proof1}} + \dfrac{S^2_{proof2}}{n_{proof2}}}} \quad [4.16]$$

Where $\bar{E}_{proof1}$ and $\bar{E}_{proof2}$ are the mean values of the first day considered and the second day considered (in fact this test is performed considering 2 days for each calculation), S_{proof1} and S_{proof2} are the standard deviation of the first and the second day considered, n_{proof1} and n_{proof2} are the number of values in the first day and the second day (it is always 24, because data sample has always all values). All parameters were calculated by using functions "mean" and "std" already present in MATLAB's database.

The calculated t-value, T_{cal}, was compared with the critical value, T_{cr}, which can be obtained from literature for different degrees of freedom and significance level, a. There are many possibilities in literature to find T_{cr} values. It was decided to use the paragraph "critical values of the Student's t distribution" from the book Exploratory Data Analysis of NIST. This comparison may result in three potential scenarios, as outlined below:

- If $T_{cal} \leq T_{cr}(n_{proof1} + n_{proof2} - 2, a = 0.05)$ then mean values of the first and the second sample are similar.
- If $T_{cal} \geq T_{cr}(n_{proof1} + n_{proof2} - 2, a = 0.01)$ mean values of the first and the second sample have relevant difference, so they are not similar.
- If $T_{cal} \geq T_{cr}(n_{proof1} + n_{proof2} - 2, a = 0.05)$ and $T_{cal} \leq T_{cr}(n_{proof1} + n_{proof2} - 2, a = 0.01)$ mean values of the first sample and the second one may be similar, but it is impossible to say with this analysis and final decision should be done by researchers.

To avoid any possibility of error it was decided to not consider the last point, keeping the focus of the analysis on the first two possibilities. This point, anyway, is explained better in the next lines.

The other method used to evaluate similarity between two different days is the Fisher's criterion. The parameter calculated is f_{cal} and its expression is reported on the equation [4.17].

$$f_{cal} = \frac{\max(S^2_{proof1}, S^2_{proof2})}{\min(S^2_{proof1}, S^2_{proof2})} \quad [4.17]$$

Where S_{proof1} and S_{proof2} are the standard deviation of the first and the second day considered, as happened before. Even in this case there is a comparison between the calculated f_{cal} and a standard value called f_{cr} to check the possible similarity. Values for f_{cr} can be found in the literature. In this analysis it was found these values on the F distribution tables in the site of SOCR (Statistics Online Computational Resource). In this case there are two possible scenarios:

- If $f_{cal} \leq f_{cr}(n_{proof1} + n_{proof2} - 2, a = 0.05)$ then the variances of the first and the second samples are similar.

- If $f_{cal} > f_{cr} (n_{proof1} + n_{proof2} - 2, a = 0.05)$,then the variances of the first and the second samples have significant difference.

		TUE	WED	THU	FRI	SAT	SUN
		1	2	3	4	5	6
MON	1	0.7327	0.6803	1.9178	2.2205	1.4061	0.0527
TUE	2	0	1.2899	2.4028	2.6523	1.9468	0.7412
WED	3	0	0	0.9867	1.3692	0.6253	0.5911
THU	4	0	0	0	0.5808	0.3176	1.6982
FRI	5	0	0	0	0	0.7931	2.0157
SAT	6	0	0	0	0	0	1.2596

Figure 21 An example of matrix with Tcal values. Name of day are reported.

The function "stud_fish" present in the appendix is only for calculating T_{cal} and f_{cal}. These parameters were calculated for every day contained in every matrix, one week per time. It was repeated the implementation of the function "stud_fish" for every week of the summertime period, from the 1st to the 13th. At the end of this operation there are 13 matrixes for T_{cal} and 13 for f_{cal}. Each matrix is 6x6. On rows there are days from Monday to Saturday, on column from Tuesday to Sunday. In this way there are no Monday-Monday or Tuesday-Tuesday… calculation because, working within the same week, it is obvious that Monday is equal to Monday, being the same day. An example of matrix for T_{cal} is reported on figure 21. There are, also, name of days for clarity of exposition.

The two profiles are deemed similar if both the Student's t-test and Fisher's exact test yield consistent outcomes. In this case they can be considered as a unique day in terms of pattern of consumption. However, if either test indicates that there are significant differences in mean values or variances between the two sample profiles, it can be inferred that the profiles are dissimilar and must be analysed independently.

As said before, the initial statistical sample, which includes data of DHW consumption during summertime period, was divided into different weeks to analyse them once per time. For each week, it is necessary to conduct a comprehensive comparison of all possible combinations of daily DHW profiles using both the Student's t-test and Fisher's exact test.

This entails comparing the profiles of each day of the week against one another, such as Monday vs. Thursday, Monday vs. Wednesday, Saturday vs. Sunday, and so on. While there is the function "stud_fish" for calculating T_{cal} and f_{cal}, the function "match" performs the comparison between values calculated and critical values (T_{cr} and f_{cr}).

Subsequently, once the statistical tests have been performed on all possible combinations of days, it is possible to determine the number of weeks in which the DHW profiles of the selected pairs of days are considered similar. In order to conduct further analysis, the percentage of matches between the DHW profiles for each combination of days of the week can be calculated using the following formula [4.18]:

$$n_{i,j} = N_{i,j} * \frac{100}{N_{total}} \quad [4.18]$$

Where $n_{i,j}$ is the number of matches in percentage, $N_{i,j}$ is the number of weeks where the day i is similar to the day j and N_{total} is the total number of weeks in the sample analysed, in this case 13. i was the day of the week of the first comparable profile (from 1 to 7). j was the day of the week of the second comparable profile (from 1 to 7). For higher clearness, an example of matrix of matches can be found in the figure 22.

	Mo.	Tu.	We.	Thu.	Fr.	Sa.	Su.
Mo.	$n_{1.1}$	—	—	—	—	—	—
Tu.	$n_{2.1}$	$n_{2.2}$	—	—	—	—	—
We.	$n_{3.1}$	$n_{3.2}$	$n_{3.3}$	—	—	—	—
Th.	$n_{4.1}$	$n_{4.2}$	$n_{4.3}$	$n_{4.4}$	—	—	—
Fr.	$n_{5.1}$	$n_{5.2}$	$n_{5.3}$	$n_{5.4}$	$n_{5.5}$	—	—
Sa.	$n_{6.1}$	$n_{6.2}$	$n_{6.3}$	$n_{6.4}$	$n_{6.5}$	$n_{6.6}$	—
Su.	$n_{7.1}$	$n_{7.2}$	$n_{7.3}$	$n_{7.7}$	$n_{7.5}$	$n_{7.6}$	$n_{7.7}$

Figure 22 The form of matrix of matches

Being percentage, all numbers in this matrix are between 0 and 100. On the diagonal there are, of course, all 100 because for every week, within the week in analysis, Monday is equal to Monday. This means that there are 13 weeks where Monday is equal to Monday, and by applying equation 4.18 can be found that the value for $n_{1,1}$ is always 100. For making another example we can consider $n_{5,2}$. This number give us the number of weeks where the Friday is similar to Tuesday. If there are 9 weeks where both T_{cal} and f_{cal} for Friday-Tuesday pass their respectively test, here we can find that $n_{5,2} = 69.23\ \%$.

Using the matching matrix, we can identify groups of weekdays with comparable patterns of DHW heat usage. Specifically, weekdays with $n_{i,j} \geq 100 - error$ should be considered similar in terms of their DHW heat usage and analyzed together as a group. The error value takes into account the precision of statistical tests such as Student's t-test, Fisher's exact test, as well as the percentage of days in a year when the building is not operational, such as holidays. To estimate the error, the accuracy of the Student's t-test, Fisher's exact test, and the percentage of days in a summer when the building's operation was atypical (such as holidays) were taken into account.

The precision of Student's t-test and Fisher's exact test is considered to be 5%. Furthermore, considering the number of atypical days in DHW heat usage, the acceptable error value was estimated to be 14% in the research of Ivanko D. et al. However, their analysis considered an entire year, while in this case there is the consideration of only summertime period. For the lower number of weeks considered and for the increasing probability of having users on holiday, creating atypical profiles in DHW consumption, it was increased the value of error. By checking various analysis on different buildings, it was assumed that a reasonable value for error is 25%. This means that, using this criterion, it is possible to identify the days of the week in the building with statistically similar patterns in over 75% of the analysed weeks.

By analysing the matrix of matches it is possible to subdivide days into groups. Within the group, days can be considered similar, so it is reasonable to assume that they have the same pattern. This operation must be done by analysing match matrix by hand. The small

number of weeks considered may create some doubt on the creation of group. For this reason, a check by hand is necessary.

It is important to say that if there is a doubt, but the value $n_{i,j}$ is close to the limit imposed by error, it is possible to insert this day in the group where there are a major number of days similar to it. To keep the analysis simple, it was decided to create maximum two groups per each building, but it can be implemented the possibility to have even more or less groups. For me, two is a good number for the purpose of this analysis.
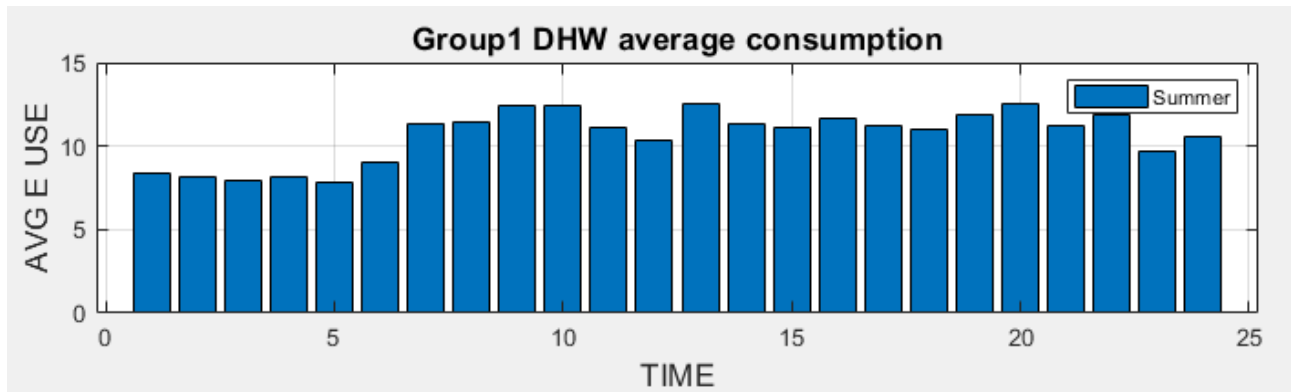

Figure 23 An example of an average profile of the group 1, obtained after checking similarity between different days.

At this point, matrixes that contain samples of different days of the week, the ones mentioned in the introduction of this paragraph, named "DHW_monday" etc… are inserted in group one or group two according to the considerations done from the previous analysis. This first step of detailed analysis is concluded by calculating mean profile for every group. Mean profile is obtained calculating mean values hour per hour between all samples in the corresponding group. At the end, there is the creation of a 24-values mean profile for both group 1 and group 2. The graphical representation of these average profiles is in the example of the figure 23.

4.9.2 Step 2: Determination of zones with minimum, medium, and peak consumption.

Identifying hours with minimum maximum and peak consumption is useful not only as a starting point for the operation of flattening of DHW profile in winter, but also for energy management in buildings. To effectively manage energy usage in buildings, in fact, it is crucial to identify the time zones with peak load, minimum load, and average heat load during the day. This operation could simplify the management of heating plants or could provide essential information to the manager of the district heating network, allowing him, for example, to regulate better heat production.

To address this challenge, we propose a statistical grouping approach for the hourly heat use of the DHW system, based on the method presented by Nakhodov et al., originally used for identifying tariff zones in the power system. This method was, then, improved and its applicability for the minimum, medium and peak hours identification is proven in the article by Dmytro I. et al. They have adapted this method for analysing DHW heat use in buildings, enabling them to categorize hours of DHW heat use into multiple groups, each with a statistically distinct mean value.

The process involves an iterative procedure and analysis of mean values using the already mentioned Student's t-test. In this case, the DHW heat use profile is considered a statistical sample, containing 24 elements (hours) with DHW heat use values e_j in each hour (where j represents the element number in the sample). This sample is the average consumption profile for group 1 or group 2 (the analysis is conducted for both groups) built in the previous paragraph. The detailed algorithm for this method is expressed in the next lines. The analysis is conducted for one group per time.

Step 2.1: Order elements in data sample

The initial dataset contains element of average DHW consumption during summertime period for the considered group. This means there are twenty-four values, but they are in chronological order. The first operation consists in sorting this vector form the smallest value to the biggest one. The elements in the vector are now called E_i where $E_{i+1} > E_i$ due to the effect of sorting operation. For doing this step it was used the function "sort," present in MATLAB's database.

Step 2.2: T_{cal} calculation

Now the starting point is the ordered vector E. In the code it is possible to find it with the name "DHW_mean_group1_E" if we are considering group 1's analysis. Starting from it, step 2.2 consists in determining T_{cal} values trough a sort of iterative procedure. This iterative procedure is used to generate two statistical subsamples, R_1 and R_2, with varying numbers of elements, based on the sample E.

At each iteration, R_1 contained M elements while R_2 contained M+1 elements, with the elements in both subsamples taken consistently from the initial sample E. There is an increase in the number of elements in R_1 and R_2 by one with each iteration, with M ranging from 1 to 23. After that, the value of the Student's t-test for R_1 and R_2 at each iteration is calculated, using the equation [4.16]. In this case, of course, subscripts "1" and "2" reported in the elements of the equation do not represent two different days, but the vector R1 and R2 for each iteration. At the end, there is the presence of a 23-elements vector for both group 1 and group 2, which contains T_{cal} numbers.

As an example, to explain better this step:

- In iteration 1: $R_1 = [E_1]$; $R_2 = [E_1\ E_2]$; M= 1; T_{cal1};
- In iteration 2: $R_1 = [E_1\ E_2]$; $R_2 = [E_1\ E_2\ E_3]$; M= 2; T_{cal2};
- In iteration 23: $R_1 = [E_1\ E_2\ \dots\ E_{23}]$; $R_2 = [E_1\ E_2\ \dots\ E_{24}]$; M= 23; T_{cal23};

The code for the implementation of all passages necessary to do this operation can be found in the function "time_interval."

Step 2.3: Eventual merging of close hours using Student's t-test.

At this point, the result of the previous operations is a vector that contains all T_{cal} values. There are twenty-three numbers. According to the discussion conducted by Dmytro I. et al., if a sample of hourly DHW heat use is ordered and monotonous, the numerical values of its elements increase evenly. This is the case of this analysis: in step 2.1, the initial sample was sorted by ascending order. If this happens, the t-criteria series obtained by the iterative procedure will also be monotonous. This means that the values of t-criteria obtained by

equation [4.16] will decrease monotonically with each subsequent iteration ($T_{cal1} > T_{cal2} > ... > T_{calM}$).

However, if the ordered sample of hourly DHW heat use is uneven, the monotonic decrease of the calculated values of the t-criteria will be interrupted by periodic sudden growth (it happens that $T_{cali} < T_{cali+1}$). Continuing to use information from the analysis of Dmytro et al., identifying the points of growth of the calculated values of the t-criteria can help determine between which hours there is a noticeable statistical difference in DHW heat use. This assumption allows me to divide the hours in the profile of DHW heat use into several groups, with each group being a sample of data where DHW heat use data varies monotonously. To avoid any possibility of misunderstanding, from this moment it will be called these groups of monotonously data variation "set." In this way there is no confusion with group 1 and group 2, which refer to similar days put together.

The neighbouring sets of hourly DHW heat use can be checked for the possibility of further merging. For this purpose, the data samples of two neighbouring sets are checked using Student's t-test, another time (refers to the equation [4.16]). As a result, the calculated value of the t-criteria, T_{cal}, can be compared with the critical value, T_{cr}. The possible scenarios are reported in the sub-paragraph 4.9.1 of the current chapter. As happened before, even in this case it was considered only the first two possibilities. If the resulted T_{cal} belongs to the last case, it was considered the merging not possible.

It is important to underline the fact that Tcal analysed in this phase are different to the one contained into Tcal 23-elements vector obtained in the previous step. They are used only to identify the corresponding index M where there is a sudden growth, so it means in correspondence of this M there is a possible difference between data before and after this index. Values calculated in this step refer to other data. In particular, when it was applied the equation [4.16] to evaluate Tcal in this step, subscripts "1" and "2" refer to all E values before the M and after the M (until the next growth), where M is the index when the abrupt growth happens.

As an example, we can analyse the plot contained in figure 24. Here there is the representation of the 23-elements Tcal vector obtained from the iterative procedure highlighted in step 2.2. It can be seen that there is a monotonous decrease from the first value up to value number 3. This means that M=3 is the first index of abrupt growth. At this point, it was performed Student's t-test using equation [4.16] but as sample 1 there are all corresponding values of demand from E1 to E3 and for the sample 2 there are value of demand from E3 to E5 (E values are the ones contained into the sorted vector of average demand of the corresponding group).
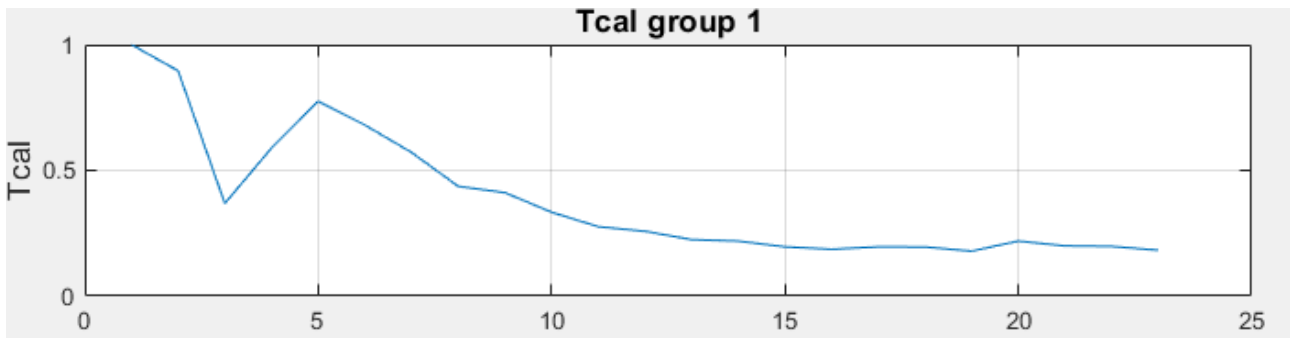
Figure 24 An example of Tcal plot versus index of iteration M.

After this splitting process, there is the creation of a matrix which contains in each column values of average demand of the corresponding group sorted E before the index of period growth M. Considering the previous example, in the first column there is E values from the 1st to the 3rd, in the second column from the 3rd to the 5th, in the third column from the 6th to the 16th and so on.

For the merging of the two sets of experimental DHW consumption data it was used the comparison between Tcal and Tcr. In case of positive result from the Student's t-test, the merging of two sets is possible. For the merging, there is the creation of a unique set. If both initial sets have a value in the considered index, the mean values are taken, on the contrary, only the value of the set that contains a value for the index considered is taken for the merged set. For example, if there is a possible merging between set 1 which has 5 values and set 2 which has 2 values, the merged set contains 5 values. The first and the second are mean values between set1(1), set2(1) and set1(2), set2(2) and the other 3 values are taken from the set 1 alone.

Once there is a merging in the sets based on the conditions explained above, a new ensemble of sets is created. To account for these changes, it was repeated the calculations of step 2.3 from the beginning with the new ensemble of sets in the sample. It was continued to iteratively calculate step 2.3 until the t-test indicated that no additional sets could be merged, and the total number of sets could not be reduced any further.

The operation of splitting mean values contained in E vectors according to the values of T_{cal} obtained from step 2.2 is performed with the function "Tcal_splitting." The merging process explained in the step 2.3 is done with function "iterative_merging_process."

It is important to underline the fact that it was decided to also insert a control to account the difference between two matrixes created in two following iterations. In fact, even if Student's t-test suggests a merging, if the relative error between matrixes that come from two following iteration is small, iteration process can be stopped. It was decided to consider a tolerance of 2%. This is useful to avoid a too long iteration. After this tolerance value, in fact, there is no more benefits in continuing the merging operation, it is only a time consumption.

Step 2.4 Identification of critical values, the boundaries of minimum, medium and peak zones.

At this point, the result is a matrix, where in each column there is experimental average consumption of DHW in summertime period of the considered group analysed, where there is no possibility of merging close columns. Thanks to the sorting operation done in step 2.1,

the first column represents the lowest values, and the last column represents the highest values of demand that occur in an average day in summertime. So, by using these columns is possible to obtain values that represent boundaries between minimum and medium load and between medium and peak load. The equation for determining them (E_{min} and E_{max}) are reported in [4.19] and [4.20].

$$E_{min} = \bar{E}_{col,1} + T_{cr,1}(M_{col1} + 1 - 2, a = 0.01)\sqrt{\frac{S^2_{col,1}}{M_{col,1}}} \quad [4.19]$$

$$E_{max} = \bar{E}_{col,k-1} + T_{cr,k-1}(M_{col,k-1} + 1 - 2, a = 0.01)\sqrt{\frac{S^2_{col,k-1}}{M_{col,k-1}}} \quad [4.20]$$

In the equations, $\bar{E}_{col,1}$ and $\bar{E}_{col,k-1}$ represents mean values of column 1 and the second-last column of matrix after merging operation, $T_{cr,1}$ and $T_{cr,k-1}$ are the critical values for Student's t-test of the column 1 and the second last one, with degrees of freedom and significance level, a, indicated in brackets, $S_{col,1}$ and $S_{col,k-1}$ are standard deviations of the first and he second last column and $M_{col,1}$ and $M_{col,k-1}$ represent the number of elements.
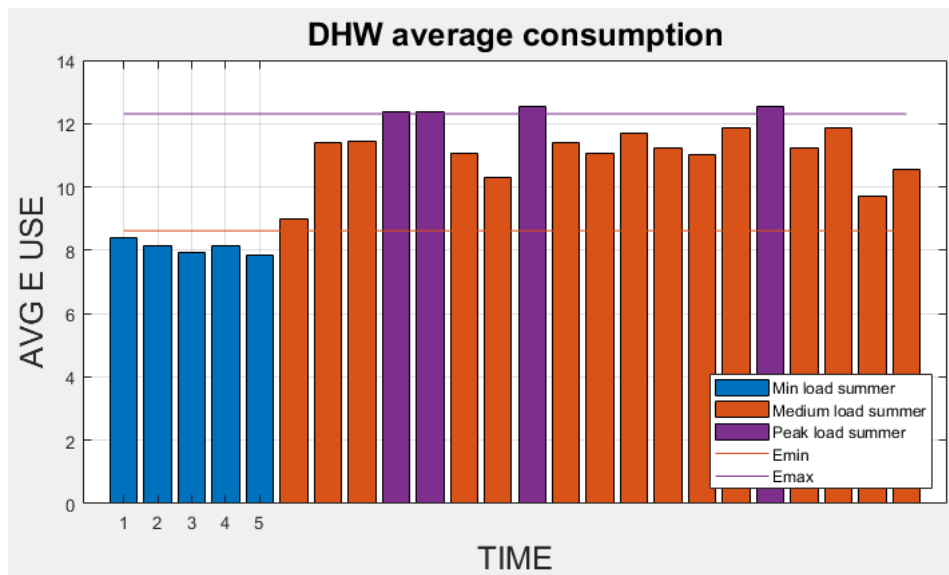

Figure 25 An example of average DHW consumption profile in summer, with subdivision in hours with minimum, medium, and peak demand.

The time periods during which the DHW heat use was below the minimum threshold E_{min} should be considered as a zone with the lowest DHW heat use. If the DHW heat use was between E_{min} and E_{max}, it could be assumed that during these periods, the DHW heat use was in a zone of average heat use. The time periods with DHW heat use higher than E_{max} should be considered within the zone of the maximum heat use.

The last phase of this part of the analysis consists in representing the mean DHW consumption profile for both group 1 and group 2 found before, an example can be found in figure 23, but with the subdivision into minimum, medium and peak consumption hours using E_{min} and E_{max} values. An example of this plot is reported in figure 25.

From many analyses done by applying this model, it was found that it could happen that E_{min} and E_{max} values are too low or too. In particular, if E_{min} value is too low it could be that no

hours have consumption lower than it. On the contrary, if E_{max} is too high, no hours have a demand higher than that, resulting that there are no peak hours, even if by observing the pattern it is clear that they are present. it was observed that it happened especially if experimental demand is low. This error could be produced by inaccuracies in the model, especially in the merging operation process and in the Student's t-test. There is, also, a dependence on the number of values in the first and in the second last column after merging process. There is no control on that numbers, because they depend on merging process and on the values in the initial dataset.

For this reason, it was decided to implement a correction: if no values are below E_{min}, E_{min} value is increased by 7%. On the other hand, if no hourly mean demand values are above E_{max}, E_{max} value is decreased by 7%. The choice of 7% is linked to analyses done in about 30 buildings, seeing that in some cases it was necessary this small correction for an accurate representation of the hourly profiles.

All passages reported in this last step are implemented in MATLAB using function "min_med_peak_v2" that can be found among other codes that it was used.

4.9.3 Step 3: creation of the summer standard profile.

The idea at this point is to create a standard demand profile using information about E_{min} and E_{max} values and by knowing which hours present minimum consumption, which ones have medium consumption and which one present peak demand. This information is
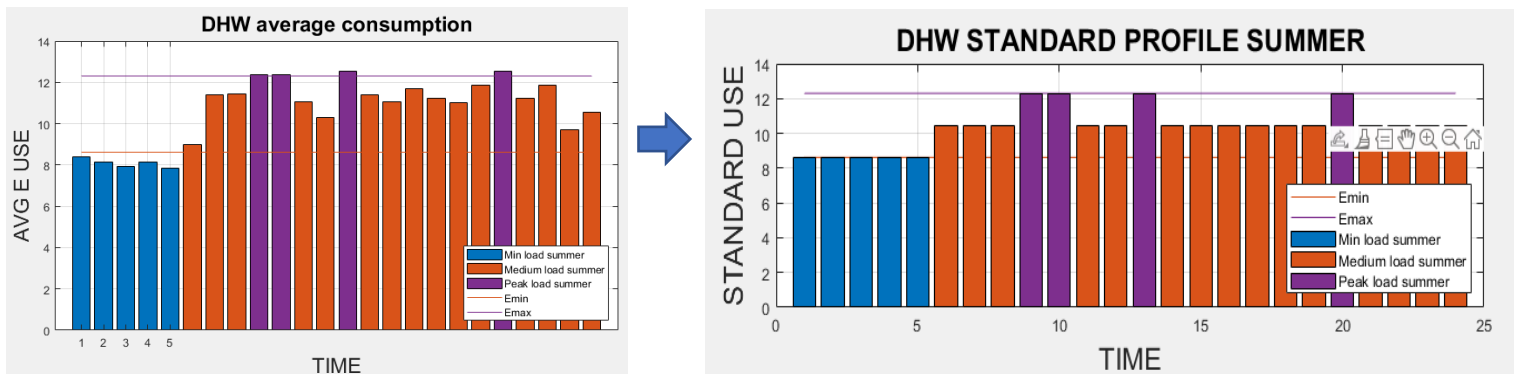
Figure 26 Passage from DHW average consumption profile to DHW standard profile. Both refers top summer period.

obtained from the previous analysis. Using this knowledge, it was built a standard profile of DHW consumption in summer for both group 1 and group 2. So, at the moment we know which days of the week are similar in terms of DHW consumption profile and we know we can consider them as a unique day. We know, also, in which hours there is minimum, medium, or peak consumption.

For the creation of the standard profile the following procedure has been followed.

- For hours of minimum consumption, namely they have a demand lower than E_{min}, in the standard profile the demand in these hours is fixed equal to E_{min}.
- For hours of medium consumption, namely they have a demand between E_{min} and E_{max}, in the standard profile the demand in these hours is fixed equal to $(E_{min} + E_{max})/2$.

- For hours of peak consumption, namely they have a demand higher than E_{max}, in the standard profile the demand in these hours is fixed equal to E_{max}.

The passage from average DHW consumption in summertime period to standard summertime profile is represented in figure 26. This operation was done to simplify the analysis when there is the passage to the winter standard profile and have values as standard as possible. There is no interest in specific value hour per hour, but it wants to know if the considered hour presents minimum, medium, or peak consumption.

4.9.4 Step 4: creation of the winter standard profile

This step consists in using the standard profile in summertime period to create the same thing but in winter. As emerged in many studies (work of Dmytro I. et al., Lomet A. et al., Gerin O. et al. are just few examples mentioned before) and from analyses done, there is a difference between DHW consumption in summer and in winter. Even if it was assumed that this type of load is temperature independent, because of lower temperature of inlet cold fresh water during winter season. However, the cold inlet water temperature has a slow response to the outdoor temperature. Further, this effect leads to seasonal variation of the DHW heat use in the buildings.
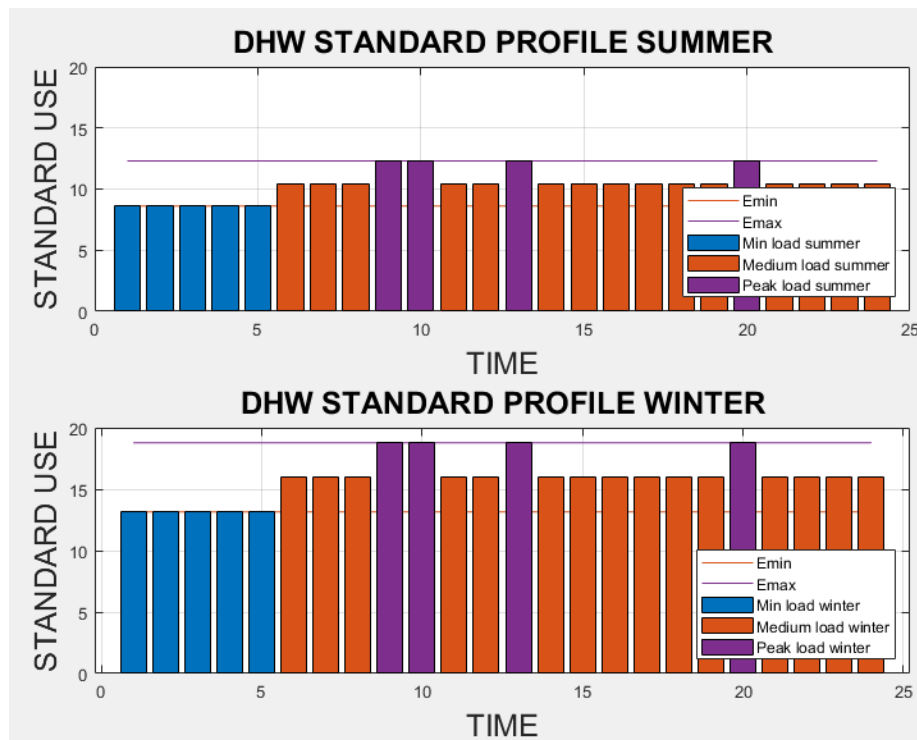
Figure 27 Difference between summer standard profile and winter standard profile. It is evident the fact that all values have just been upscaled by a factor.

To adapt standard summer profile to winter consumption and create, in this way, a good reference for DHW demand in this period, it was decided to upscale summer consumption profile by a certain factor. This factor is the difference in % between mean values of DHW consumption during summer and during winter. Data for DHW consumption in winter are taken trough the operation of splitting between DHW and SH heat load demand described before. By considering mean values of consumption in a long time period, it was able to be

independent from hourly variation of the demand. In this was it was possible to exclude, for example, too high values caused by sudden spikes of DHW consumption.

 All the necessary operations can be found in the corresponding part of the general script "DHW_model_10922_v3". At the end, it is possible to plot the winter profile. An example of this result can be found in figure 27. During this phase it was had, also, separate winter consumption from the rest of the year and demand was subdivided into days of the week. In this way it is possible to apply winter standard profile of group 1 or group 2 according to the membership of the considered day to one group or the other. For example, by keep dived Monday, Tuesday, etc. it is possible to use standard profile of group 1 on Monday if Monday belongs to group 1 and standard profile of group 2 on Sunday, if Sunday belongs to group 2. This operation is necessary for the general applicability of the script. In fact, the only thing that must be changed is which days belong to group 1 and which to the group 2.

4.9.5 Step 5: Flattening DHW profile

This is the last step of detailed DHW consumption analysis. In this phase it was compared winter DHW demand obtained from splitting operation with standard profile in winter built in the previous sub-paragraph. This step is composed by different operations.

First of all, it was calculated the difference between every hourly experimental point in winter period (from January to May and from September to December) and the standard profile of this season. This operation is done for every day. The operations were done using matrixes that have 37 columns of 24 values per each column. 37 because every column corresponds to the considered day in winter season. For example, in the matrix called "E_DHW_Y_FRI_win" there are the experimental hourly values of 37 Fridays of winter season. In particular, there are 21 weeks from January 1st to May 31st and 16 weeks from September 6th to December 31st. This fact is true if year considered is 2019, if there is a different year in analysis, it must be changed the first day of the year in the script implementation.

After percentage difference evaluations, it was found that the maximum difference between standard values and real values is in the range of 70%. It is significant to underline the fact that this is the maximum difference. The mean difference between DHW splitted demand points in winter period and standard demand profile in winter season goes between 0 and 15%, as it was found in the analysis of various buildings. This highlight the fact that the standard profile for winter is a good approximation for average consumption and it can be used for flattening demand in points where hourly consumption is much different from the standard consumption.

From these considerations, it was calculated a new profile for DHW consumption in winter. It checked one by one the difference between the DHW hourly consumption obtained from splitting operation in analysis and the standard hourly demand for the same hour of the same day. It was decided a percentage, in the analysis it was fixed 30%. All values that present a difference in percentage higher that 30% are subjected to correction. It is important to emphasize the fact that 30% is taken considering that the maximum difference is around 70%, the mean around 10% and it was taken an average between them. This percentage can be easily changed. In the chapter dedicated to future work (chapter 8) It will be explained

better how this tolerance percentage can be determined in a more precise way. For now, with available data and information, the only possibility is to assume a value based on the experience of the researcher.

The correction applied to these values consists in a reduction by a percentage of the value. Even this percentage must be chosen by experience of the researcher. In future work section it will be explained that both "tolerance" percentage and this "decrease" percentage could be implemented in an iteration loop to find the best values for describing better the profile of the building. However, this operation requires at least a period where demand is measured already splitted, between DHW and SH. As said before, it will be focused the attention on this point in chapter 8. At the moment, it was decided to fix this correction percentage, called "decrease," at 25%.

After this operation of correction, it was rebuilt the whole hear DHW consumption, assuming for winter demand corrected values and for summer demand the experimental demand (in fact, as demonstrated before, in summer there is only DHW consumption).

The last operation consists in balancing SH demand. In fact, by changing DHW values also SH of the correspondent hour must be changed, to preserve the global demand equal to the one measured from meter. It was calculated the value subtracted to the DHW point with a percentage difference higher than tolerance, then, after obtaining new correct value for DHW, it was added the difference value to SH demand. In this way the value subtracted from DHW goes to SH and the global heat consumption hour per hour is preserved and is still equal to experimental value.
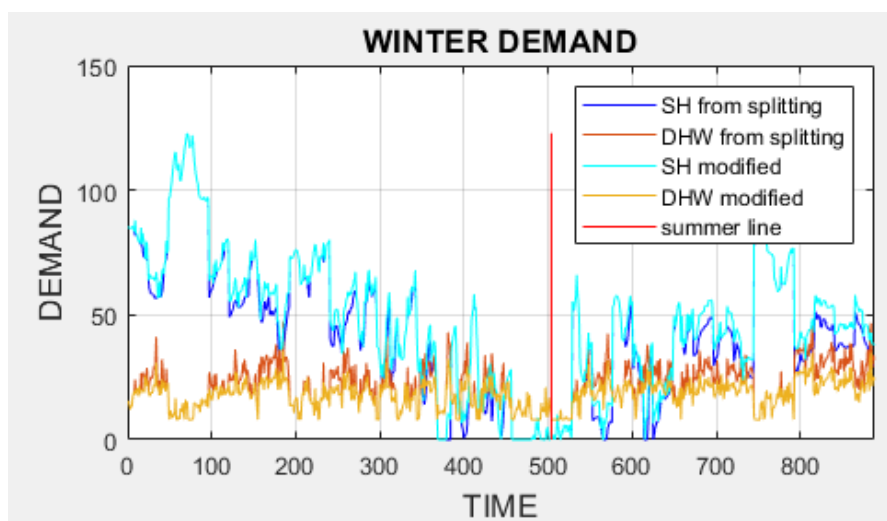
Figure 28 An example of DHW and SH demand before and after flattening operation. In this figure only Fridays of the year are represented. Here there is only winter season. Red line is the boundary between May 31st and September 6th.

After that, it was plotted new values for SH and DHW of the year in chronological order, together with values obtained from splitting operation. In this way, a comparison between them is possible. It was plotted every day of the week distinguished one from the other and then the entire year. It is possible to see in figure 28 and 29 that this operation is a success.

With this operation the detailed analysis of DHW heat consumption is conclude and also the entire analysis of a building with complete database. Tools presented here gave information about heat consumption of the building, in particular SH and DHW heat load, patterns of typical use, correction possibilities, clues about intended use, atypical days consumption, and more. Comments, considerations, and possibility to use these data are explained in chapter 7, where it will be presented some results of the case studies analysed. The next chapter will talk about how a similar analysis can be performed in case of there is a partial database of the building's heat demand.
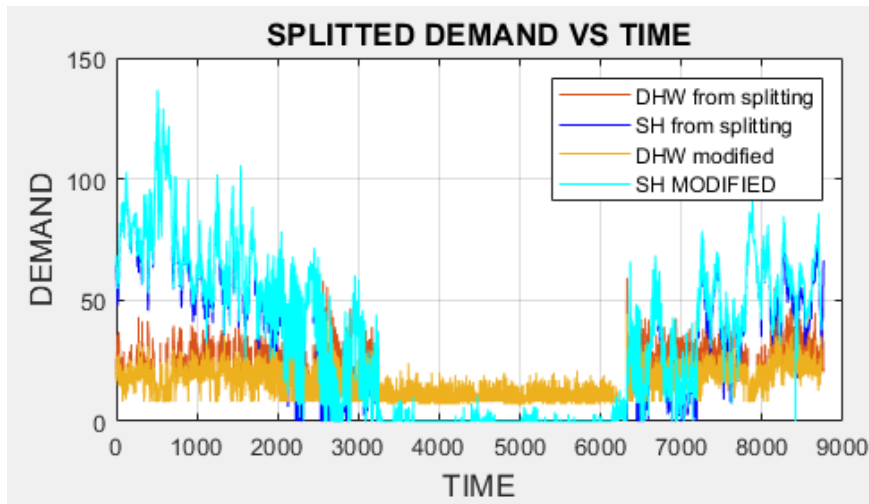

Figure 29 An example of DHW and SH demand of the entire year. Here there is the representation of values before and after correction operation. It can be noticed that after correction DHW demand profile is more regular, as it would be.

Chapter 5: Description of method with partial database.

After the analysis conducted on a complete year database, it is wanted to apply similar tools also in case database is not completed. There is, in fact, the possibility to analyse even partial databases with reliable results. In practice, it is much easier to measure consumption for a limited period than an entire year. It common, in addition to that, finding hour where meter did not work for any reason, maybe it was an error or maybe it was simply turned off.

As it will be shown in this chapter, it is possible to analyse these cases by reconstructing energy signature curve and it is possible to split DHW and SH heat load consumption. However, as it will be shown in the next paragraphs, it is impossible to perform the detailed DHW analysis in summertime period because in the analysed database there were no data about summer consumption. It is important to underline that this is a common case in practice. In fact, demand is usually measured during wintertime period to obtain information about consumption. Results obtained from this analysis are consistent with the ones from the previous analysis.

This analysis was applied to a different dataset. The case study related to this partial database discussion is the district heating network of the historic centre of Verona. This means climate are completely different from Tartu, with different rules and mode of operation. This fact becomes relevant especially when it was evaluated DHW circuit losses, in the sub-paragraph 5.7.1 of the current chapter.

It begins the illustration of the workflow followed from the data pre-processing operation, as done in the previous chapter. All figures and examples made in this chapter, useful to explain better the method, are taken from the analysis of condominium number 10 of the Verona's dataset.

5.1 Data pre-processing.

Pre-processing operation is even more important when there is an analysis of partial dataset. Due to the smaller number of experimental points available, in fact, their quality becomes more important. For this reason, particular attention must be done in this phase, above all when there is a partial database. With partial database it was meant that not an entire year of experimental measurements is available. In the case analysed in this site, the Verona's database, only winter season was available. In most building period of measurements goes from end of November to the end of February.

In addition to that, in this database there were a lot of missing values, probably due to meter's fault or because it was turned off. By not doing an appropriate pre-processing, a lot of zeros would be present in the experimental global demand values, creating big inaccuracies. They are, in fact, not a representation of the building's behaviour, but only caused by measurements error. In this paragraph it will be shown the operations done in the initial dataset for improving quality of data.

First of all, it was controlled that the database begins with a complete day. There were cases where first values were, for example, 13 hours of a day. This is a problem, because it makes impossible an accurate hourly division, as performed in paragraph 4.2 of the previous chapter and as it is described in paragraph 5.2 of the current chapter. As it will be explained

later, in building located in mild climate, as it happens for Verona, hourly division in an essential operation for understanding building's behaviour. In case there was an incomplete day as first values of the dataset, it was decided to not consider it and delete it from the experimental values.

The next operation consists in calculating hourly thermal consumption. it was used the same equation implemented for analysis of complete database, which can be found in [4.1]. In case there were more than one values of the interested data per hour, it was decided to consider always mean values. In Verona's database, for example, there are 3 values per hour for each variable. The first one represents the maximum in the hour, the second one the minimum and the last one the average. This happens for volume of water used, for power, inlet temperature of water and return temperature of water. In figure 30 there is an example of database, in red the values it was considered are highlighted.

time	flow_min	flow_max	flow_avg	power_mi	power_max	power_avg	flowTemperature_min	flowTemperature_max	flowTemperature_avg
mercoledì 27 novembre 2019	0.047333333	0.04733	0.04667	1.20867	1.208666667	1.21	44.03333282	44.03333282	44.05333333
mercoledì 27 novembre 2019	0.056	0.06775	0.0625	2.36225	2.805	2.5825	56.34999847	57.92499924	57.1375
mercoledì 27 novembre 2019	0.045	0.045	0.045	2.2515	2.2515	2.25	63.45000076	63.45000076	63.45
mercoledì 27 novembre 2019	0.042	0.07633	0.06	2.079	3.884333333	2.983333333	63.63333511	65.09999975	64.36666667
mercoledì 27 novembre 2019	0.933	1.238	1.09	31.391	46.121	38.76	73.19999695	73.80000305	73.5
mercoledì 27 novembre 2019	0.55275	0.57175	0.5625	16.8	17.749	17.275	72.52499962	73.10000038	72.8125
mercoledì 27 novembre 2019	0.543	0.61133	0.57667	18.4203	20.99633333	19.70666667	75.06666565	75.16666412	75.11666667
mercoledì 27 novembre 2019	0.011	0.011	0.01	0.348	0.348	0.35	64.30000114	64.30000114	64.3

Figure 30 An example of Verona's dataset. In red average values are highlighted. Only few variables are reported here, not the complete set.

Another operation of pre-processing consists in individuating the period of measurements. It is typical that measures are taken during a singular winter season. This means that 2 different years are considered: from November to December and from January to February, in this case. This is important when temperature dataset is considered. To draw plot of demand versus outdoor temperature, in fact, data about outdoor temperatures are necessary. In case of partial database with a heating season between two different year, two years of dataset for outdoor temperature are needed. Lucky, these types of data are usually quite simple to find.

The next step consists in empty values remotion. This operation is performed in MATLAB, where it was implemented a function for the remotion of all empty values from the vector that contains data about hourly energy consumption. It uses the fact that when an empty box of a table is imported in MATLB, it is read as a "NaN" value. So, it was done a NaN remotion from the energy vector.

It was decided to follow two different approaches in this remotion. The first one consists in removing only NaN values. In this way there is the creation of a vector with only real experimental data about hourly consumption. The vector with hourly consumption data pre-processed in this way is called "ENERGY_N_NaN." The problem is that in this way it is impossible to split data on hourly base. Even if this is still possible, there would be big differences between hours, where some hours may have vast number of values and other a smaller number, by creating some problems in accuracies of the model.

For this reason, it was decided to also implement another method. In this case there is a remotion of all days that contains at least one NaN value. The vector with hourly consumption data pre-processed in this second way is called "ENERGY_D_NaN." By applying this method there are less experimental day available, but there is the certainty that all days included in the dataset are complete and without errors in measurement operation. For the entire analysis it will be used as a starting point, data that come out after this second

method of pre-processing. Data from the first approach are still available and they were used mostly to control if the considerations done with dataset from the second approach are consistent with a larger number of experimental values (from the first approach).

With these operations, the phase of data pre-processing is over. At the end of this step, vectors which contain hourly heat consumption pre-processed and a vector that has hourly outdoor temperature are available. From this moment, the data analysis begins.

5.2 Weekdays and Weekend splitting.

This is a special step, necessary in some buildings and optional or useless in others. From many analyses and previously considerations described on the previous chapter, it is clear that in some buildings there could be a significant difference in terms of heat consumption between workdays and weekend. So, for buildings used as offices or school, quite different profiles of consumption occur during workdays or weekends. By considering all together, there is the creation of big inaccuracies. During weekends in this type of buildings, in fact, no consumption is present, so demand is zero or almost zero. So, subdividing workdays and weekends it is essential for obtaining a good model of the real building's behaviour.

For all these reasons, it was decided to implement a function that can split days between weekdays and weekends. It is called "NaN_removal_v2" and can be found among other functions and codes used in the analysis. As input, it uses the vectors with hourly energy values, a vector of one variables of the dataset to use for controlling when there is a NaN number, and a number which represents the first day of the dataset (Monday=1, Tuesday=2…).

It is easy to decide when subdivide workdays and weekends and when, on the other hand, consider all days together because in Verona's dataset there is the information about the intended use of the building in analysis. In case this information is absent, the choice can be done by plotting experimental demand versus time and seeing if there are periodical decreases in demand, clue of a weekend lower demand. In addition to that, it emerged from analyses conducted on the buildings that errors in measures are common during weekends, maybe because the building is closed, or the plants are turned off.

5.3 Hourly data splitting.

The operation done in this phase is similar to the one done for the complete database, explained in paragraph 4.2. For this reason, it will not be explained in detail all passages because it would be a repetition of what already said in the mentioned paragraph. It was, also, used the same function, the one called "HOUR_split."

The only difference is that this time it was performed the hourly data division not only on global heat load demand, but also on outdoor temperature values. It was also calculated vector for daily mean temperature. It was decided to represent points of experimental demand versus daily mean temperature and not hourly outdoor temperature to keep the same approach followed for the analysis with complete database. However, it was performed hourly division also on temperature values, because for following phases it is

necessary a separation between demand in various hour of the day, and together with it also of outdoor temperatures.

As said before, for the entire analysis it was used the vectors that contains only complete days of hourly demand values. This means starting data for this operation have been pre-processed with method number two. Thanks to this operation, the subdivision into hours of the day is quite easy.

5.4 Profile splitting and interpolation.

One of the main aspects of this type of database, and in particular with all mild climate, is that space heating in winter season is not always in function, on the contrary to what happened in colder climate, Tartu for example. There are, in fact, some hours where space heating is not in operation. It is not fair, for that reason, consider a unique demand profile during winter season. It happens that not in every time there is SH + DHW, but in some hours there is only DHW consumption, if present, or only losses in the circuit if there is no request for DHW.

In addition to that, rules of municipality of Verona must been considered. For 2019-2020, the years of measurements, municipal rules said that in historical centre, space heating can stay in operation for maximum 14 hours and its period of working must goes from 5.00 am to 23.00 pm. It also said that the period of the year where space heating is allowed goes from October 15th, 2019, to April 14th, 2020. Some considerations about that: days in the dataset analysed belong to this period and from analysis of hourly energy demand, hours where space heating is turned on are evident.

Using this information, it was developed an algorithm to automatically split hours into two diverse groups: hours when space heating is on operation and when it is off operation. The division is done by tracing a line between the upper part of the profile of demand and the lower part. This operation is done by interpolating all data. Thanks to this particular behaviour of demand, the interpolation line is a good boundary for two zones. An example of plot with hourly demand versus outdoor temperature and boundary line is reported in figure 31.
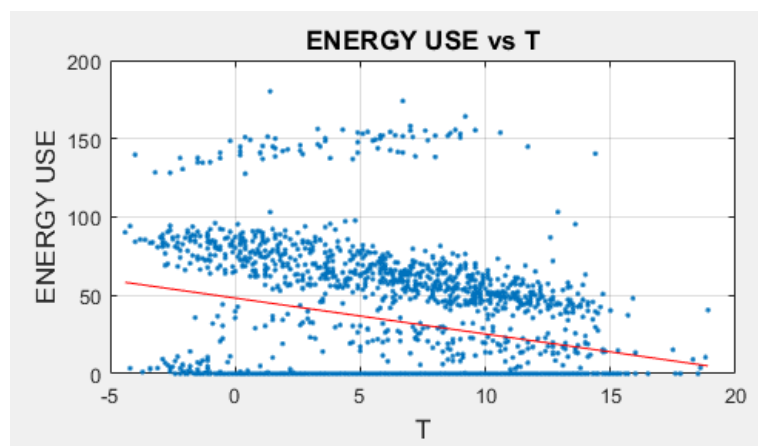

Figure 31 An example of Hourly energy demand versus T. Red line represents the boundary between moments when SH is in operation and when it is turned off.

After that, it was represented every hourly vector versus corresponding daily mean temperature, inserting in the plot also this boundary line. Thanks to this operation it is possible to find when space heating is on, i.e., demand is above the boundary lines and when space heating is turned off, i.e., demand values are below red boundary lines. In the example of figure 32 there are the plot of two different hours. It is clear that between 0.00 am and 1.00 am (the so called "ENERGY_0am", the same name as in previous analysis, as explained in paragraph 4.2 of chapter 4) space heating is not in operation. On the other hand, between 5.00 am and 6.00 am ("ENERGY_5am") space heating is turned on.
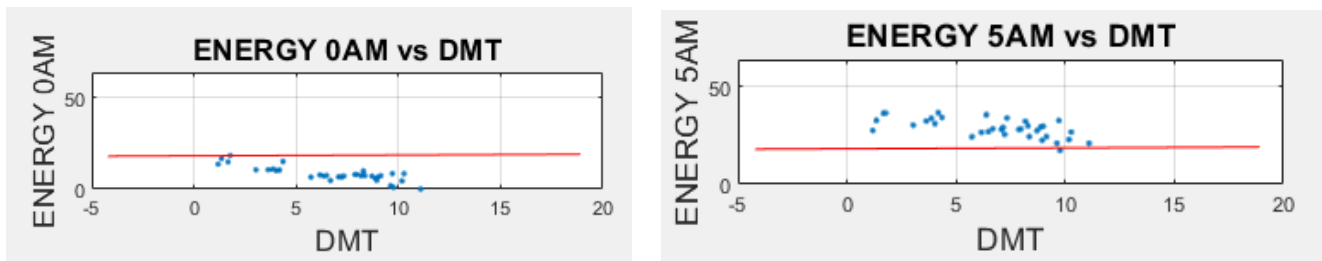

Figure 32 Representation of demand in hour between 0.00 am and 1.00 am (left) and between 5.00 am and 6.00 am (right). It is evident the difference in terms of size of heat load consumption, due to the presence or the absence of SH.

By working with experimental data, it could happen that some hours have some points above and some points below this boundary line. For this reason, it was decided to consider a percentage of tolerance. Vectors which have a certain % of points above the boundary line are considered hours of SH on, on the contrary they are considered hours without space heating. That level of percentage depends on the configuration of heat demand profile of the building and on how clear is the separation between hours with SH or without SH. In general, it was decided to keep it around 70%. The percentage value is correct when there is maximum 14 vectors in the matrix containing hours with SH on, because this is the rule of Verona's municipality.

After this point there is the creation of two matrixes: one for SH on period and the other for SH off period. In the first matrix each column represents one hourly demand vector, in case of in that hour SH in turned on. For example, if the matrix for SH on is 35x14, it means that there are 14 hours of the day where SH is in operation and that per each hour there er 35 measured values of hourly heat consumption.

In almost every building of this dataset number of experimental points is low, much less than 365 values of a complete year. This is due to the restricted period analysed and to the errors of meters.

At this point, the following passage is about interpolation of experimental data. From this point and for the rest of the analysis, group where SH is in operation is analysed separately from the group where SH is turned off. This means that in this phase two different interpolations are necessary, one for each group. In addition to that, in the previous passage there was the divisions of hourly outdoor temperature vectors with the same procedure. This operation is necessary to represent in an accurate way global heat demand hour per hour with corresponding values of outdoor temperatures.

In this phase it was used function "polyfit" from MATLAB's database to perform a linear interpolation of first degree of global heat demand with SH on and, separately, with SH off. The fitting lines have the same equation as [4.6] but in this case only the part < CPT is considered. In fact, there is no possibility to analyse summer consumption part of the demand because it is absent in the initial dataset. An example of the results of this operation can be found in figure 33, where green points represent demand of that hours with SH on and blue points, on the contrary, when SH is not in operation.
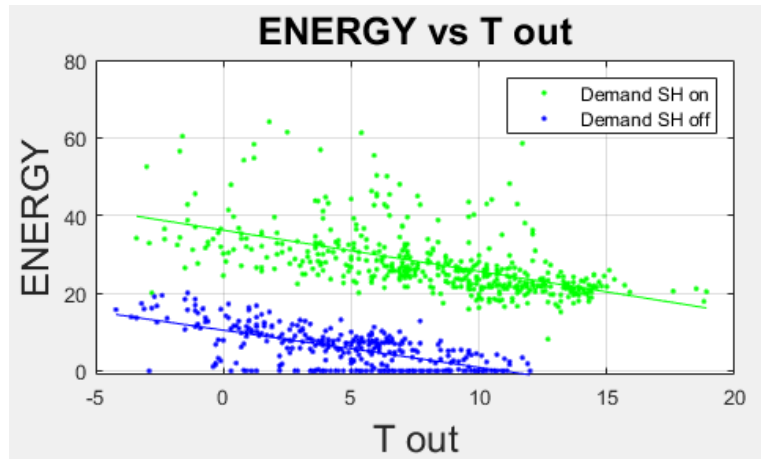

Figure 33 An example of interpolation of experimental points of global heat demand when SH is in operation (green) and when it is turned off (blue).

Up to this point, coefficients for interpolation of green and blue points are known and they are the starting values for the next parts of the method. They are, in fact, essential for the individuation of the CPT and the estimation of summer heat consumption, as illustrated in the next paragraphs.

5.5 Application of Confidence Intervals.

In this phase, the same laws described in paragraph 4.6 of chapter 4 were implemented. Literature base is the same and the code used for implementation is the same. The only difference between functions CI, CI_v2 and CI_v3 is about graphical things, such as title or axis' labels. Equations used for the calculation of upper and lower CI is the same as reported in [4.13] where parameters used are the same reported in [4.11], [4.12], and the paragraph 4.6 in general.

On this phase it was applied confidence intervals on both green points (SH on) and blue points (SH off) of the figure 33. This operation is important to have an idea of presence of outliers in the original dataset. In this first part on analysis confidence intervals have a role of visual analysis. In the second part it will be used them to exclude points outside confidence intervals and by doing this operation it was found more regular and consistent results. An example of application of confidence intervals is reported in figure 34. Here there is the same dataset as figure 33, with application of CI on experimental values with SH in operation on the left and without SH on the right.
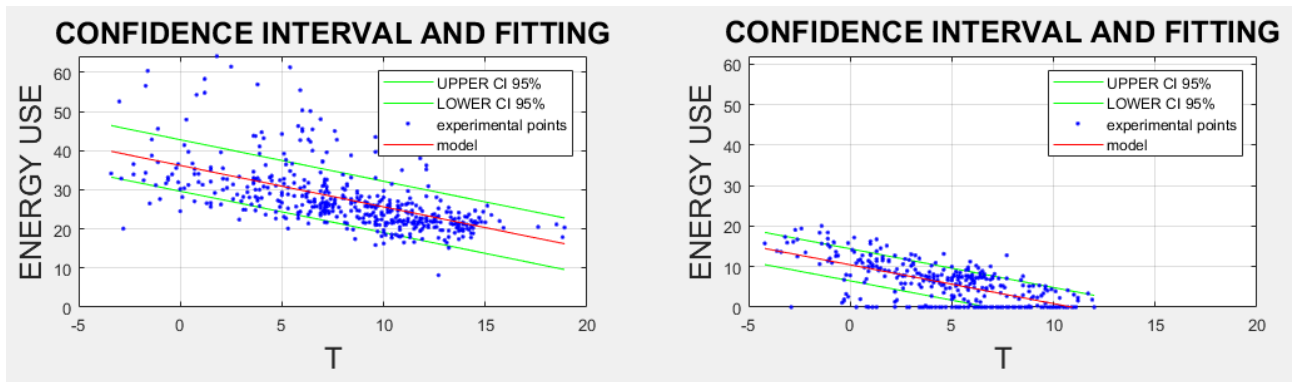
Figure 34 An example of application of CI in two different group of experimental points. This is the same building, but on the left, there are values when SH is in operation, on the right values where SH is turned off.

Also, for this case comments on results are done in chapter 7. Here it is important to underline the fact that, as can be seen in figure 34, application of CI gives interesting information about heat use of the building. In particular, it is evident that points of higher demand are outside confidence interval, which mean they belong to particular profile of operation. In the comment section it will be used this information to also explain this special regime of work for heating plants. With these considerations it is over the part of explanation of CI implementation.

5.6 Reconstruction of Energy Signature Curve.

This is the crucial phase of the analysis of partial database. As explained before, in fact, there is no presence of summer consumption. However, to obtain a good representation of building's consumptions even in summer and to apply the model for splitting DHW and SH consumption it is necessary to also find the temperature independent part of the ESC. Even if it would not be an absolutely precise representation, in this section it was proposed a statistical way to estimate that part of the energy signature curve.

In addition to that it will be discussed a method for the evaluation of CPT in this case. In fact, it is impossible the application of the method described in 4.3.1 because the temperature independent part of the demand is unknown. To solve this problem, it was decided to determine CPT in a statistical way, using, of course, only the available data about outdoor temperatures and heat load consumption.

The general idea at the basis of this procedure is to determine CPT firstly by analysis of wintertime period of global energy demand, and then use this value of temperature and interpolating curve found in the previous step to obtain a value called change point demand (CPD). This is the demand of the building when there in an outdoor temperature equals to CPT. This point also represents, in a theorical building, the constant part of the demand during summertime period. This fact is also highlighted in many studies, the work of Milić V. et al. about screening of thermal characteristics in a resident district, just to mention one of them.

In fact, assuming that the ESC part on the right of CPT is perfectly temperature independent, in a plot where global heat demand versus outdoor temperature is represented, this zone has the shape of a perfectly horizontal line. For studying the building in this case, it was assumed that on the right of CPT demand is perfectly temperature independent, so it is a horizontal line. An example of theorical ESC line is reported below, in figure 35.
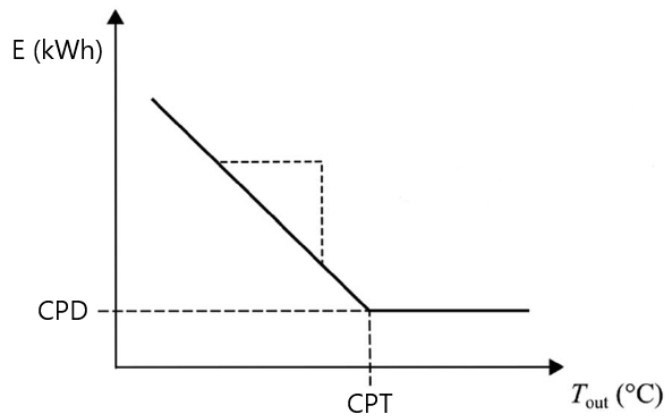

Figure 35 Theoretical ESC curve.

So, the first step consists in the determination of CPT. It would be possible to consider the maximum temperature registered in the heating season as CPT, but this could lead to possible errors. In fact, it is possible that in one single day temperature was particularly high, but this do not make us sure that this is the real CPT. To exclude this exceptional value of temperature and, consequentially, of demand, it was decided to not consider from the database values present in the upper tail of distribution. This is a pure statistical operation, with the goal to exclude points where temperature would be too high because these points would have too high influence on CPT values. If, for example, during a particular hot day, outdoor temperature is over 20°C, taking this maximum value as CPT would be a clear error. By not considering these extraordinary outdoor temperature values it is possible to obtain a better estimation of the real CPT.

So, the next step consists in identifying which values must not be considered. To do that, it was controlled that outdoor temperatures follow a normal distribution. By having this type of distribution, it is ok to assume that all points with CDF (cumulative distribution function) >95% are extraordinary points. They stay, in fact, in the upper tail of normal probability distribution function. In this phase it was implemented three different controls to be sure that outdoor temperatures reflect a normal distribution.

The first one compares a histogram of outdoor temperature values with the normal distribution with the same standard deviation and mean value as temperature distribution. An example of this plot can be found in figure 36.
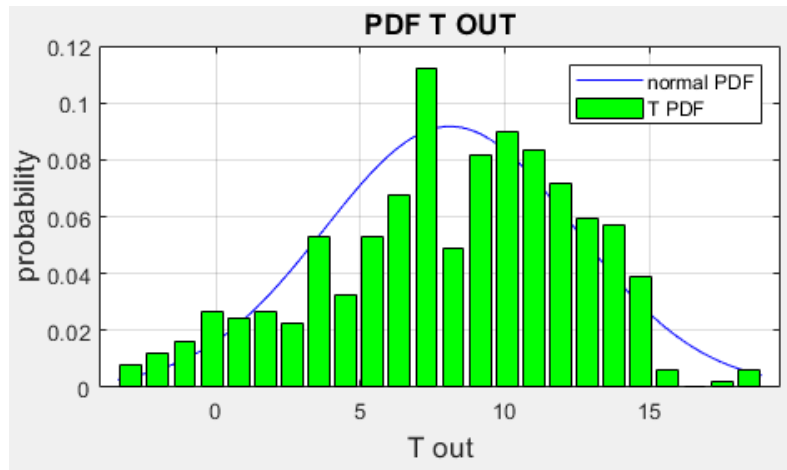
Figure 36 An example of outdoor temperature distribution, in green, and normal PDF with the same mean value and standard deviation.

The second tool it was used for this control is the comparison between CDF on outdoor temperature and CDF of normal distribution. This plot is also useful to identify which is the temperature value that has CDF equal to 95%. This is, as explained before, the boundary of temperatures: value higher than that are not considered. This means that temperature values which has CDF equal to 0.95 is the CPT needed. An example of this plot is reported in figure 37.
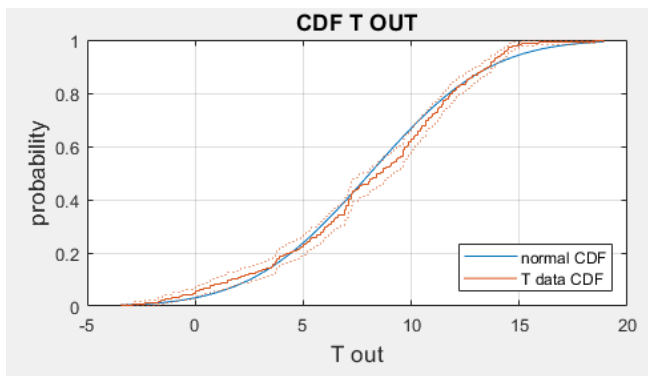
Figure 37 CDF of outdoor temperature in red and CDF of normal distribution in blue
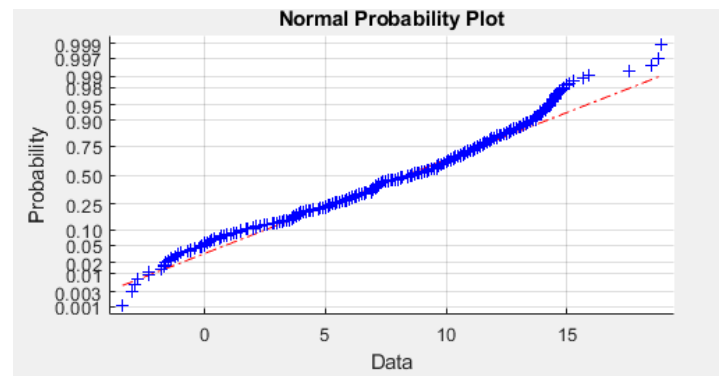
Figure 38 An example of output from the function "normplot" in MATLAB.

The last tool for controlling the normal distribution of outdoor temperature is the function "normplot" present in MATLAB's database. If the red line drawn by this function follows blue straight line, the distribution is normal. An example of that is in figure 38 and it is clear that the normal distribution is confirmed.

It is important to clarify the fact that all these operations are done on temperature, even if CPT depends on building operation, because this is a sort of filter for particular and extraordinary value of outdoor temperatures that could produce inaccuracies in the CPT estimation and could lead to underestimating summer consumption. By excluding these higher temperature points, also the lowest part of temperature dependent part of demand curve is not considered. With this fact the consumption during summer is possibly a little overestimated, but this phenomenon is preferable to underestimation or to a completely meaningless CPT definition.

Thanks to these considerations, it is possible to calculate CPT as the temperature values that has a CDF equal to 0.95. As explained, in this way all extraordinary temperature values are excluded. After its determination it was used this value to calculate CPD, assuming that this is the global heat demand of the building when the temperature is equal to change point temperature. This value of demand is calculated by using interpolation line and its coefficients, found in the paragraph 5.4 of the current chapter. The expression of CPD is expressed in the equation below, the [5.1]:

$$CPD = \alpha_{SH,on} + \beta_{SH,on} * CPT \quad [5.1]$$

Where CPD is the change point demand, also equal to consumption in summer, $\alpha_{SH,on}$ and $\beta_{SH,on}$ are the interpolation coefficients of the fitting line for the part of hours with space heating in operation and CPT is the change point temperature.

CPD is also the value of heat consumption during summer, and it was assumed constant, so that there is no dependence of temperature in this case. It was assumed that in every other hour and day of the year, demand is equal to CPD. This is, of course, a clear simplification, but it is the only feasible way to operate if other data are missing.

Of course, this would not be a detailed representation of consumption during the rest of the year, but only an estimation to perform the SH and DHW splitting in days and hours where global energy demand is measured (winter period). As maximum temperature it was considered the maximum temperature of 2020. For the summer season it was decided to use temperature values for 2020 from March to November. This means that a year for the buildings of this dataset goes from November 2019 to November 2020. In figure 39 it is reported an example ESC reconstructed and with a red line CPT is highlighted.
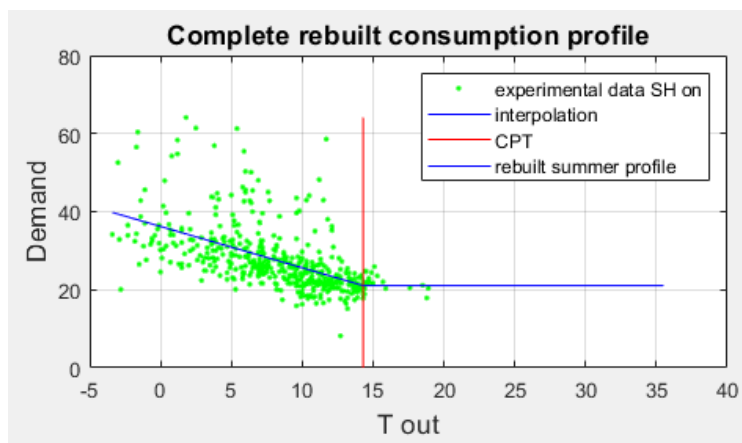
Figure 39 An example of ESC reconstructed using the method described in this paragraph. CPT is highlighted with a red line.

From figure 37 it is clear that it would be an error using the maximum value of temperature registered in the winter season as CPT. It can be seen that, in this case, this value is around 20°C, while CPT is 14.3°C. Of course, the second one is a more reasonable value for CPT, also according to other studies and from the previous analysis on Tartu database. This is an evidence that the proposed method is sensible.

After this phase there are all the elements to proceed with the SH and DHW demand splitting, described in the next paragraph.

5.7 Splitting of DHW and SH energy consumption.

The DHW and SH heat consumption splitting is performed in an equivalent way to the one done in the previous chapter. Literature references are the same, i.e., the study conducted by Dmytro I. et al. about the DHW and SH splitting in a hotel in Norway. It is important to underline the fact that all following operations, as already specified, are done with hours where SH is in operation.

The only substantial difference stays in the method for the determination of losses into the DHW water circuit. By not having the summer consumption profile, in fact, it is impossible to apply method developed by Marszal A. P. et al., previously used and explained in section 4.9.1 of chapter 4. It is explained how losses for DHW circuit are determined in section 5.7.1 of the current chapter. In the next lines it will be explained the method used for obtaining a division between DHW and SH heat load consumption.

First, it was created the vector that contains values of outdoor temperature for the summer period, as described in the previous paragraph. After that it was taken from the vector that contains experimental values when SH is on during winter, only points with corresponding outdoor temperature lower than CPT. Then it was created, using function "polyval," the demand in winter period using the regression line found before. Up to this point all is set up for starting. Results of this operation of selection and evaluation are presented in the blue points of the example in figure 40.
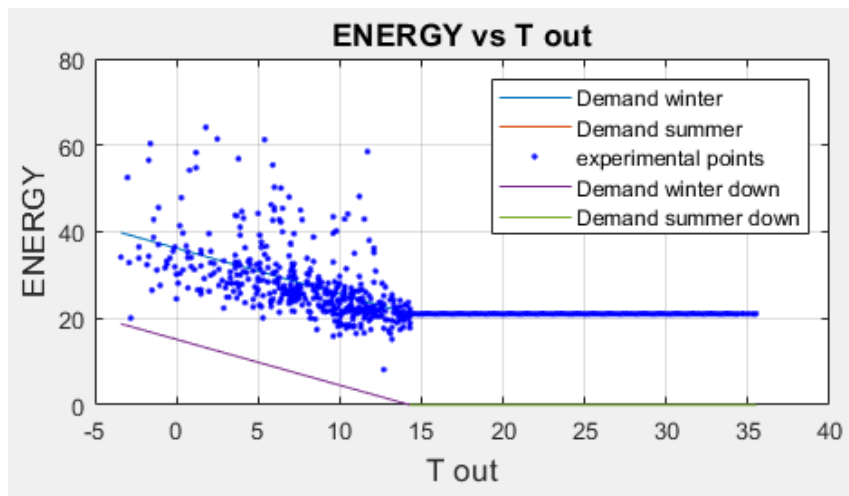
Figure 40 An example of experimental points, ESC and ESC shifted down by shifting coefficient.

After this phase, as noticed in paragraph 4.9, ESC is shifted up by a coefficient. To obtain a model for SH heat load consumption it is necessary to move down the curve by this coefficient. This value is calculated in the same way as happened before, using the same considerations. In this way, a model for SH can be found by using ESC, and it is represented by equation [5.2], where $f(x)$ represents the points of the ESC and E_{SH} is the hourly energy demand for space heating.

$$E_{SH} = f(x) - \min(f(x)) \quad [5.2]$$

After this operation of shifting, the new ESC curve is obtained. In the example in figure 40, it is represented by purple line for winter period and green line for summer period. In this

case, differently to what happened in chapter 4, ESC is continue for construction in CPT, so no operations to assure continuity are necessary.

After that, experimental values of demand are represented din chronological order. This is an important operation, useful to have a different representation of consumption and to calculate SH and DHW heat load consumptions. Due to the fact that there are so many values in a year where demand is equal to CPD, it was decided to represent only a part of the entire year in the plot implemented in the code to obtain a clearer representation of demand profile. An example of that can be found in figure 41.
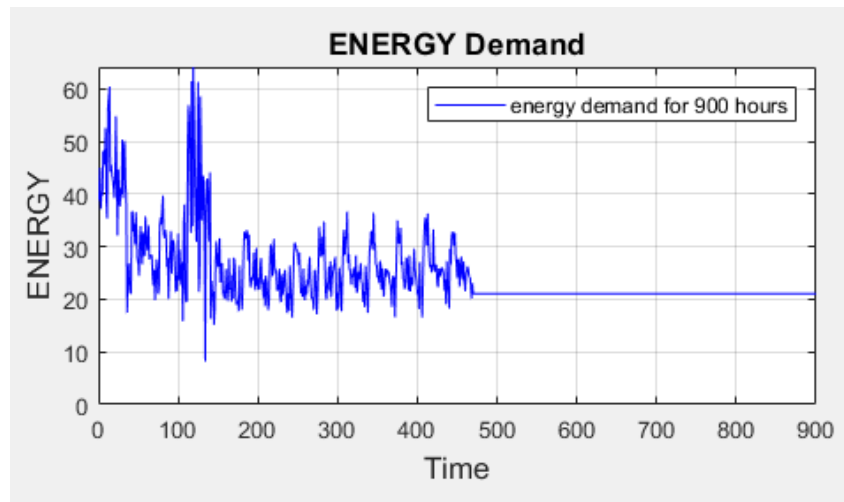
Figure 41 An example of energy demand plotted in chronological way.

To perform DHW and SH splitting by following the method explained in the paragraph 4.9 it is necessary to have an estimation of the losses in the DHW circuit. Even in these buildings, in fact, they are connected to a district heating network, and this produces losses. Knowing these losses is crucial for a good DHW estimation. The following sub-paragraph is dedicated to the evaluation of these losses.

5.7.1 Evaluation of DHW circuit losses.

DHW circuit losses are necessary to know to calculate DHW heat load consumption. In chapter 4 it had calculated them by considering the mean profile of consumption during summer and using two methods, the ones explained by Marszal A. P. et al. in their article. The idea at the basis of the first method was that the minimum value of average summer consumption corresponds to losses. On the other hand, according to the second approach, losses must be calculated considering average consumption during night hours. If database is partial, summer consumption is absent, and these methods cannot be applied. So, it was developed a new way to evaluate these losses, by using available data.

The idea is to use hours when SH is not in operation. By analysing which hours belong to this category in many buildings and considering, also, rules from Verona's municipality, a major part of these hours is during night. While there could be some differences between buildings in terms of hours of operation of the space heating plant, in general it was found that during night SH is always turned off. This is due to the fact that most of users are

sleeping if we are considering a residential building or they are not present inside the building, if we are talking about offices or schools, for example.

Starting from these considerations, it was decided to use energy demand between 0.00 am and 4.00 am of hours without SH to estimate losses in DHW circuit. In fact, as proven by statistical survey K. Winqvist, more than 90% of Europeans sleep in this time range. So, because users are sleeping, they do not use DHW for purposes like showers or dishwasher. For this reason, the first operation that has be done in this phase is to select hourly vectors that contain energy demand between 0.00 am and 4.00 am.

After that, it was imposed a filter on temperature. In fact, it was decided to consider only the highest values of temperature. It was noticed that for these hours the fitting line has a decreasing behaviour. This fact could be caused by the seasonal difference in DHW losses and their dependence to outdoor temperature. It is possible to notice that, as should be, when outdoor temperature is lower, demand in these hours is higher. This is because DHW must be available at a certain temperature and if outdoor temperature is lower, energy to heat up water up to the wanted value of temperature is higher, according to an equation similar to [4.1]. Considering the entire experimental values of these hours leads to a big overestimation of losses during summer period, because in this case outdoor temperatures are higher, so losses are lower. This approach is consistent with analysis of buildings, as it will be explained in the results chapter (7).

To consider only highest outdoor temperature it was decided to use the normal distribution filter. It was checked that even these values of temperature follow a normal distribution. Of course, this is verified because also temperature values when SH is in operation follow a normal distribution, as showed in the previous paragraph. After the normal distribution check, it was considered for losses evaluation only outdoor temperatures with CDF > 90%. This percentage is decided based on the analysis of many datasets. 90% allows to consider the highest values of temperature but without the risk to have too few points. It is not shown an example of tools for normal distribution checking because they are the same used in the previous paragraph.
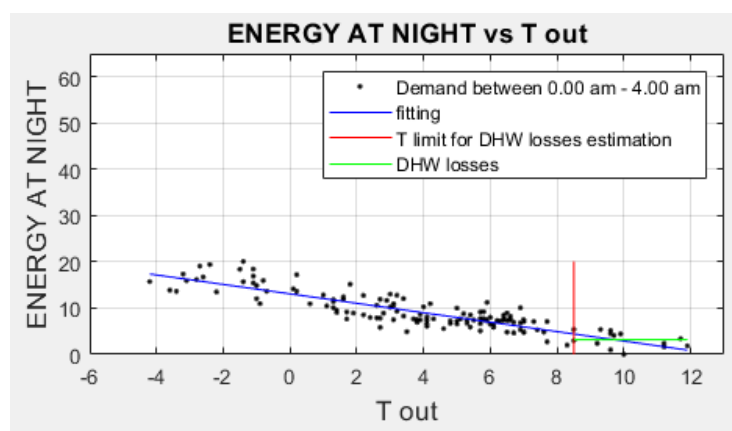
$$E_{loss} = mean(E_{exp,90\%}) \quad [5.3]$$

Figure 4236 An example of DHW losses evaluation using experimental points of hours when SH is turned off.

Where E_{loss} is the value for losses in the DHW circuit and $E_{exp,90\%}$ are a vector that contains all experimental points in the considered period with corresponding outdoor temperature with CDF higher that 90%.

After that, considering these points of the hottest part of the period without SH, it was found an estimation for losses by considering mean value between them. After that it was represented both boundary for outdoor temperature considered (red line in figure 42) and value for losses in DHW circuit (green line in figure 42).

Thanks to this method for the evaluation of DHW circuit losses it is possible to proceed with SH and DHW splitting. Passages are the same conducted for the complete database, described in paragraph 4.9. As a recap, first E$_{SH}$ is obtained by shifting down the ESC by a value called shifting coefficient. This is equal to the minimum demand in summer period, in this case it is equal to CPD. The equation of this step is reported in [5.4]. It is similar to equation [5.2], but for clearness of exposition here it was substituted the minimum value of ESC with CPD.

$$E_{SH} = f(x) - CPD \quad [5.4]$$

Where $f(x)$ is the reconstructed ESC equation and CPD is the change point demand (the value of global energy demand corresponding to CPT and assumed to be summertime period constant consumption of the building).

After that there is the determination of E$_{DHW}$. There are reported below, in equations [5.4], formulas used for the determination of this heat load consumption.

$$E_{DHW} = \begin{cases} E_{TH} - E_{SH} + E_{loss}, & E_{TH} > E_{SH} \\ E_{loss}, & E_{TH} \leq E_{SH} \end{cases} \quad [5.5]$$

Where E_{DHW} is the heat load for DHW, E_{TH} is the experimental point, the global thermal energy demand, E_{SH} is the heat load for SH, obtained in the previous lines, from equation [5.4], E_{loss} is the value for losses into the DHW circuit.

Then, there is the balancing phase, to be sure that in every moment the sum of DHW and SH heat load consumption is equal to experimental measurements. After that, the last operation consists in a correction process to have all values, both hourly SH heat load and hourly DHW heat load not negative. If there is a negative value in one of them, it becomes zero and that part of the load is given to the other hourly load. These same operations were done also for the complete database, in paragraph 4.9. An example of results is reported in figure 43. In this image there are both splitted demand versus outdoor temperature and splitted demand versus time. Comments and considerations on chapter 7.
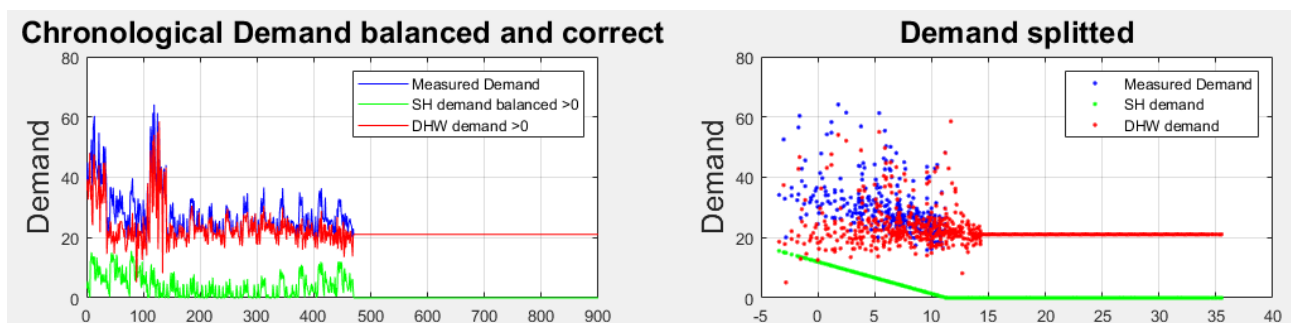

Figure 43 An example of results after splitting operation. Demand vs time on the left and Demand vs T out on the right.

It is relevant to underline the fact that all this analysis provides satisfactory results for the part of the year where experimental data are available. After CPT, in figure 43 where red line becomes horizontal, this type of analysis does not have sense and so results given. It would be meaningless try to estimate in a precise way demand in this zone because there are zero information about summer consumptions, so the estimation would be too inaccurate. However, results for the experimental values analysis are consistent.

5.8 Analysis with outliers' remotion.

Due to the high dispersion of experimental values in the analysed partial datasets (belonging to Verona's case study), it was decided to repeat the previous analysis but after an outliers' remotion. It was used confidence intervals, whose implementation is described in paragraph 5.5 of the current chapter, to identify outliers. For this part of the analysis, in fact, it was considered only points inside confidence interval. Results obtained follow a more regular profile, but in this way some information about extraordinary heat demand are lost. Although, this is an interesting operation to see if this strange demand points influence the general profile of consumption and what is they influence.

The only difference between this part of analysis and the previous step is the initial phase. In this site, in fact, it was used only measured data inside confidence intervals, both for hours of SH operation and for hours without SH. The phase of points selection happens after SH on and SH off period identification. In figure 44 it was reported considered values in green and excluded values in blue. It can be noticed that remotion has an effect mostly on highest values of demand. As a results, there is the absence of extraordinary peaks of demand consideration.
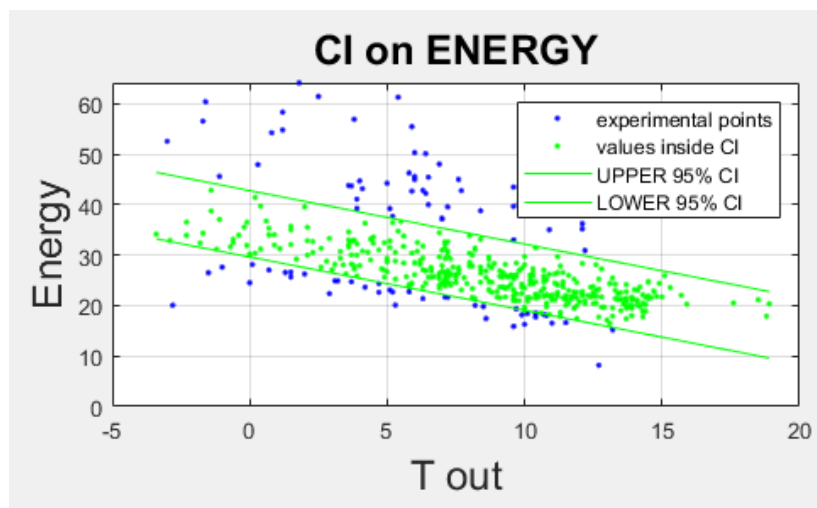

Figure 44 An example of outliers' remotion conducted with Confidence Intervals.

All other phases are not repeated here, because they are the same described in previous paragraphs. Comparison between results with and without outliers' remotion is done in chapter 7, when it will be discussed results of Verona's partial databases analysis.

Chapter 6: Case studies.

In this chapter it will be provided a general description of the two datasets where it was applied tools described in the previous chapters. A total of 26 buildings from Verona and 51 buildings from Tartu were analysed. For Tartu's dataset it was applied method described in chapter 4, because every building has an entire year dataset, while for analysing Verona's buildings it was applied method for partial database, as described in chapter 5. The next two paragraphs talk about Tartu and Verona ad provide as much information as possible, paying attention to observe privacy rules.

6.1 Tartu.

This dataset is about some buildings located in Tartu, a city in the east side of Estonia. These data were already used in other publications. Some previous work about the same buildings were conducted by Lumbreras M., Garay-Martinez R. et al. where they built a data-driven model for heat load prediction or by Eguiarte O., Garay-Martinez R. et al. about a data-driven assessment for the supervision of district heating networks. In this paragraph it will be used information taken from these works because they provide a good description of Tartu's dataset configuration.

First, it is important to specify that all buildings are connected with a district heating network. The DH network in Tartu, Estonia is operated and owned by GREN, delivering heat between temperature ranges of 75/45ºC as minimum temperatures and 110/60ºC as maximum water temperatures to over 1500 consumers, producing up to 500GWh yearly. Approximately 40-60 new connections are made to the grid annually, with 49% of consumers being collective housing, 33% industry and commercial buildings, and 18% individual housing.
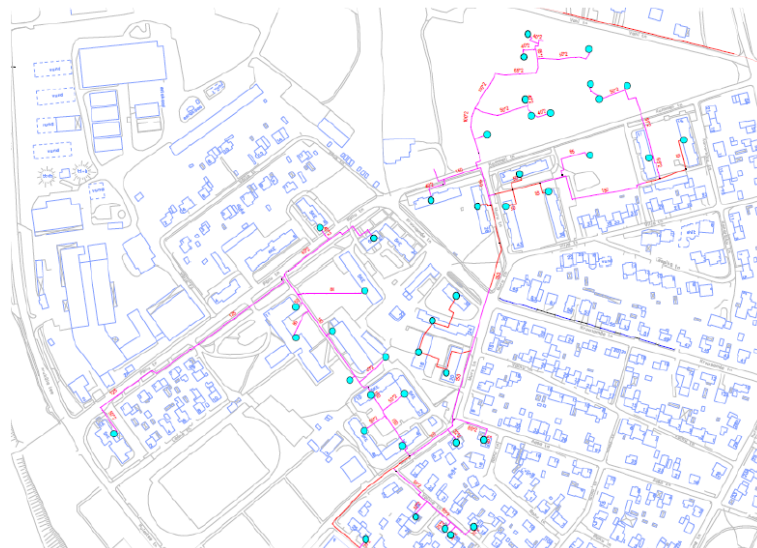
Figure 45 Map of district heating network in Tarkon area, in Tartu, Estonia.

The dataset used in this thesis is related to the subnetwork and consumer substations within the Tarkon-Tuglase area, in the northwest of Tartu. The network covers a distance of 5.34 km, serving over 54 consumers with ongoing connections. In fact, as said before, it was analysed 54 different buildings, the ones with available measurements from this zone. The delivered heat ranges up to 4.3 MW and is used for residential, commercial, educational,

and office buildings. The total consumption is around 8.2 GWh. In figure 45 there is the map of the Tarkon's district heating subnetwork. Connections are represented with red lines and light blue dots are for users.

Hourly measures come from smart energy meters located in each building. Every meter has constantly measured different variables of the system and has sent them to the DH operator, GREN hour per hour. As showed in the pre-processing step in paragraph 4.1, every smart meters has measured the cumulative energy delivered to the building (so, hourly energy use is calculated by subtracting the current hour's reading from the previous hour's reading), the cumulative volume of water consumed, the hour of the day, the hourly power, hourly flow, water inlet temperature, return temperature, hot water user temperature, secondary hot water flow temperature and secondary return temperature. A scheme of location and layout of the smart energy meters in the district heating network of Tartu is reported in figure 46. Letters T indicate temperature meters, while "m" is the flow sensor. Five different temperatures (in °C) and volumetric flow (in m^3/h) are measured. For this study, the total
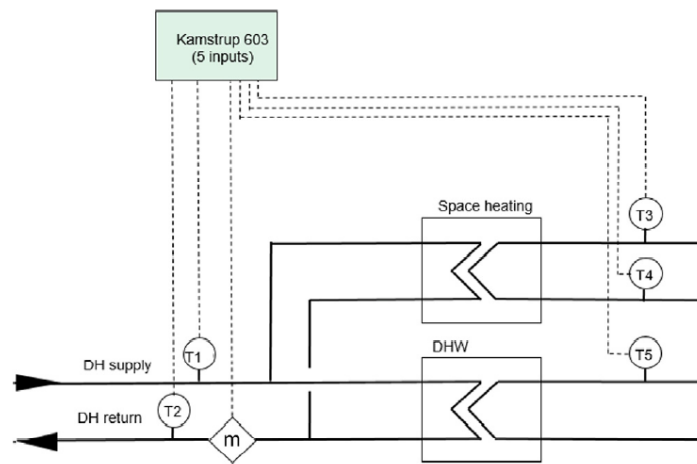
Figure 4637 Scheme of smart meter and sensors' location in DHN of Tartu.

heat consumption is calculated using T1, T2, m, and the specific heat of water.

Each building is assigned an ID code for identification purposes, which is unrelated to the building's physical location. This measure is taken to protect users' privacy and prevent identification issues. However, as highlighted several times during the discussion, this fact makes impossible to know any information about physical parameters or users' behaviour of buildings in analysis.

The energy meter used in the buildings is the Multical® 603 from Kamstrup, technical parameters and documentation are available in the manufacturer' website reported in references. Its accuracy surpasses the European directive (EN-1434e1:2015) for this purpose, and the measuring error stays under 5% in all variables recorded. Each substation, which is each smart meter, corresponds to one building. The studied substations encompass various thermal zones based on their ultimate use. Residential apartments, offices, educational buildings, and commercial buildings are included.

The other important variable to consider is the outdoor temperature. Its measurements, together with other climatic variables, were obtained from a weather station managed by the University of Tartu. It was used only outdoor temperature, but in the weather variables dataset there were also global solar irradiance on a horizontal plane (GT) measured in watts

per square meter, wind speed (WS) measured in meters per second, and wind direction (WD) measured in degrees.

It was used weather data already pre-processed by Lumbreras M., Garay-Martinez R. et al. They combine the heat load and weather data sources, a calendar-based process. The weather data was obtained at a frequency of 15 minutes but was resampled to hourly intervals to match the frequency of the data from the smart energy meters. This resulted in 8760 readings representing each hour of the year. Both data sources were coupled using the exact date and time (month, day, and hour) for the year 2019.

The authors also did an operation of identification and remotion of outliers and reading errors contained in both datasets (from smart meters and from weather data). Reading errors are directly removed from the original dataset, resulting in a reduction in the total number of data points available. To identify outliers, quartiles of each variable are calculated using boxplots. The interquartile range (IQR) is used as a criterion. This is a good method, also used in other studies, for example by the authors Li Aihua et al. or Schwertman N. et al. For daily data, there was also a remotion of incomplete days (i.e., those with less than 24 measures), and there is no need to calculate missing values.

7.2 Verona

This is the second case study analysed in this thesis. This case is hugely different from the Tartu's database. First, climate is vastly different. Tartu has a colder climate, with almost no influence of relative humidity of air. Verona, on the other hand, has a milder climate and relative humidity is extremely important in terms of comfort conditions. However, for the purpose of this investigation, relative humidity influence is not taken into account, while higher outdoor temperature is a relevant data for the heat load consumption of the buildings.

This fact plays a significant role especially for establishing the beginning and the end of the heating season, which is the operation of the space heating plant and its intensity. In particular, colder climates have more consumption than milder climates, of course. Regarding CPT, instead, not big differences were noticed during the analysis. This is due to the fact that CPT depends on the outdoor temperature: when it becomes higher that a certain value, space heating is no more necessary, and it is turned off. Fort this reason, CPT is similar in both climates.

Another significant difference between the two case studies is the experimental data available. As said on many occasions during the previous chapters, Verona's database contains data about winter demand only. This means summer part of consumption is completely absent in every analysed building. This fact makes the application of certain tools. For example, the detailed analysis of DHW summer consumption profiles or the analysis of DHW boxplots are unfeasible.

In addition to that, Verona's database was not already pre-processed, in contrast to Tartu's database. As explained in the previous paragraph, in fact, in Tartu's values authors of previous studies already performed a good operation of outliers' remotion and incomplete days' remotion. In Verona's case, instead, there were a lot of missing values and some extraordinarily high points of heat demand. In this case the operation of pre-processing was extremely important to obtain convincing results. All operations done in this data for this

purpose are described in paragraphs 5.1 (empty values and reading errors remotion) and 5.8 (outliers' remotion with confidence intervals).

However, in Verona's database, some information about the intended use of buildings and their physical parameters, in certain cases, were available. The following description of Verona's district heating network and information about buildings are taken from the analysis performed in the same dataset by Quaggiotto D. in his PhD thesis.

Experimental measurements come from the district heating network of Verona Centro Città. Verona, with a population of 259,610 inhabitants, is the largest city municipality in the Veneto region. As stated by Quaggiotto D., Verona's climate characteristics are taken from UNI 10349-1:2016. The degree-days is 2,469; climatic zone is E; external design temperature is -5°C (notice the highest temperature compared with Tartu); the annual heating period is composed by 183 days.

In figure 47 there is a satellite photograph that shows Verona's district heating networks, which consist of five networks with a total length of 220 km and 1318 user substations. The thermal load peak is 170 MW. Two of the networks, Forte Procollo and Borgo Trento, are not hydraulically connected but can exchange heat via a heat exchanger. These two networks are not subjects of analysis.
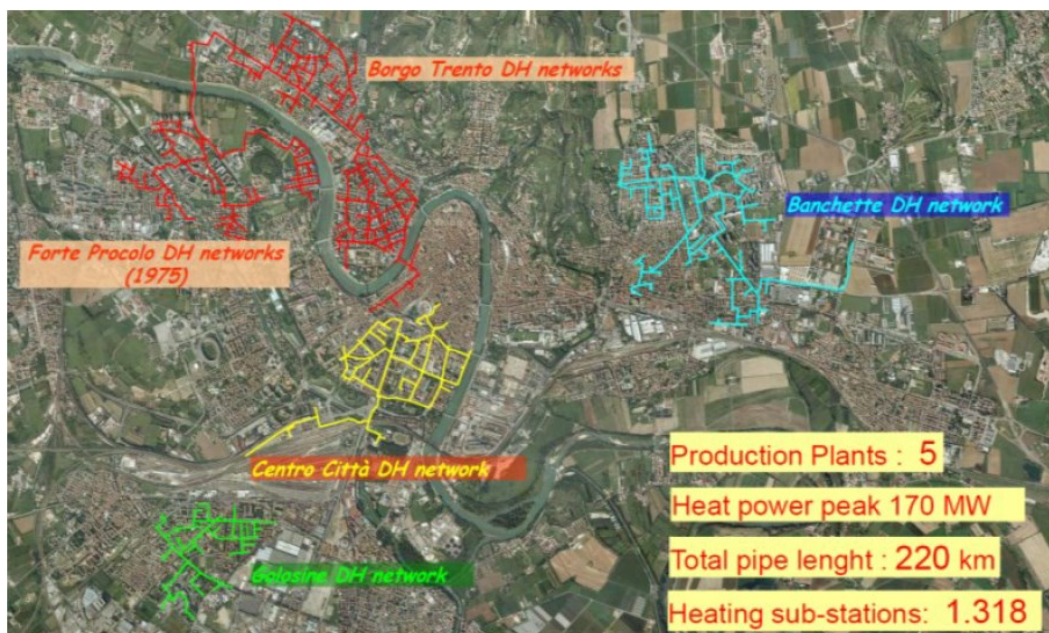
Figure 47 Satellite's photo of Verona's district heating network. Taken from Quaggiotto D.'s work.

The focus of this case study is on the district heating network (DHN) of Verona Centro Città, as depicted in figure 48. The DHN operates with a constant supply temperature and variable flow rate, and serves 247 users with space heating and, in some instances, domestic hot water. The network covers a distance of approximately 25 km, with a volume of 652 m^3. The total volume of the substations connected to the network is about 3.2 Mm^3, and they require around 70 GWh/year of heat, with a load peak of about 38 MW.

In figure 48, the user substations are depicted as blue dots, while the supply stations are represented by three green dots. The primary supply station, CCC, utilizes five CHP units, heat pumps (which are supplied by the cooling circuit of the CHP), and three auxiliary gas

boilers to produce and supply heat. The other two supply stations, CRV and CSD, have three auxiliary and reserve gas boilers, with CRV also recovering heat from a foundry. They

DH network length	25 km
User substations	247
Volume of connected buildings	3.2 Mm^3
Annual energy demand	70 GWh$_t$
Thermal peak load	38 MW$_t$
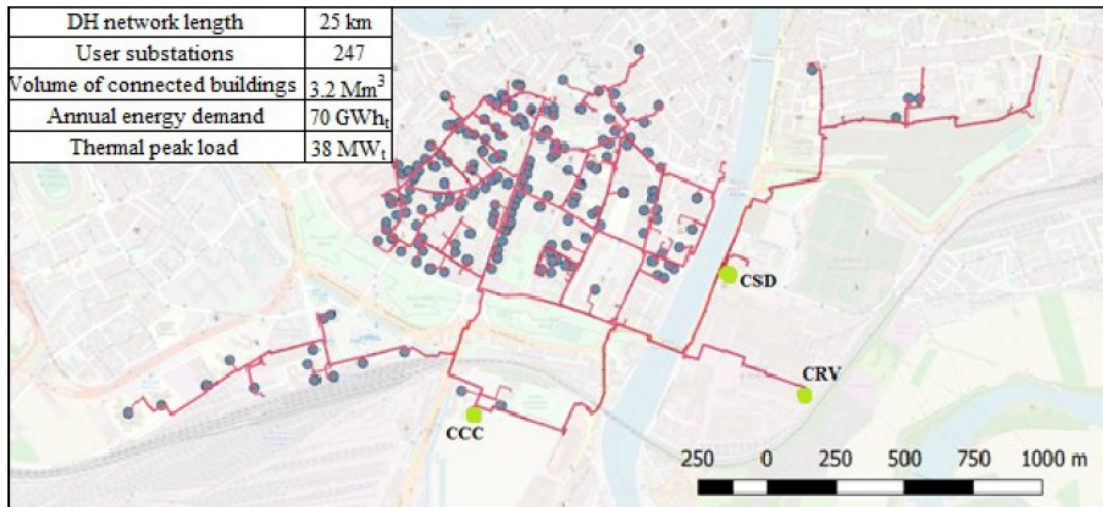
Figure 48 Plan of the DHN of Verona Centro Città. red lines are pipes, blue dots are users and green dots the primary supply stations.

have different priorities in terms of heating supply during the year, due to their differences in power capacity created by distinct types of generators. However, primary supply stations are not analysed in this work, so no more information is necessary.

Regarding the buildings analysed, there are some information available. There are 26 buildings where consumption data are available, but information about year of construction, floor area and other general information are available for 16 of them. They are reported in the table in figure 49. For every building it is available the intended use. In fact, it is indicated in the name of the dataset. The analysis was performed for: 12 condominiums, one gym, 2 hotels, 2 offices, one police station, one rehab centre, 5 schools and 2 train stations.

Typology	Age	Floors	Height [m]	F. area [m^2]
Office A	1990	4	13	3350
School A	1870	3	9	1080
Office B	1980	3	9	445
Condominium A	1960	6	22	788
Condominium B	1960	6	22	788
Condominium C	1930	6	19	167
Condominium D	1930	5	16	148
Hotel A	1950	6	19	119
Condominium E	1960	6	19	540
Condominium F	1920	5	16	729
Hotel B	1910	4	13	534
Office C	1940	4	13	324
Condominium G	1930	6	19	162
Condominium H	1970	4	13	94
Condominium I	1970	6	19	292
School B	1960	3	9	2325

Figure 49 Table with general information about 16 buildings of the Verona's dataset.

As specified some lines ago, Verona Centro Città's DHN is connected to 247 user substations, but only 26 of these have been monitored. Monitoring for these 26 substations was conducted from October to November 2019 until March 2020 by Gizero Energie, a specialized energy monitoring company. The monitoring included measurements of maximum, minimum, and mean power in kW for each hour, water flow in m^3/h, and supply and return temperatures in °C. For some substations, pressure data in the supply and return pipes was also provided, but they are not of interest for this analysis. All of these measurements corresponded to the primary circuit of the heat exchanger installed at each user.

Ultrasonic compact energy meters were installed to measure energy consumption in the district heating systems. These meters operate on the principle of transit time measurement and do not have any moving parts, which prevents mechanical wear of the metering components. They also have low-pressure losses, a low start flowrate, and good tolerance to suspended particulates in the water flow. Overall, the use of ultrasound technology ensures stable and accurate measuring results.

For the knowledge of general information about buildings it must be thanked AGSM Energia Spa, the energy supplier, which provided key information for these buildings, such as the call sign, ID number, address, heating and DHW heat exchanger sizes in kW, year of plant installation, and invoiced energy demand for 2017. The measured data and nominal data were analysed in the work by Quaggiotto to evaluate the level of reliability of the energy supplier's information. After his pre-processing operations, higher quality datasets were obtained, and this is the starting point of this analysis.

Chapter 7: results and comments.

After the illustration of the tools used for the analysis of complete and partial databases and the presentation of real case studies done in the previous chapters, in this one it will be illustrated analyses' results and provide some comments and explanations about them. As said before, it was analysed a total of 77 buildings (51 form Tartu and 26 from Verona).

From their results it was found some similar characteristics and behaviour among them and also some cases where the method followed does not produce satisfactory results. In this chapter It will not be presented all buildings, it would be too time consuming and useless, especially for buildings that have similar behaviour. In next lines there will be shown the most interesting buildings for each dataset, in terms of results or considerations. In each title the building's name and the corresponding case study are indicated. The presentation begins by presenting some cases of Tartu's buildings, each one with a complete dataset.

In each paragraph it will be shown results of the implemented code for the analysis, providing explanations, interpretations, and considerations where they are necessary.

7.1 Building "10922" in Tartu.

In building 10922, but this is true for every building with complete dataset due to the fact that the method used is the one presented in chapter 4, there are a total of 28 figures as code's output. Starting from the first ones, figures from 1 to 24 represent a different hourly vector analysis. This mean that figure 1 of the output represents the analysis for every experimental point where demand is between 0.00am and 1.00 am. It is the so called "ENERGY_0am" vector mentioned in chapter 4, paragraph 4.2 as a result of one of the first operations done in the dataset. In figure 50 reported below, it was reported the "figure 1" that comes from the output of the code. It is unnecessary to report all 24 figures, so it was insert only the first one as an example.
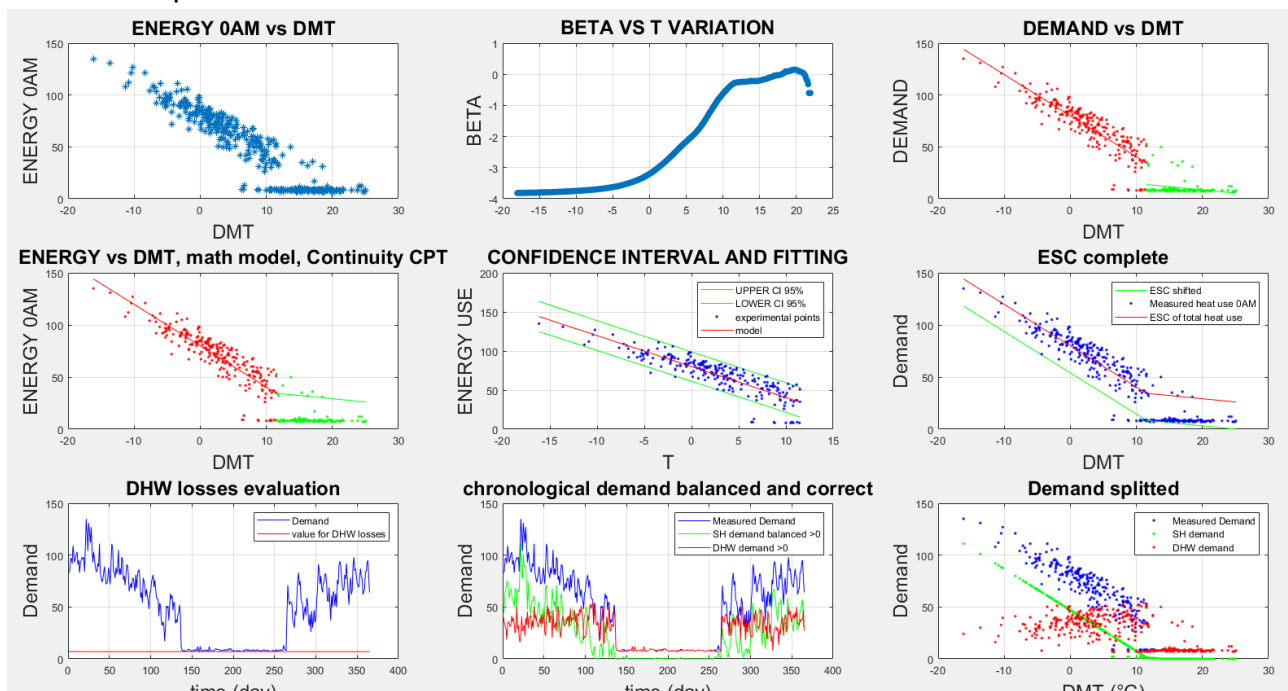

Figure 5038 Image 1 of the output of the building 10922's analysis.

In figure 50 there all plots previously mentioned during the explanation of method. Starting from the upper left corner we can find the energy demand plotted versus outdoor temperature, only of the hour between 0.00 am and 1.00 am. By analysing also other hours the pattern is similar, there is always a temperature dependent part and a temperature independent part. In this building no particular consumption is noticed in specific hours: more or less the general pattern is the same for every hour. This happens because Tartu has a cold climate, where the necessity for space heating is present all over the day in winter season and sometimes also during hotter months of the year.

This fact can be noticed also by observing values of demand. This is the consumption of the building, and the maximum is almost 150 kWh, that means, because it is referred to one hour of measurement, the maximum power registered in the building is 150 kW. This is a high value, even if this is not a specific consumption, so it is possible that the demand is high because the building is particularly big.

Moving to the right we found the coefficient beta plot, for estimating CPT point, as described in paragraph 4.3.1. After that there is the separation between temperature independent part and temperature dependent part, with the separate interpolation of both groups of points. Proceeding in the central part of the figure there is the ESC and, in the centre, the application of confidence interval in the temperature dependent part of the ESC. It is possible to notice that almost every point is inside a 95% confidence interval, so it means that demand in this building is regular all over the year. No extraordinary points of demand are noticed.

It is important to underline the fact that usually in colder climates demand appears more regular. This happens because space hating demand is higher and present during every hour of winter season. For these reasons peak of demand, typically caused by abrupt requests for DHW, are confused among these high SH consumption points. This fact emerges in Verona's analysis, where demand peak caused by DHW are much more evident, and where municipality rules impede the continuous operation of SH plant.

After this plot there is the representation of the ESC shifted down by the shifting coefficient. In this way the green line on the 6th plot represents the model for space heating consumption. It has a shape like that in every analysed building: space heating demand is zero with high outdoor temperatures (>CPT) and it increases when outdoor temperatures decrease. In all real cases, the right side of CPT is not perfectly temperature independent. A small effect of reduction of demand is still present due to reduction also in DHW heat load when outdoor temperature increase. This happens because, even if more or less the temperature at which hot water is requested from the users is constant all over the year, water that arrives from the aqueduct has a higher temperature because the external environment is hotter during summer. This fact produces a decrease in energy demanded to heat up domestic water.

On the lower side if the figure there is the chronological representation of the demand. Here it is evident the difference between winter and summer. The separation between two periods is sudden, there is no smooth passage. Furthermore, difference in terms of consumption size is relevant between the two periods. This highlights the fact that space heating has an especially significant role in this building and gives an important contribution in terms of power required and total energy consumed during the year from the building. From this plot, and the successive information about DHW and SH consumptions, it is possible to say that with high probability thermal insulation of this building is not at the top. In well insulated

buildings, in fact, there is less need for high space hating power, while domestic hot water consumption becomes more relevant. Here space heating is much bigger than DHW. However, it is important to keep in mind that a high space heating demand is still explainable by the cold climate, with outdoor temperatures up to -15°C.

In the same plot, in red, heat losses in DHW circuit are highlighted. As explained in paragraph 4.9.1, they are equal to the average consumption during night-time in summer. Compared to winter consumption, summer consumption seems to be much more regular and, also, much lower. Summer consumption represents only DHW heat load, so we can assume its consumption follows a quite regular pattern, without sudden spikes. This fact, combined with the high consumption, can suggest the fact that the building in analysis is a big building with the necessity to keep SH during all hours. By analysing the other hours of the day, it is possible to notice that DHW mean value is higher between 20.00 pm and 22.00 pm, and has a small peak also in the morning, between 8.00am and 9.00 am. All these clues could lead to the interpretation that this is a big residential building. This thesis is made stronger also by the difference in profile patterns between workdays and weekends, as explained in the following lines.

By proceeding with the last two plots, we can find the splitted demand versus time and splitted demand versus outdoor temperature. Starting from the first one, here it is evident the fact that in summer there is only DHW consumption. This is tight, due to the fact that during the hottest part of the year, the ESC becomes temperature independent, so the heat load consumption is almost constant, and it is only for DHW. During the rest of the year SH becomes and important heat load. Because it was calculated from a linear model, and this fact is evident in the last plot, where SH heat load consumption follows a line during winter, SH has a small relevance in peaks of consumption.

However, this fact is consistent with physical behaviour of the building: in case of sudden peaks in demand, it is more likely that the cause is DHW than SH. This happens because DHW is decided by users, and it is necessary a big amount of power to heat up water up to the desired temperature in relatively small time. Imagine having a shower: the hot water must be available in a few seconds, and it must be provided for the entire shower's duration. However, it is possible to notice that even in summer, when there is only DHW consumption, there no relevant peaks of demand during the day. There are some hours where mean consumption is higher, as mentioned before. This fact could be due to, for example, the generous size of the building: if there are many apartments, especially if inhabited by users with different habits, the peak in load consumption due to a shower of one user cannot be noticed in the overall demand profile of the entire building. Nevertheless, it is important to specify that this is just the most likely interpretation of the building's intended use, but it is not the only one.

It is hard to say something about physical characteristics of the building because its floor area is unknown. Without his information, it is impossible to calculate specific heat consumption and, of course, this value cannot be compared with typical values for buildings in the same geographical area. If floor area of the building would be known, a lot of further considerations could be done.

It would be possible to obtain information also about losses of the building, in particular by knowing power of heating system hour per hour it would be possible to determine the sum

of losses for transmission, ventilation and infiltration. However, to calculate them, internal temperatures hour per hour are necessary. This is a crucial data, without indoor temperature value for losses cannot be calculated. It would be possible to estimate indoor temperatures, using standard values for similar buildings in the same area, but this estimation would be too inaccurate. Previous research works used this technique, for example the study by Milić V. et al. or by Eriksson M. et al., but they started from a certain data of indoor temperature.

Another information to obtain regards the slope of the temperature dependent part of the energy signature curve, i.e., the model for SH heat load consumption. This data provides an interesting estimation of the total heat loss coefficient. Of course, it represents a specific value, measured in (kW/°C). Its expression can be found in equation [7.1].

$$Q_{tot} = Q_{trans} + Q_{vent} + Q_{infilt} \quad [7.1]$$

This expression is reported in the previously mentioned studies. In this equation Q_{tot} represents the specific total heat loss coefficient (in kW/°C), Q_{trans} are losses for transmission, Q_{vent} losses for ventilation and Q_{infilt} losses for infiltration. It is impossible to separate Q_{tot} into these three components without making strong hypothesis and approximations about heat loss coefficients, users' behaviour, physical and geometrical parameters of the building. It was decided to not approximate them because no information in this field is known, and the inaccuracies would be too much.

However, for every building the coefficient Q_{tot} is known and by its quantity it gives an idea of the insulation degree of the building. With high values, it means that the building is poor insulated, if outdoor temperature increases a little, losses increase a lot if the coefficient is high. On the contrary, with well insulated buildings, this coefficient is little. These considerations are valid for every building analysed in this thesis.

For the building in exam here, the 10922 in Tartu, $Q_{tot} = -4.6\ kW/°C$. Taken alone does not give much information, but compared with values from other buildings in the same dataset it is possible to understand, for example, which is the one with the highest coefficient and find in this way where to focus the attention for an eventual retrofitting intervention. Going further, it is possible to check if every building has the same coefficient of if there are significant differences among them. Another important consideration about this data is a comparison between this coefficient obtained for different hours of the day in the same building. By performing an hourly splitted analysis, in fact, it is possible to compare values of Q_{tot} calculated in all the 24 hours of the day. In the table reported in figure 51 there are values for Q_{tot} for the 24 different models obtained by analysing the hourly splitted energy values. In figure 52 there are the same values plotted for clearness of illustration.

The value mentioned before for Q_{tot} (-4.6 kW/°C) was obtained by considering complete yearly model of the building. The output of the analysis of all 8760 experimental values for demand considered together is presented in figure 53 and discussed in the following lines. In this site it is relevant to underline the fact that the value found for the complete model is an average value between Q_{tot} reported in table in figure 51. It can be noticed that there are few hours where this coefficient is higher in absolute terms, and this means that there is a stronger dependence on temperature for SH consumption, i.e., for global losses.

The higher values are in the morning, with a peak between 9.00 am and 10.00 am. This means that in this time range, losses for transmission, ventilation and infiltration are higher. Due to the fact that materials of the building or its geometrical parameters cannot change between different hours, this variation could be caused by users' behaviour. In particular, there could be an influence of inhabitants on aeration of environments or simply on indoor temperature, by regulation of thermostat. This is only a hypothesis and cannot be confirmed because no information on this sense is known.

N hour	Hours of the day	Qtot
1	0.00 am - 1.00 am	-3.96
2	1.00 am - 2.00 am	-3.87
3	2.00 am - 3.00 am	-3.83
4	3.00 am - 4.00 am	-3.82
5	4.00 am - 5.00 am	-3.81
6	5.00 am - 6.00 am	-3.81
7	6.00 am - 7.00 am	-4.17
8	7.00 am - 8.00 am	-4.53
9	8.00 am - 9.00 am	-4.75
10	9.00 am - 10.00 am	-5.08
11	10.00 am - 11.00 am	-4.69
12	11.00 am - 12.00 am	-4.45
13	12.00 am - 13.00 pm	-4.63
14	13.00 pm - 14.00 pm	-4.69
15	14.00 pm - 15.00 pm	-4.54
16	15.00 pm - 16.00 pm	-4.71
17	16.00 pm - 17.00 pm	-4.69
18	17.00 pm - 18.00 pm	-4.58
19	18.00 pm - 19.00 pm	-4.52
20	19.00 pm - 20.00 pm	-4.32
21	20.00 pm - 21.00 pm	-4.25
22	21.00 pm - 22.00 pm	-4.2
23	22.00 pm - 23.00 pm	-4.3
24	23.00 pm - 0.00 am	-4.16
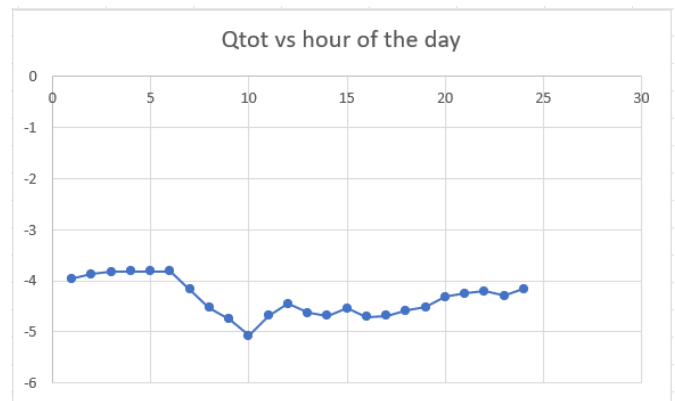
Figure 52 Plot with Q_{tot} hourly values vs hour of the day for building 10922 in Tartu.

Figure 51 Table with Q_{tot} hourly values for building 10922 in Tartu.

In addition to that, variation in the total heat loss coefficient is small (between complete model value and peak hour value there is a difference of around 9%) so, it could be produced by inaccuracies of the models and distribution of experimental data points. Another explanation could be given by the working hours of the SH plant. Despite it is in operation during all hours in winter period, of course there are moments when the power delivered by the system is lower, because there is less power need from users. So, the peak in Q_{tot} can be caused by the increase in power delivered by SH plant after a period where there was less necessity of energy. There is also a possible influence of DHW consumption that could lead to some inaccuracies.

The previous explanation could also give an interpretation of the smaller values (in absolute term) for Q_{tot} in night period. Values are lower, highlighting less dependence of consumption to outdoor temperatures for these hours. This could be explained by the fact that during night users sleep, so indoor temperature could be lower than day value. For this reason, space heating plant can provide less energy. In equations [7.2] and [7.3] there are reported the expression for Q_{tot} as written by Eriksson M. et al. in their work.

$$Q_{tot} = \frac{P_{DHN} + P_{ihg} + P_{solar} - P_{loss,DHW} - P_{DHW}}{T_{in} - T_{out}} = \frac{P_{tot}}{\Delta T} \quad [7.2]$$

In the previous law, Q_{tot} is the already mentioned specific total heat loss coefficient (in kW/°C), P_{DHN} is the delivered space heating power from the DHN, P_{ihg} is the power for internal heat gain, P_{solar} the power of solar heat gain, $P_{loss,DHW}$ the value of losses in DHW circuit, P_{DHW} power delivered for domestic hot water (DHW), T_{in} the indoor temperature and T_{out} the outdoor temperature. In case we are considering demand during night, of course solar heat gain disappears and so DHW, for the hypothesis explained in paragraph 4.9.1, regarding DHW loss calculation method. During the night, Q_{tot} has the expression reported in equation 7.3.

$$Q_{tot} = \frac{P_{DHN} + P_{ihg} - P_{loss,DHW}}{T_{in} - T_{out}} \quad [7.3]$$

Q_{tot} is determined from experimental hourly values of demand, which is a representation of the hourly global heat consumption and with outdoor temperature values, assuming that T_{in} is constant during the hours, with a possible reduction during night-time.

A critical issue must be specified regarding previous equations. They are mentioned in many studies that talk about ESC models and obtainable data with them. In addition to the already mentioned works by Eriksson M. et al., Milić V. et al., it was analysed also work by Rhodin P. et al. In these studies, a physical interpretation of the slope of the temperature dependent part of the ESC line (what here is called Q_{tot}) can be found, and it is the one that it was illustrated in these lines. However, parameters that appear in the equations [7.2] and [7.3] have a specific meaning and they cannot be confused with E_{SH} and E_{DHW} that it was found in the results of the splitting procedure.

To understand this difference, it is necessary a comparison between two different expressions of the ESC. The first one is the energy signature curve expression that it was used in chapter 4 for describing the energy necessary for space heating (equation [7.4] comes from a simple passage of the term $\min(f(x))$ from the second member to the first member of the equation [4.14]). The second one (equation [7.5]) is obtained by multiplying both members of equation [7.2] by difference of temperatures $T_{in} - T_{out}$.

$$E_{SH} + \min(f(x)) = f(x) = ESC(x) \quad [7.4]$$

$$Q_{tot} * (T_{in} - T_{out}) = P_{DHN} + P_{ihg} + P_{solar} - P_{loss,DHW} - P_{DHW} = ESC(x) \quad [7.5]$$

In the previous equations, $ESC(x)$ are the points of the energy signature curve. By comparing the two formulas there is a similarity between terms. As described in articles mentioned before, the sum of P_{DHW} and $P_{loss,DHW}$ is the so called "shifting coefficient" mentioned in paragraph 4.9. This fact is evident also by observing figure 53, where it was reported an example of an ESC curve taken from the articles, where they are indicated with terms P_{dhw} and P_{dhwc}. To be precise, in the mentioned articles there is the hypothesis of constant DHW consumption during the year. It is a simplification choice that it is not done in this analysis, otherwise it would be useless to perform splitting between SH and DHW consumption because results would be obvious. As a first approximation, and for exposition clearness of the parallelism, it is possible to assume that the sum of P_{DHW} and $P_{loss,DHW}$ is equal to the shifting coefficient for the entire year.
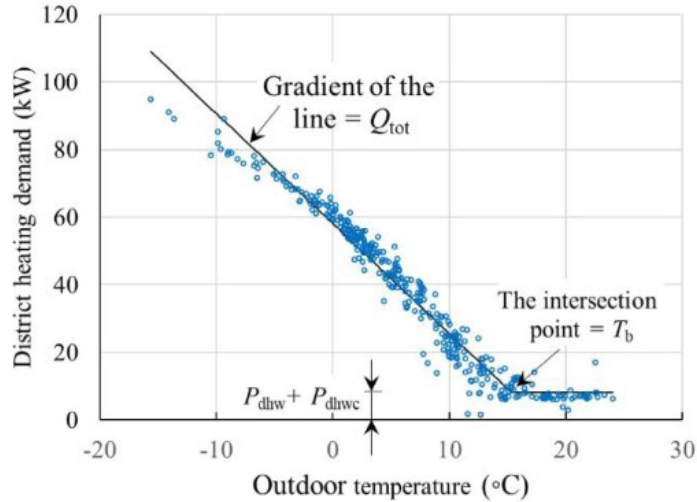
Figure 53 An example of ESC curve taken from Eriksson et al. article, useful to check similarities in terms of names and parameters with the method implemented in this analysis.

So, as previously said, the sum of these two terms is equal to $\min(f(x))$, the so-called shifting coefficient, as highlighted in equation [7.6]. Thanks to this consideration, it is possible to say that what in the model is called E_{SH}, in this discussion it represents the sum of $P_{SH} + P_{ihg} + P_{solar}$, as reported in equation [7.7]. Of course, T_b is the change point temperature (CPT). Thanks to this parallelism, it is possible to give a more detailed physical interpretation to the results of the model.

$$\min(f(x)) = P_{loss,DHW} + P_{DHW} \quad [7.6]$$

$$E_{SH} = P_{DHN} + P_{ihg} + P_{solar} \quad [7.7]$$

Equation [7.6] is in accordance with the equations used to build DHW demand, reported in [4.15]. However, the E_{SH} value reported in equation [7.7] is not the final SH consumption obtained from the model. It is the E_{SH} before the rebalancing phase and before the correction of negative values. In poor words, it represents point of the ESC shifted down. This fact would be true if, as happened in the mentioned research, experimental points used to create the building's model are referred to demand of the building. However, in this case, the situation is unluckier. In fact, experimental points are measured on the DHN side of the building's heat exchange. This means that we have measurements of directly P_{DHN}. In reality we cannot say anything about real demand of the building, because all analysis is conducted using these data. So, equation [7.7] is true only if demand is known, but this is not the case.

From this consideration it is possible to conclude that information about solar heat gain (P_{solar}) or internal heat gain (P_{ihg}) are impossible to obtain from datasets. They could be estimated, but there is no necessity now.

Because we are considering night period, the considerations are done taking into account equation [7.3]. Solar heat gain disappears, and internal heat gain is still present. However, no differences should be noticed in our model, because, for the reasons illustrated before, none of them are considered.

It is possible to notice that, if there is a change in the indoor temperature (imagine that during night indoor temperature is lower), the temperature difference at denominator becomes smaller. If this happens, also heat delivered by SH plant becomes lower. Assuming that internal heat gains and losses in the water circuit do not change, also numerator tends to decrease. If this happens, Q_{tot} should remain almost constant, as said before. From results reported in figure 51, it is possible to notice that the difference between values of Q_{tot} during the night hours and the value for the complete model there is a difference about 15%.

This difference could be explained by approximations due to the statistical approach of the model or by the decrease in thermal losses caused by a difference in indoor temperature (and a decrease in difference between indoor and outdoor temperature, too). To see if T_{in} plays a role in the determination of Q_{tot} as slope of the temperature dependent part of the ESC it was tried to plot demand vs difference between indoor and outdoor temperature, instead of only outdoor temperature. It was performed an interpolation of experimental points of global heat demand, but the same results were obtained, so there is no influence of T_{in} in this particular operation. In addition to that, a difference in the order of 15% is still acceptable and perfectly coherent with also similar cases present in the studies of Rohdin P. and Milic V. Without more information it is difficult to give a more accurate interpretation of this fact.

The information obtained from this discussion and from the ESC model of the building is extremely useful for many purposes. Compared to relying solely on specific energy use in kWh/m^2, a typical parameter when there is the analysis of a building, this model provides additional valuable information. This includes the building's total loss term, Q_{tot}, as well as change point temperature and heat for DHW, losses in water circuit and SH heat load consumption.

As a further consideration about results, it is important to spend some words for the DHW heat load consumption during the year. One of the results of the analysis is, as already mentioned, the vector that contains hourly DHW heat consumption. To check if results obtained are feasible, it was decided to sum all these values to obtain the total heat consumption for domestic hot water in one year of operation. it was decided, also, to calculate the heat consumption pro-capita for DHW for one year. It was used the equation reported in [7.8].

$$E_{DHW,std} = m_{DHW} * cp_{water} * (T_{user} - T_{aque}) * 365 \quad [7.8]$$

In the previous law, $E_{DHW,std}$ is the standard consumption pro-capita for DHW, m_{DHW} is the standard value for mass consumed per person per day (in case of residential building) and it was assumed it equals to 40 l/(day * person), cp_{water} is the specific heat of water, 4.186 kJ/(kg * K), T_{user} is the temperature at which domestic hot water must be available, it was assumed 46°C (this data was taken from the dataset since it is available, but this is a standard data) and T_{aque} is the temperature at which water is available from the aqueduct (it was assumed 3°C because Tartu has a cold climate).

With that law it was calculated the DHW consumption in one-year pro-capita and this standard value it is equal to 730 kWh/person. With this value and the total DHW consumption of the building (E_{DHW}) in one year it was performed a division to calculate the number of people, using equation [7.9].

$$N_{people} = \frac{E_{DHW}}{E_{DHW,std}} \qquad [7.9]$$

For this specific building it results from calculation that there are around 220 people living there. This is in accordance with the considerations done before, where it was said that this should be a big residential building. However, it is a pretty high number of people, assuming 3 people living in an apartment, there would be more than 70 apartments. It could be, of course. Proceeding with the analysis of results, it will be shown that with the correction procedure of DHW demand in winter described in paragraph 4.9, the number of people decreases up to 190 (so around 60 apartments).
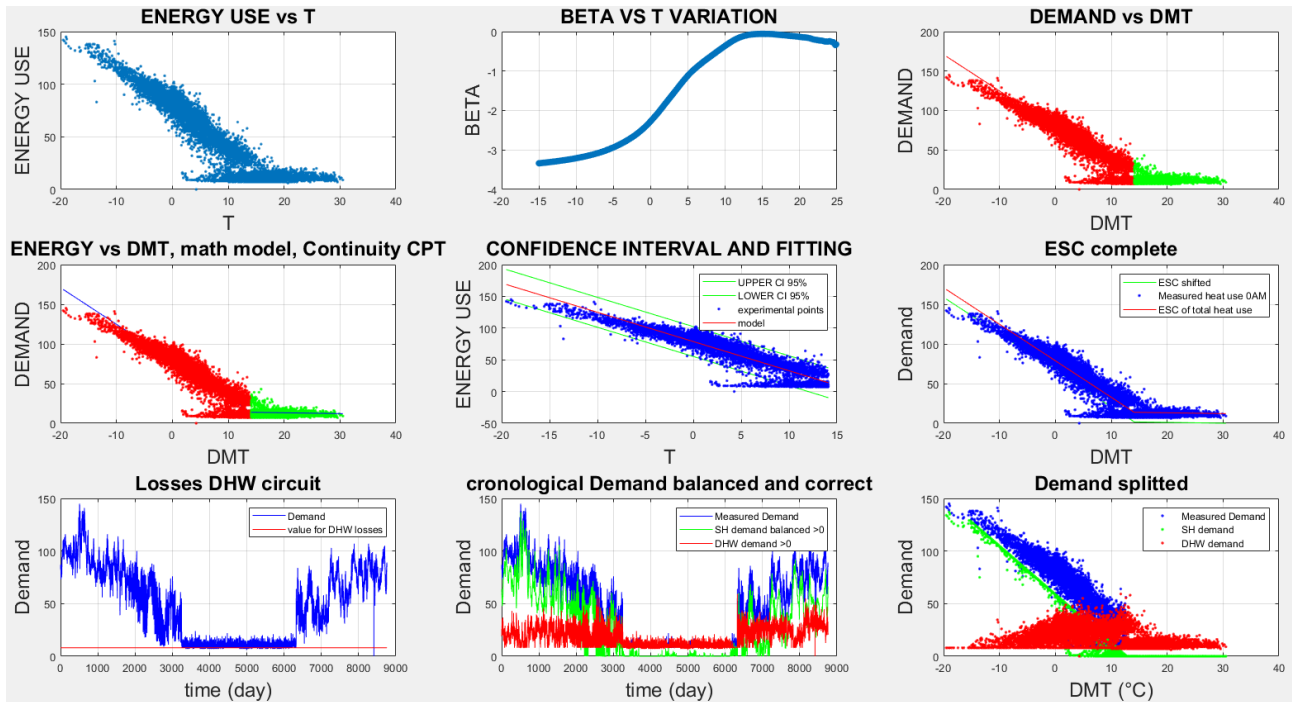
Figure 54 Results of the analysis of the 8760-elements vector of demand for building 10922 in Tartu.

Regarding the other hours of the day, no particular differences are present. Profiles are quite similar one to each other and any eventual points of interest are described in the previous lines. The next figure shown in this site is the figure number 27 in the output of the code. In this paragraph, it is reported in figure 54. It represents the same identical plots and data shown for hourly values, but this time the vector considered is the 8760-elements vector which contains all hours of the year. Meaning of every element is the same as described for the figure 50. As can be seen, there are much more points, and the reading is more complicated. However, the presence of a higher number of experimental values contributes to reduce variations and oscillations in the demand profile. By considering plot number 8, in fact, DHW demand (the red line) seems to be more similar between winter and summer than before, even if high differences in terms of oscillation of profiles are still present.

Talking about some peculiar things about this analysis, confidence interval plot shows that almost every experimental point is inside the 95% CI area. There are only some exceptions, due to the normal times where there a request for a particular energy. An important zone of points outside the CI is present in the lower-right area of the plot. However, these points represent hours with DHW consumption only, without space heating, even if there are low

temperatures. This is a perfectly normal event: sometimes there is no necessity for SH even if temperatures are low, so only DHW or losses in the DHW circuit are present.
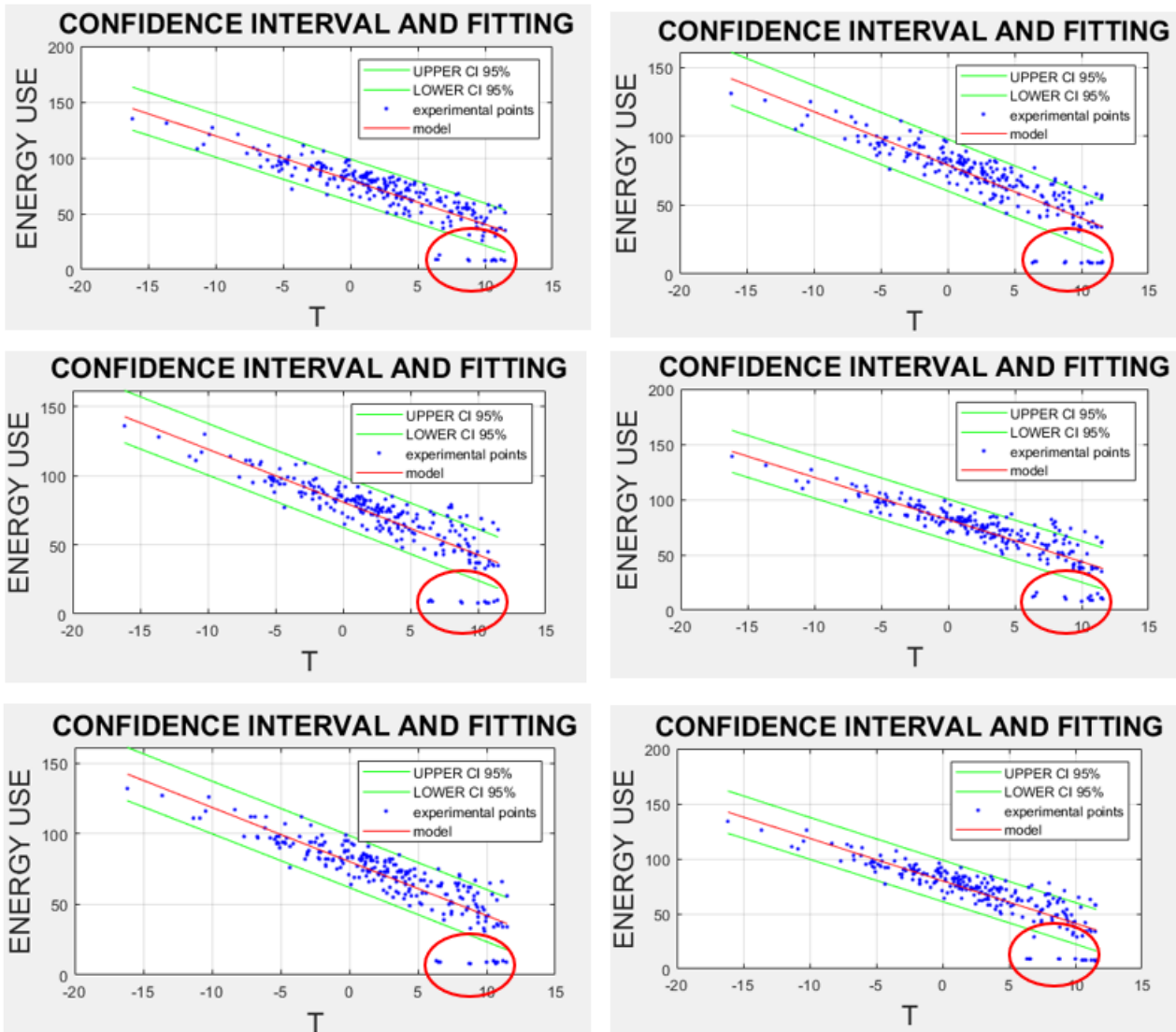
Figure 55 Hours between 0.00 am and 6.00 am (from top left). With red circles the hours without SH are highlighted.

From the analysis of singular hourly vectors, almost all these points of lower demand are present in night hours (between 21.00 pm and 5.00 am). In figure 55 it was reported CI plots of the mentioned hours of the building in analysis to show this fact. This characteristic of demand is consistent with the previous interpretation of the reason why these points are outside from the confidence intervals. During night, in fact, users sleep, and space heating becomes less important, when temperatures are not so low. Of course, if the outdoor temperature is many degrees below zero, thermal losses is truly relevant and space heating is necessary to keep good indoor comfort condition. These points of zero demand could also be explained by the presence of low outdoor temperatures that could occur in some days when SH plant is already turned off.

Let us proceed with other results of the code. The next image from outputs is the number 25, reported here in the figure 56. Here there are the boxplots of energy consumption divided into workdays, Saturdays, Sundays and total. For this particular building it can be said that no big differences are present between weekends and weekdays. Looking at red lines (median values for each hour of the day) it is clear that during night there is the highest consumption. This fact is consistent with the previous considerations about no space heating at night in some periods and it will be explained why.
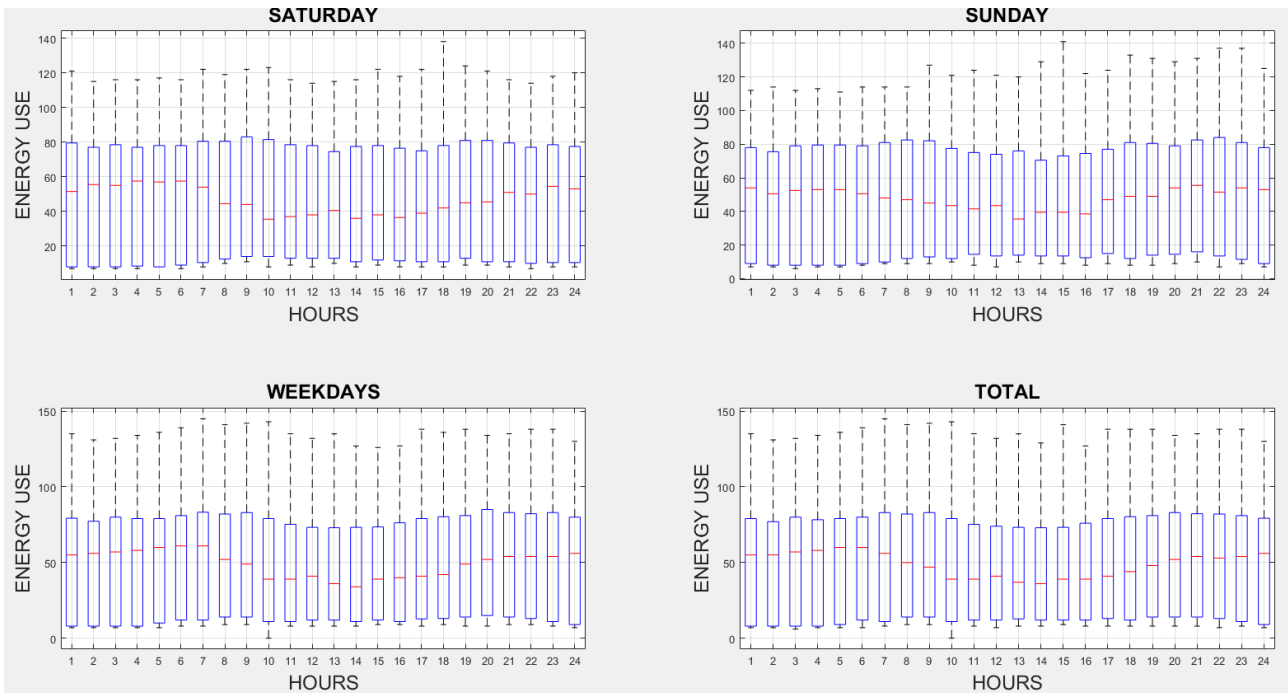
Figure 56 Boxplots of yearly energy consumption divided into saturdays, sundays, work-days and total of building 10922 in Tartu. Image 25 of the output of the code.

It is evident form plots reported in figure 55 that days without space heating demand are low and they are all situated in the higher temperatures zone. So, they are too low to have a relevant effect in median consumption. In the boxplot of figure 56 they are represented by the lowest boundary of the blue box and not by the median red line. This is the reason why they have small influence, but why consumption is higher during night?

As said before, due to the way these data are collected, there is no consideration of both internal heat gain and solar gain. As a consequence of this fact, when temperatures are low, and this happens for most of the time in winter season, higher heat energy is necessary also during night. This energy is higher than the one requested during the day because at night there is no effect of solar gain. On the other hand, during the day the solar gain plays a key role giving free extra energy to the indoor environment. It must be considered, also, that during the day outdoor temperatures are higher than night, so thermal losses are lower, with a consequence of lower demand. In simpler words, during the day the sun provides extra energy and makes the outdoor temperature higher. So, the thermal energy required from the building to the DHN (the parameter measured and collected in the dataset) is lower than the night-time demand.

During workdays, the south-west plot in figure 56, no particular peaks are noticed during the day. There is a local maximum at index 12, so between 12.00 am and 13.00 pm, maybe for the lunch time. In this period, assuming in a consistent way that this is a residential building,

some people return from work to eat at home and some extra energy is required for the dish washing. After that local peak demand median values return lower until 19.00 pm, when they start increasing. This phenomenon could be explained by a combination of sun disappearance (no more solar heat gain), users return from work (they could increase setpoint temperature in the indoor environment), high probability of showers and dish washing.

Regarding the Sunday plot, it can be seen that during morning consumption is, in general smoother compared to workdays, ant the peak between 7.00 am and 8.00 am is lower, probably because during Sunday people sleep more and there is less necessity of hot indoor environment from the early morning. The heating of indoor space is, in this way, distributed between more hours and the solar heat gain plays a key role also in this morning load. Saturday is remarkably similar to Sunday, without relevant differences.
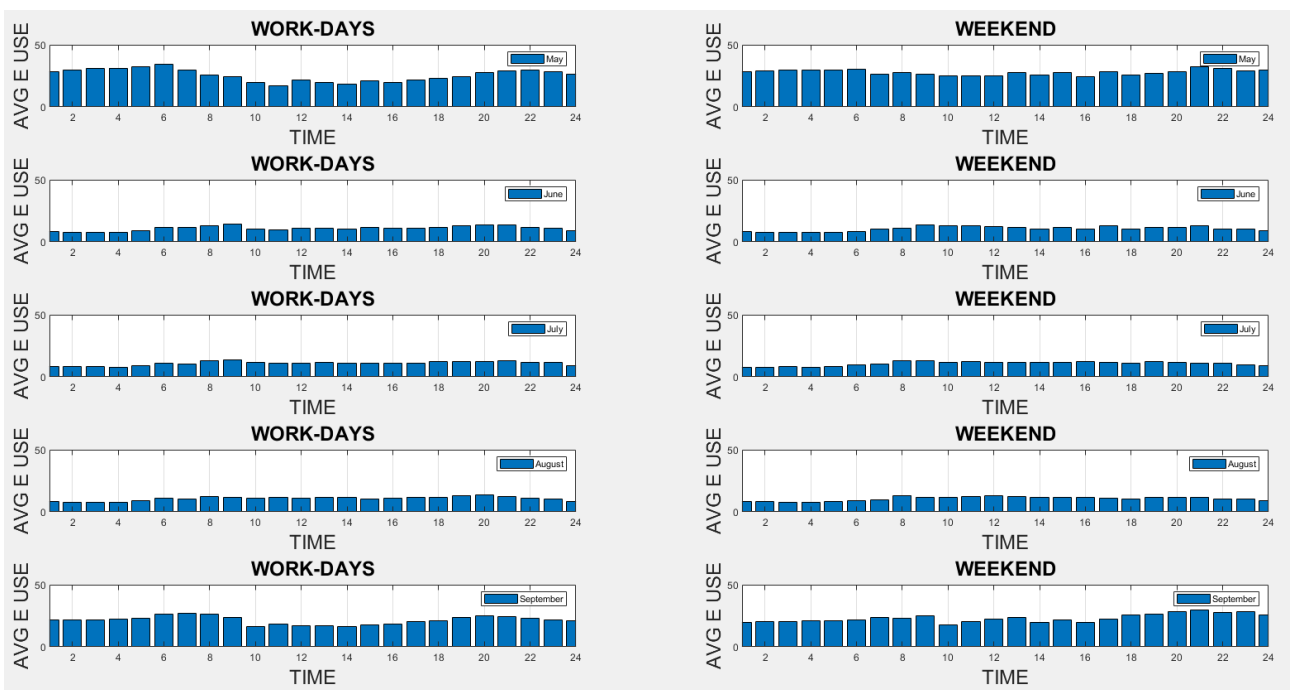
Figure 57 Global consumption in summer months (from May to September) divided in workdays and weekends for building 10922. It is the image 28 among code's outputs.

The next image taken from the outputs is about summer consumption. In figure 57 there is the image 28 from the output of the code. There are 10 plots. On the left there are workdays and on the right weekends. Each line represents a different month. It is clear that in May and in September heat consumption is much higher than the other summer months (June, July, August). This happens because, having Tartu a cold climate, in May and in September space heating is still necessary. Just to have an idea of the impact of this type of consumption we can look at the difference between May and June. During central hour of the day columns are almost two times higher in May than in June and this difference becomes much larger during night hours. In this period, in fact, May's columns are more than 3,5 times higher than June's columns. This fact is consistent with the previous considerations about space heating consumption at night.

By considering a single month it is much more evident the difference between weekend and workdays, confirming previously information. In May, for example, it is evident that during

morning the profile is smoother than June. During summer months the difference in profile between workdays and weekends is much less evident and to study them in a more accurate way it is necessary a deeper analysis. With this figure, in fact, outputs of the first part of the analysis are out.

In the figure below, number 94, it was reported results for the analysis of correlation between TMA and energy consumption. The meaning of this exam is reported in paragraph 4.5. for this building it is possible to say that the thermal inertia is quite high. Form this analysis it emerged that the highest correlation is in correspondence of the 28th previous hour. Among the analysed buildings this is the one with the highest inertia. This fact is probably due to the large dimensions of the structure.
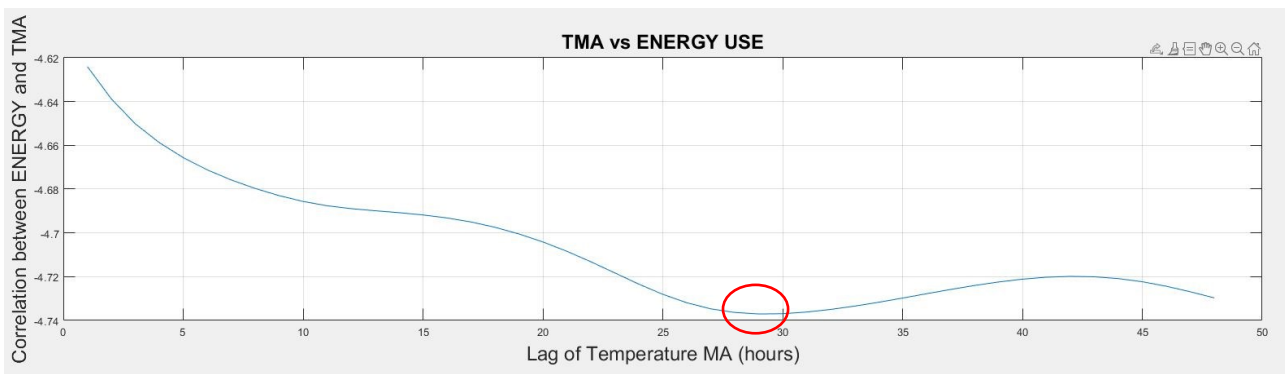
Figure 94 Effect of temperature lag on energy use of building 10922, Tartu.

From this point, all figures proposed are referred to the detailed analysis of DHW profile during summer and the last part of the analysis about the operation of smoothing of the DHW profile in winter.

As a first result of this part of the analysis, it is reported the match matrix. This is the results of all comparisons between different days of the weekend performed with Student's t-test and Fisher criterion. The matrix is reported in figure 58. It can be seen that values are not black or white. In fact, the similarities between different days are not so clear. This is due to the small number of weeks considered, the uncertainty of methods and the fact that during summer it is more likely that building has some extraordinary periods of work (because users are on holiday, for example). All these reasons and their implications were discussed in paragraph 4.9.1.

	1	2	3	4	5	6	7
1	100	0	0	0	0	0	0
2	76.9231	100	0	0	0	0	0
3	84.6154	76.9231	100	0	0	0	0
4	84.6154	92.3077	76.9231	100	0	0	0
5	69.2308	69.2308	76.9231	100	100	0	0
6	76.9231	69.2308	92.3077	84.6154	76.9231	100	0
7	92.3077	100	84.6154	92.3077	92.3077	84.6154	100

Figure 58 Match matrix for the building 10922 in Tartu.

The day that has higher difference with the others is Friday (number 5). So, it was decided to insert it in group 1 and the other days in group2. For this building, all following results are obtained considering the groups composed like that. In figure 59 there is the outputs of the

paragraphs 4.9.1 and 4.9.2. on the left there are all plots referred to group 1 days and on the right the ones referred to group 2. The first plots are the average profile during summer months (from June to August). Some minor differences can be noticed, but more or less the pattern look remarkably similar. This is because in summer there is only DHW consumption, so the differences that occur between workdays and weekends described before do not happen.

Proceeding with plots, we can find Tcal plots. They are parameters useful for method implementation and to separate hours of the day in minimum, medium and peak consumption period, but they are not linked to any physical meaning.

Last plots are more interesting. First, it can be seen that hours with minimum consumption of DHW are belonging to night period. So, it was correct to assume that during night there is no consumption of DHW from users because most of them are sleeping. Regarding this point, it is also clear that there are no particular differences between considering minimum value of consumption or the mean value among these hours. This means that both method one and method two described in paragraph 4.9.1 for the evaluation of losses in DHW circuit are correct.
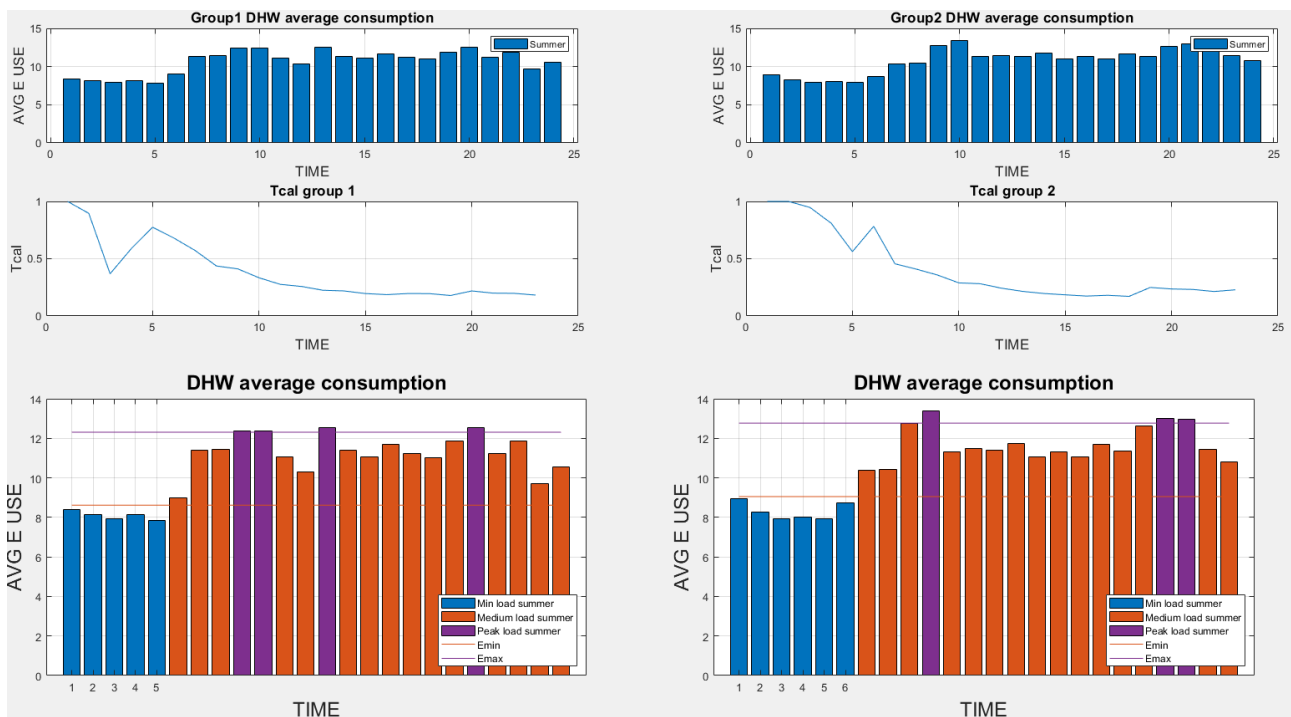
Figure 59 Results from the splitting of hours into minimum, medium and peak periods in building 10922, Tartu.

Going further, hours with medium consumption are the most diffused during the average day profile. This is reasonable. Another relevant fact is that peak consumption hours are few, only 3,4 hours in an entire day. This fact is reasonable too. Regarding them, it can be noticed that they are especially during the morning (between 10.00 am and 11.00 am) and during dinner time (between 20.00 pm and 22.00 pm). First peaks can be explained by morning showers or early lunch, while evening peaks are easily explained by showers of users and hot water necessity for dish washing, as discussed before.

Another interesting fact is that there is small gap between peak consumption and minimum consumption. Demand goes from approximately 8 kWh for minimum demand and 13 kWh

for the maximum peak. One thing to say about that is that, as highlighted before, DHW demand is much smaller than SH demand. In addition to that, an interpretation of this smooth profile during the day, without peak much higher than minimum, is that with a lot of users and apartments, especially if they have different routines, jobs and habits, there is a flattening of demand profile. Abrupt demand from one user cannot be seen in average consumption because we are considering data of the entire building. This fact makes the hypothesis that this is a large residential building stronger.

Talking about numbers, E_{min} is 8.62 kWh for the group 1 and 9.05 kWh for the group 2, while E_{max} is 12.31 kWh for the group 1 and 12.77 kWh for the group 2. As said before, and now confirmed by numbers too, there is a truly minor difference between group 1 and group 2 days. It is important to underline the fact that this is just a case, of course differences between workdays and weekends depend a lot on the intended use of the building.

The next figure, the number 60, shows standard profile for summer and for winter built with method illustrated in paragraph 4.9.3 and 4.9.4. There are not so many things to say about them. They are useful for the operation of correction of DHW profile in winter and all the considerations done in the previous lines are valid for them too.

From the differences in the columns' height, it is possible to say that the difference between average DHW in summer and in winter is about 50%. For winter season, E_{min} is 13.15 kWh for the group 1 and 13.71 kWh for the group 2, while E_{max} is 18.77 kWh for the group 1 and 19.33 kWh for the group 2. Difference is quite big, but it depends on mean values for DHW in winter, so it depends on the model used for splitting operation. This is the best result obtainable with this approach. However, after correction operation, both total energy consumption for DHW and oscillations in consumption profile are reduced.
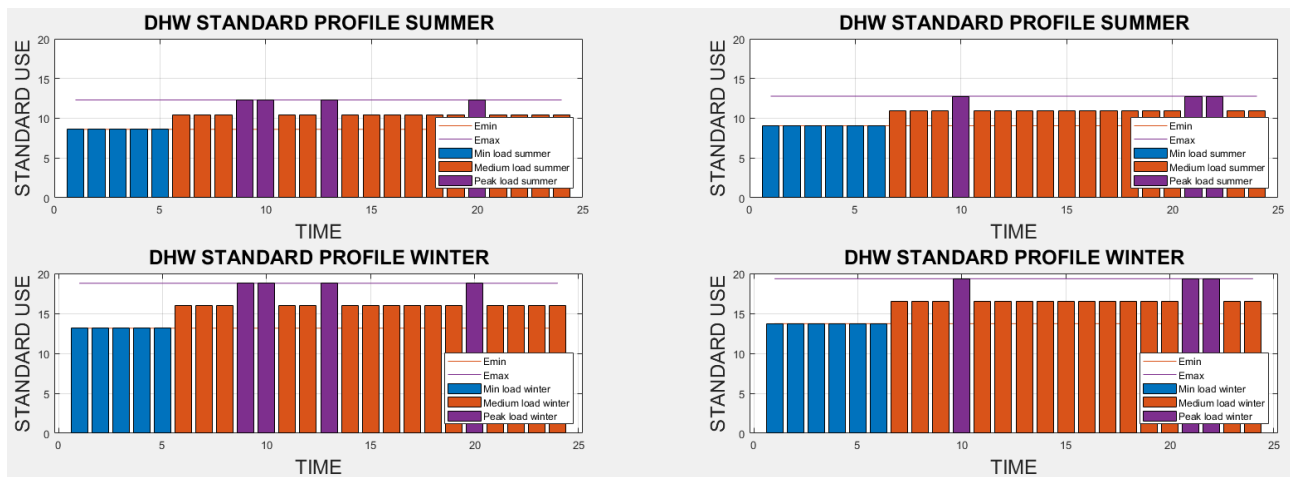
Figure 60 Standard profiles in winter and in summer for days of group 1 (left) and group 2(right) in building 10922, Tartu.

In the last image from code's outputs there are values corrected of DHW consumption, reported here in figure 61. Starting from the upper-left plot there are Mondays, Tuesdays, Wednesdays, and so on. Red lines represent the limit of summer, before and after that there are winter months of 2019 (year taken in analysis). Regarding information provided by these plots, comments done before are still valid. In this site the more regularity of DHW demand in highlighted. It is clear that oscillations of demand during winter period are much less and with minor intensity. Mean values of DHW in winter is still quite to high compared to mean

value in summertime, but this is the results of the method. By applying techniques to change this value there would be a complete change in physical concept at the basis of the method.

Of course, if DHW demand is modified for making it more regular and more like the summer demand, SH heat load consumption changes too. In correspondence of points where DHW measured value has a substantial difference with winter standard profile and, as a consequence, its value is corrected, the demand subtracted by DHW is added to SH. It happens that SH points are no more distributed perfectly along a line but follow a freer distribution. However, the general path is still a line, in accordance with hypothesis of ESC model. This fact can be noticed in the last plot of figure 61.

As said some lines ago, with modified value of DHW consumption there is a reduction of the total energy destinated to this purpose. Therefore, by applying equation [7.8] and [7.9] it is possible to calculate new number of people. With the parameters of correction implemented in the method and described in paragraph 4.9.5, it results that in the building there should be 190 people. This value is consistent with the conclusion that the building in analysis is a large residential structure, too. As explained in future work paragraph, chapter 8, if just a little more information would be available, accuracy of the model could increase a lot. With
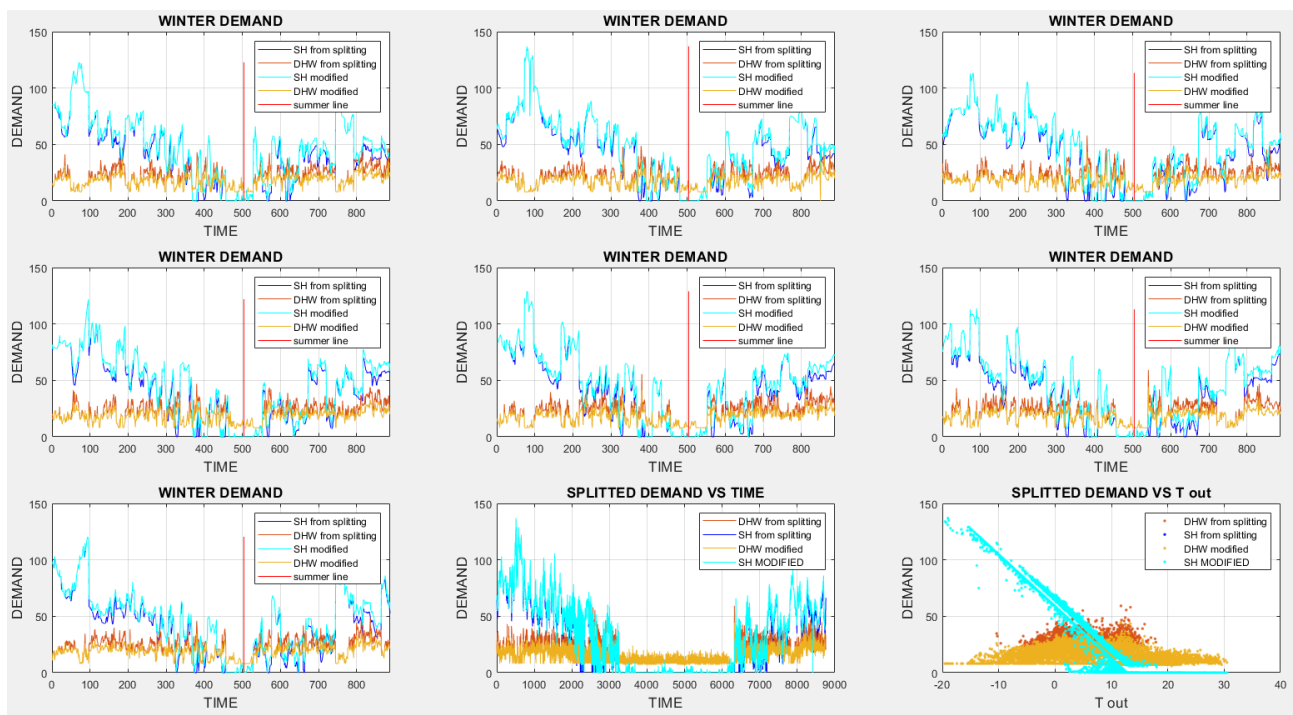
Figure 61 Representations of DHW and SH before and after correction of DHW profiles. From top left there are: Mondays, Tuesdays, Wednesdays, Thursdays, Fridays, Saturdays, and Sundays. Last two plots are about 8760-values hourly vector of year consumption.

this method, by knowing just the number of occupants it would be possible to calculate the standard consumption of the building and applying the correction procedure to make the DHW total consumption obtained from splitting plus correction operations equal to standard consumption calculated with equation [7.8]. Of course, standard consumption is not equal to real consumption of the building, but it remains a big step forward compared to the actual model.

With this figure outputs of the implemented tools are out. A lot of information can be extrapolated starting from just hourly total consumption for one year. All these

considerations can be used for many purposes: from improving DHN operation to study and apply retrofitting interventions. Detailed explanation of this analysis' potential is provided in conclusions' chapter (the number 9). The following buildings from Tartu' databased are not subjected to this detailed analysis of results because most considerations are true also for them. In the following paragraphs it was decided to present just the most relevant cases, if they are present, and it was wanted to highlight just their peculiarity.

7.2 Building "10051" in Tartu.

The next building from Tartu's dataset it was wanted to show is building 10051. The decision fell on it because in summer it has neither DHW nor SH consumption. By using this building, it was proved that this analysis provides satisfactory results even if it is impossible to obtain information from DHW consumption in summer, because it is absent. However, detailed DHW summer consumption profile analysis and correction of winter DHW profile as described in paragraph 4.9 cannot be performed. For this reason, it will be shown here only results from the first part of the analysis. It is relevant to anticipate that, due to the absence of DHW consumption in summer, and the impossibility to correct DHW profile in winter, so it presents some oscillations.
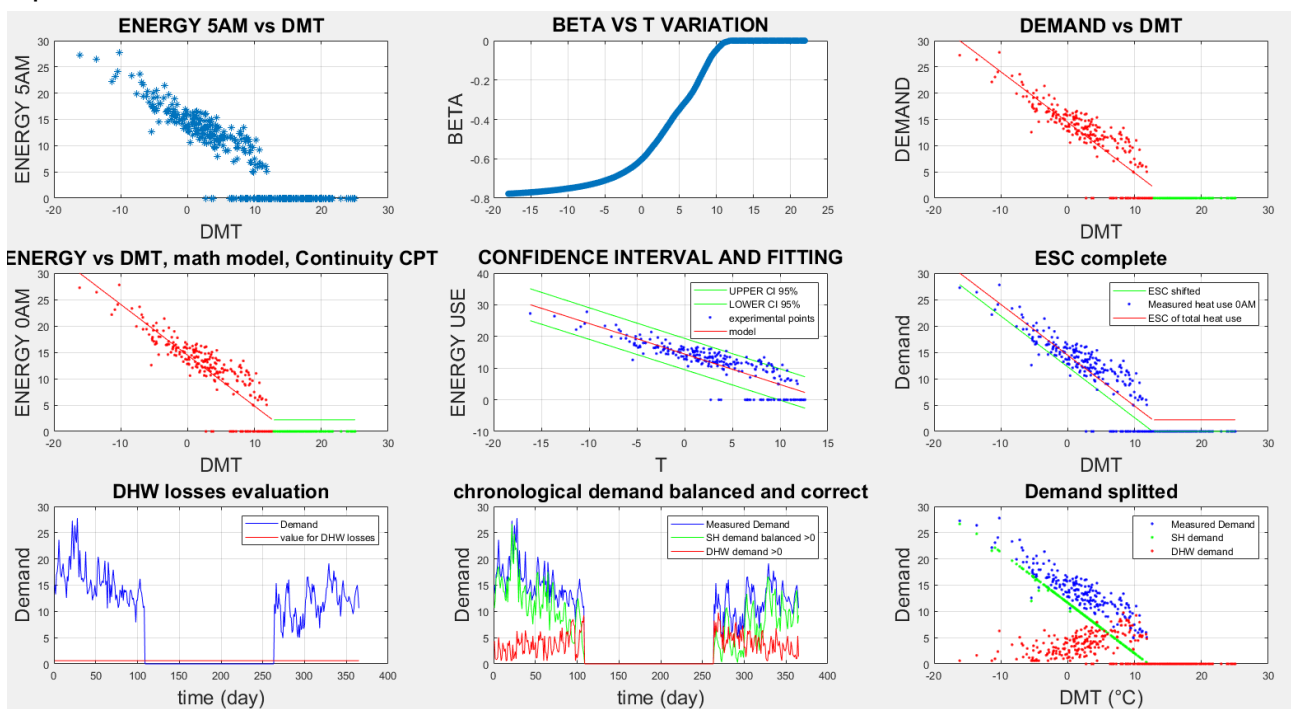
Figure 62 Figure from the output of the building 10051's analysis. Hour between 5.00 am and 6.00 am.

There are not so many differences among hourly vectors. It is reported here, in figures 62 and 63, outputs for both a night hour and a day hour.

Meanings of plots are described in the previous paragraph, 7.1, so it will not be repeated them here. In the second plot it is possible to notice that it is quite easy to determine CPT by observing variation of beta parameter (β). Its constant behaviour, and equal to 0, after CPT is due to the fact that neither SH nor DHW is present in summer period, this means that when SH disappears also DHW is absent. Because both are null for every outdoor temperature in this zone, their interpolation gives always zero as parameter β (see equation

4.6). Because of this fact, the temperature independent part of the ESC is always a horizontal line.

Talking about confidence intervals, even in this case they are applied only on the temperature dependent part of the demand. The only points outside the confidence area are the zero-consumption hours. The explanation for them is the same given in the previous analysis in 7.1. Even in this case, and ad it happens for every building in Tartu's database, the original dataset was already pre-processed.
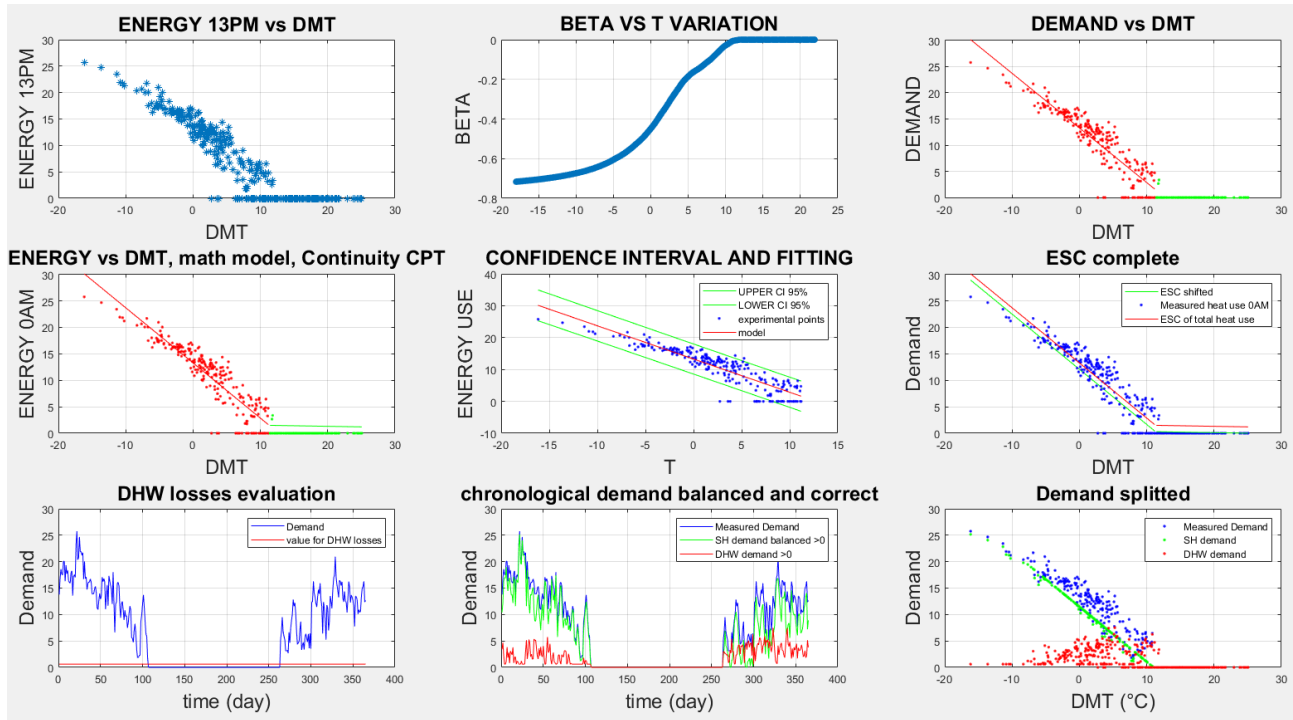
Figure 63 Figure from the output of the building 10051's analysis. Hour between 13.00 pm and 14.00 pm.

This building was already studied in the article of Lumbreras M, Garay-Martinez R. et al., together with other buildings of this dataset. However, for this particular building also the intended use is provided: it is an apartment building. From the maximum value of demand (around 30 kWh) it is possible to say that this is smaller than the previous building (10922). Nevertheless, number of apartments in the structure is quite high because no relevant peaks of demand is present in some hours. As explained before, in a single house or in a group with a few numbers of apartments, peaks, especially due to DHW, are much more evident. With many apartments, peak caused by a single user are less evident.

Another reason consumption seems to be so regular is that DHW load is not provided by DHN, so it is not measured by meters. With Tartu's dataset it is possible to understand if DHW is provided by DHN or not by analysing data from meters. As described in paragraph 6.1, there is one sensor which measures the temperature of hot water available to users. By looking at data for building 10051 it can be noticed that this temperature is always 0, for every moment of the year. The conclusion is that no water is circulating in this pipe, meaning that DHW is provided in a separate way (for example a gas boiler or solar collectors).

By taking into account this fact, it is clear that the operation of splitting between DHW and SH is meaningless. However, other conclusions discussed in this paragraph are valid and demonstrate the fact that the analysis proposed can give useful considerations even if some

of the tools cannot be applied (in this case the method for SH and DHW splitting and the detailed analysis of DHW profile).

By analysing other hours of the day, average consumption during the winter season is higher during night than during day. This, as happened for the previous building, can be imputed to the absence of solar heat gain during night hours and the lower outdoor temperature. Thermal losses are higher and extra energy is not available, so the energy provided by the DHN must be higher.

Another important parameter obtainable from the analysis is the already mentioned Q_{tot}. In the previous paragraph a detailed explanation of its meaning is provided. For this building, $Q_{tot} = -0.99 \, kWh/°C$. Coefficient of global thermal losses is much lower than the previous building. Because this value represents the slope of the interpolation line on the temperature dependent part of the ESC model for the building, a different value for Q_{tot} means that the two lines are not parallel. If they are not parallel, it means that one the demand of one building is more dependent to outside temperature than the other.— This is true, but it is important to keep in mind that we are not considering specific values. This means that Q_{tot} is extremely related to the dimensions of the building.

To have the possibility to make a comparison between different buildings, it is necessary to establish a criterion for creating specific values of demand. The most useful one would be dividing hourly consumptions by the floor area of the building. This criterion would give us also the possibility to compare their losses coefficient and consumption with standard value reported in literature. Unluckily floor area information is not available. Another possibility could be using number of occupants or number of apartments. This would be nice, but for some building this value cannot be calculated because it is obtained with equation 7.9, but not always DHW load is available (this is the case of building 10051). To have the chance to make a sort of comparison it was decided to create specific profiles of demand by dividing every measured value by the maximum value of demand. By applying this criterion, it emerged that building 10922 and 10051 have a remarkably similar response to variation of outdoor temperature in terms of heat consumption and they present an almost identical specific ESC in the temperature dependent part of the curve.

Regarding at Q_{tot}, it is possible to create also for this building a table like the one presented for building 10922 in figure 51. In this case, values for Q_{tot} are much more similar than building 10922. This is surely due to the smaller dimension of the building, which could lead to a more stable value. In figure 64 and 65 it was reported the table that contains Q_{tot} values per each hour of the day and their representation in a plot with the same axis as figure 52, to make a fair comparison.

Even in this case, but this is a general trend for this type of building, higher values of Q_{tot} are present during day hours. Even in this case difference between Q_{tot} values are in the order of 10%, as happened for building 10922. All possible explanations for this phenomenon are presented in the previous paragraph (7.1) and it would be just a repetition to present them also in this site.

As said before, last two plots in figures 62 and 63 are useless for this building. Because there is no energy provided for DHW purpose by DHN, splitting demand is a meaningless operation. As a consequence of that, DHW part of load does not exist in reality, so it must be not considered. However, this fact offers the possibility to do some considerations about the model for SH.

N hour	Hours of the day	Qtot
1	0.00 am - 1.00 am	-0.95
2	1.00 am - 2.00 am	-0.94
3	2.00 am - 3.00 am	-0.94
4	3.00 am - 4.00 am	-0.96
5	4.00 am - 5.00 am	-0.94
6	5.00 am - 6.00 am	-0.96
7	6.00 am - 7.00 am	-0.95
8	7.00 am - 8.00 am	-0.96
9	8.00 am - 9.00 am	-0.93
10	9.00 am - 10.00 am	-0.98
11	10.00 am - 11.00 am	-1.01
12	11.00 am - 12.00 am	-1
13	12.00 am - 13.00 pm	-1.04
14	13.00 pm - 14.00 pm	-1.05
15	14.00 pm - 15.00 pm	-1.07
16	15.00 pm - 16.00 pm	-1.05
17	16.00 pm - 17.00 pm	-1.01
18	17.00 pm - 18.00 pm	-1.04
19	18.00 pm - 19.00 pm	-1.01
20	19.00 pm - 20.00 pm	-0.94
21	20.00 pm - 21.00 pm	-1
22	21.00 pm - 22.00 pm	-0.99
23	22.00 pm - 23.00 pm	-0.98
24	23.00 pm - 0.00 am	-0.99

Figure 65 Plot with Q_{tot} hourly values vs hour of the day for building 10051 in Tartu.

Figure 64 Table with Q_{tot} hourly values for building 10051 in Tartu.

Because no DHW is present, all experimental points refer to SH consumption. Talking about CPT, it is evident that the mathematical approach proposed to determine it is correct even in this case. In fact, all points different from 0 are included in the temperature dependent part of the ESC. In addition to that the accuracy of representing SH consumption with the regression line can be controlled. The parameter that can give information about that is R^2. For this building, R^2 is between 0.72 and 0.82 for every analysed hour. For the complete model, with all hourly values of the entire year, R^2 for the temperature dependent part is 0.82. This fact confirms that using this type of model to estimate SH consumption, as it was done in the applied methodology, is correct.

In figure 66 it was reported analysis' output for the global hourly demand. No other specific considerations are necessary for this image. All interesting points were highlighted in the previous lines.
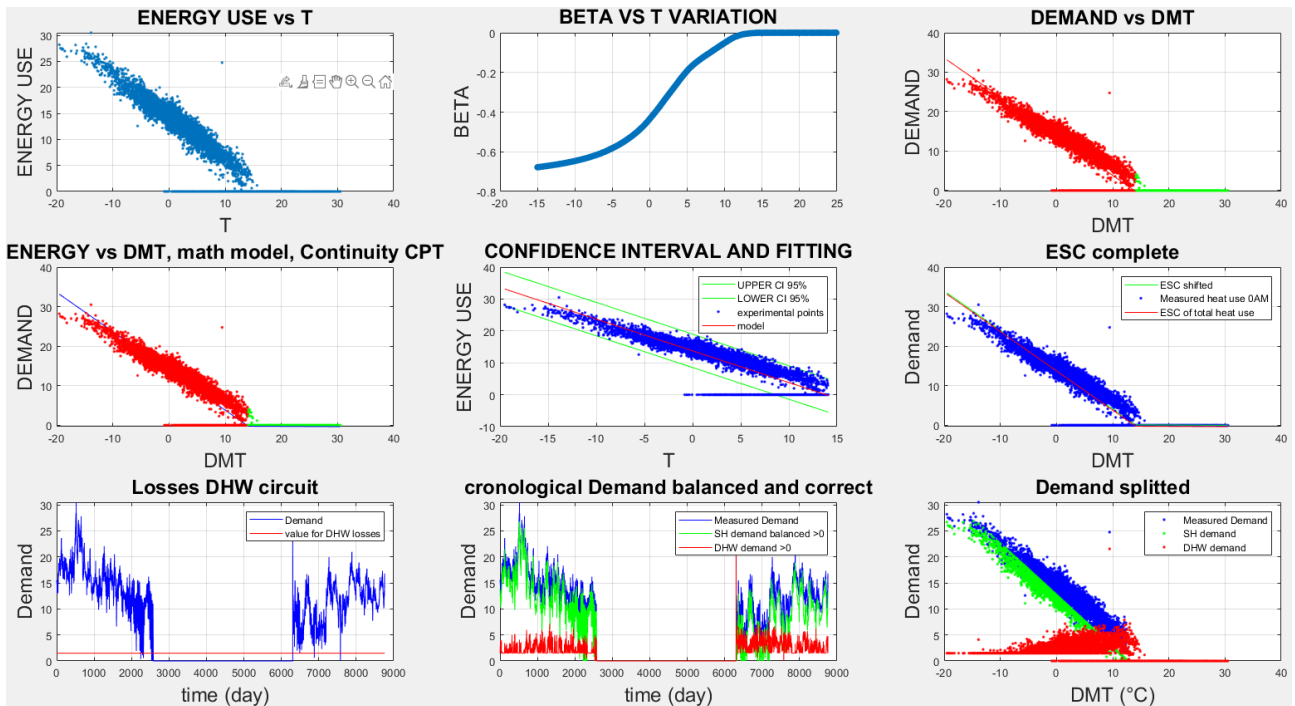
Figure 66 Results of the analysis of the 8760-elements vector of demand for building 10051 in Tartu.

By continuing the analysis, in figure 67 it was reported the boxplot of global energy consumption. Because DHW is absent, this is an extremely useful tool to understand the SH profile during winter in residential building in Tartu's zone. As can be seen, hypothesises about SH operation are still valid. During the day, the demand is lower because there are both the effect solar heat gain and higher temperatures. In weekends consumption during the morning hours is smoother and it remains high for more time because, an average, users sleep more than workdays and stay at home, not going to work. The minimum in SH consumption during the workdays happens between 15.00 pm and 16.00 pm. This fact is probably due to the presence of SH for the previous hours and the absence of users from home. Peaks are between 5.00 am and 8.00 am, probably because there is the necessity of provide energy to increase indoor temperature when users wake up. Unfortunately, no information about indoor setpoints is available, so it is possible just to make hypothesises based on average behaviour of users.

Regarding boxplot of DHW consumption in summer very few words can be said. There is no demand for DHW or SH so, as shown in figure 68, demand is always zero. It is important to notice that in this building, also in May there is no consumption. It is probably due to a different users' behaviour, or a different type of residence compared to building 10922, where also in May there was SH consumption, even if it was small.
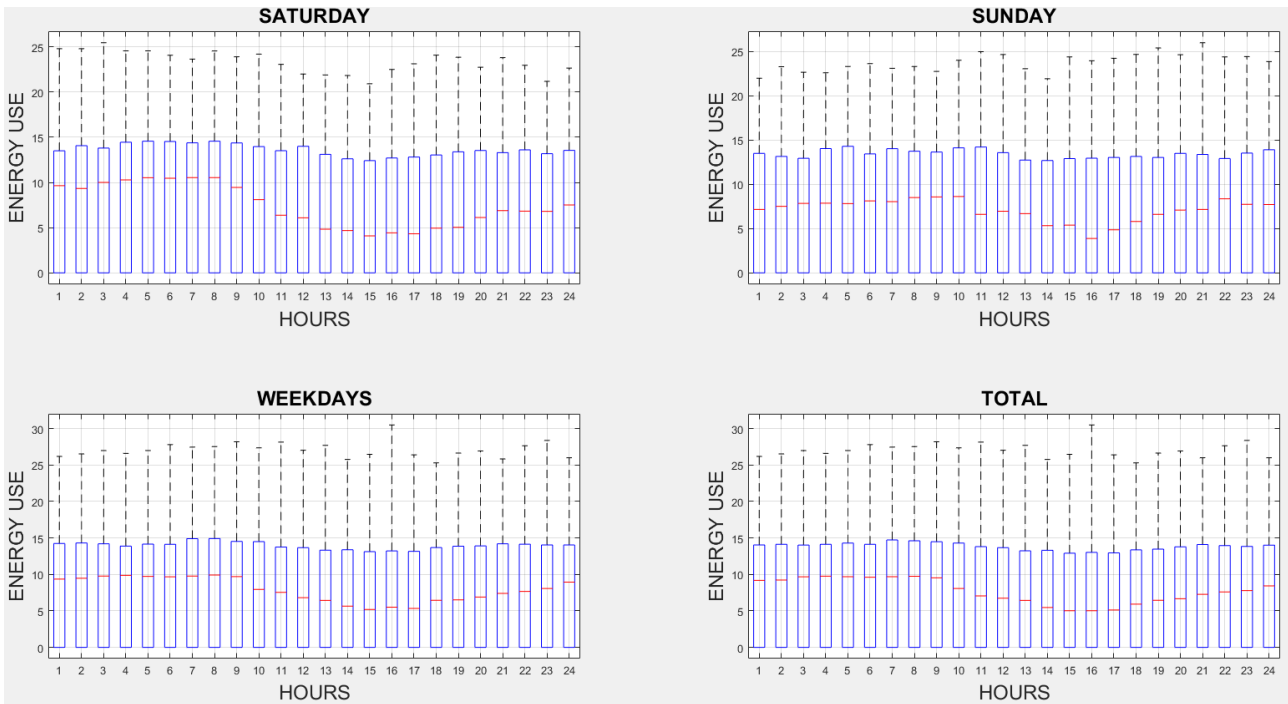
Figure 67 Boxplots of yearly energy consumption divided into saturdays, sundays, work-days and total of building 10051 in Tartu.
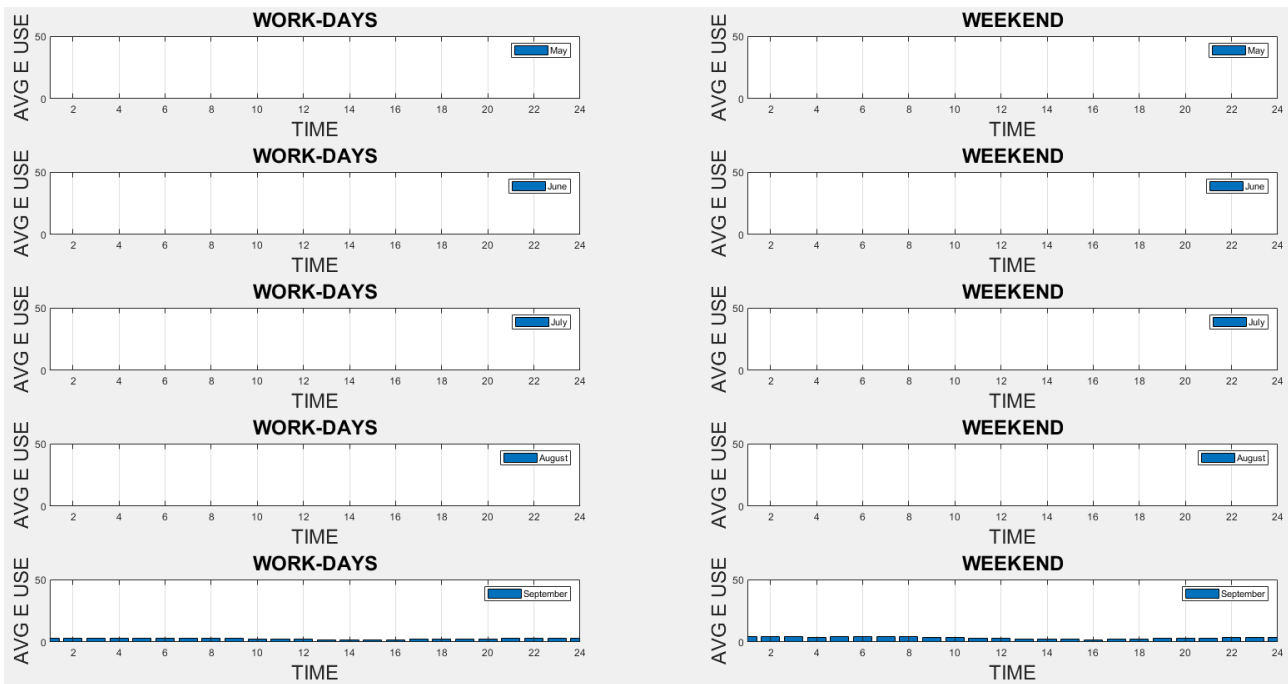
Figure 68 Global consumption in summer months (from May to September) divided in workdays and weekends for building 10051.

It can be seen some SH demand during September, but because these plots consider average demand during the entire month, it is possible that SH demand is present only in few days. In fact, by analysing September values it emerged that SH was turned on between 17.00 pm and 18.00 pm on the September 20th. So, the small demand present in the September's plot is caused by these last 10 days of the month. For this reason, this plot has no useful information.

107

In the figure below, number 95, it was reported results for the analysis of correlation between TMA and energy consumption. The meaning of this exam is reported in paragraph 4.5. for this building it is possible to say that the thermal inertia is quite high. However, the analysis in this case does not give satisfactory results, as it can be seen in the plot, where no consistent minimum points are evident.
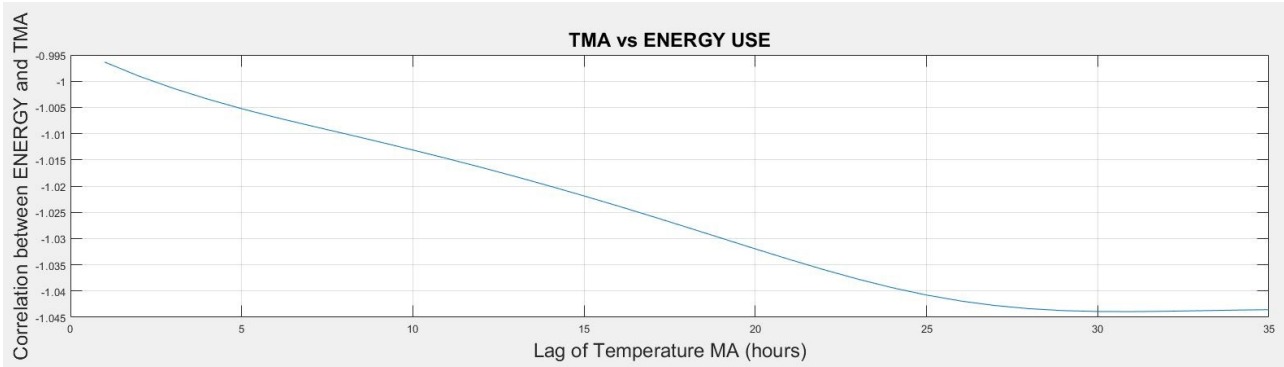
Figure 94 Effect of temperature lag on energy use of building 10051, Tartu.

As explained before, the remaining part of the analysis is useless and meaningless in this building. However, the fact that useful information can be obtained by applying these tools emerged from the presentation of results done in this paragraph. For this reason, it is possible to say that the fact of having no information during summertime does not impede the application of the method. Nevertheless, it is important to keep in mind to interpretate results to avoid completely wrong conclusions.

7.3 Building "10686" in Tartu.

After a standard example as building 10922 and a case where there is no DHW provided by DHN, here it was wanted to show results of the analysis of a building with a much more variable experimental dataset. Until this point, it was presented buildings with regular datasets, so that almost every experimental measurement stays inside the confidence area. In this paragraph it will be shown outputs of the code when it is used in a building with much more dispersive heat load and with incredibly low energy consumption during summer.
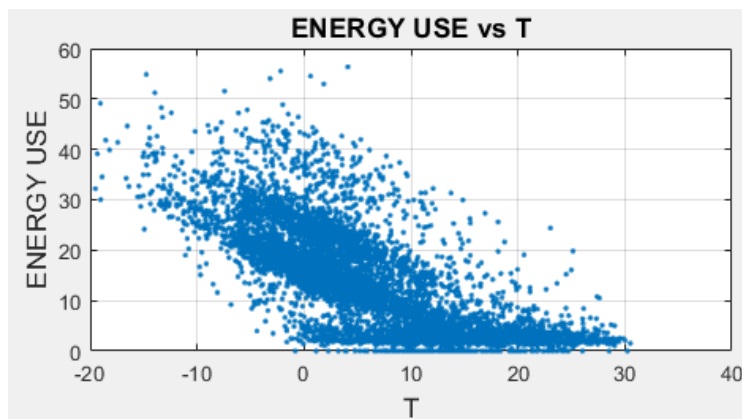
Figure 69 Hourly energy demand versus outdoor temperature in building 10686, Tartu

For this building, both SH and DHW are provided by DHN, so the complete analysis can be performed. In figure 69 it was reported the global energy demand versus outdoor temperature for every hour of the year, to have an idea of degree of dispersion of experimental points. It is important to underline that these data, like the others in Tartu's database, were already pre-processed, so dispersion cannot be imputed to shoddy quality
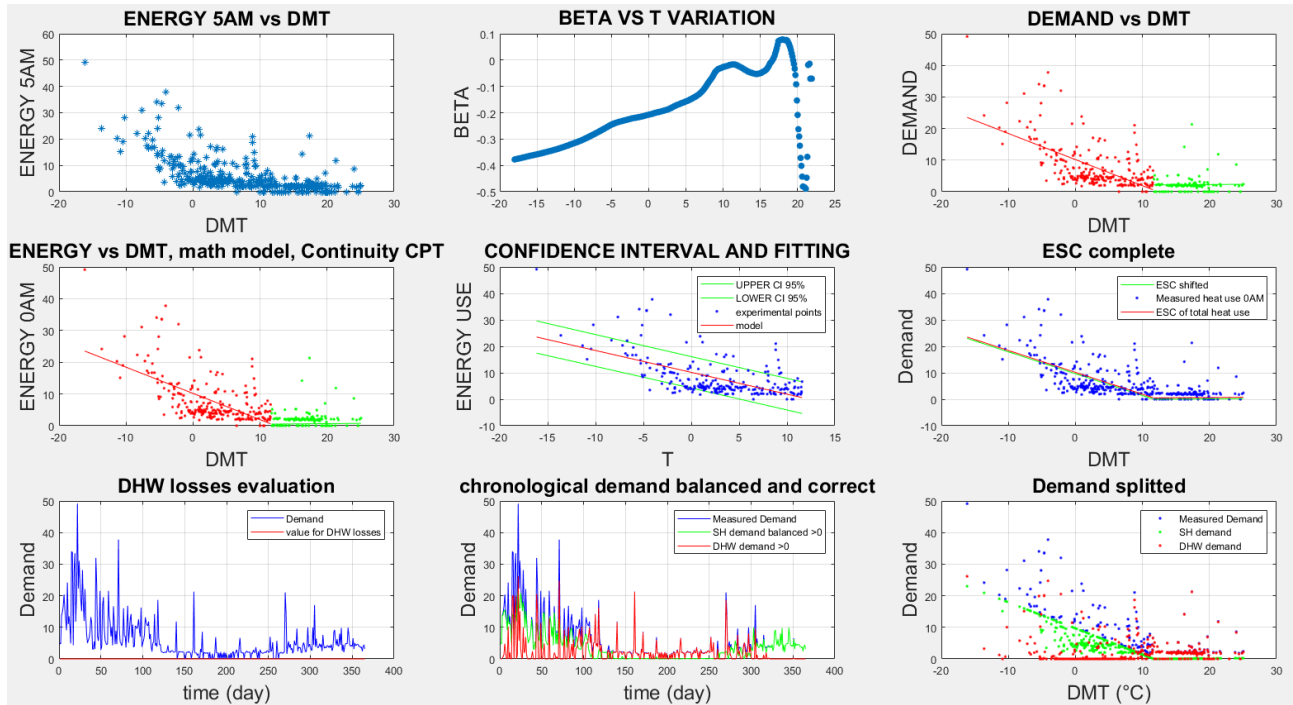
Figure 70 Figure from the output of the building 10686's analysis. Hour between 5.00 am and 6.00 am.
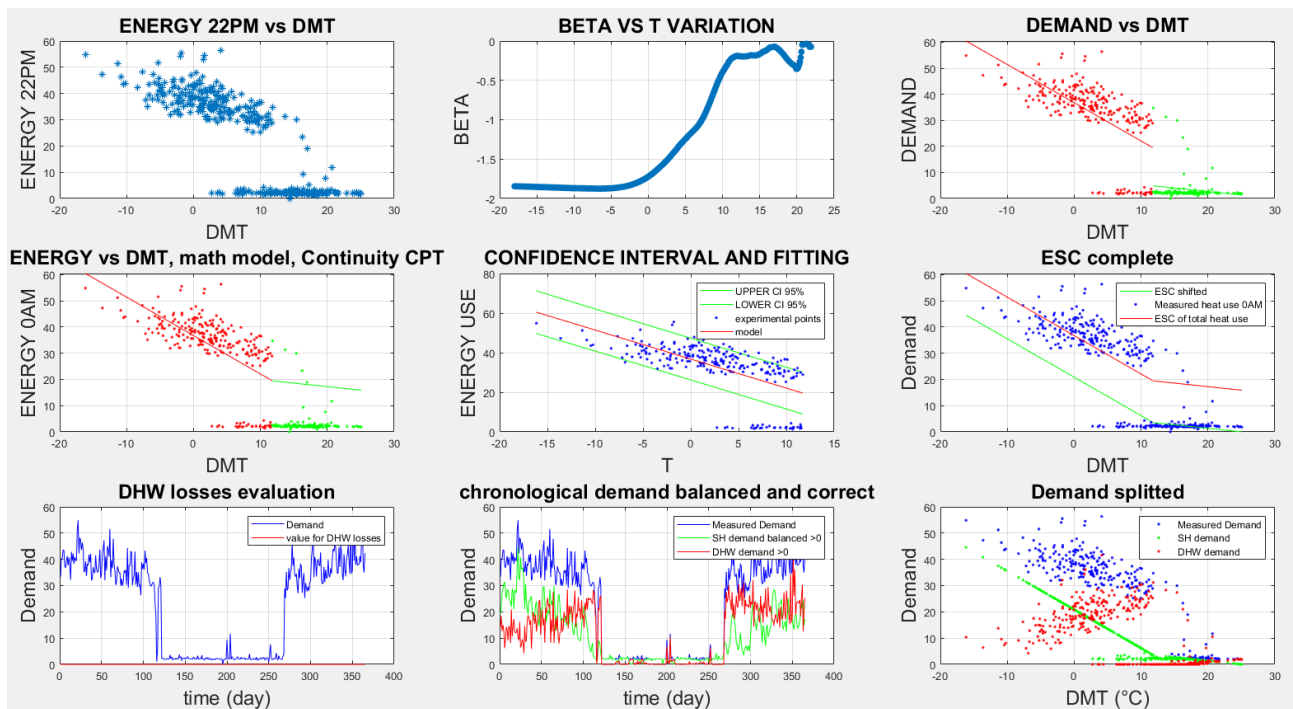
Figure 71 Figure from the output of the building 10686's analysis. Hour between 22.00 pm and 23.00 pm.

of starting data.

By observing hourly demand vector, it is possible to give an interpretation of this dispersion. Points of higher demand, in fact, do not appear randomly but they happen during a precise hour of the day. In particular it emerged that lowest point of demand occur always between 5.00 am and 6.00 am, while peaks of demand are between 22.00 pm and 23.00 pm. it is reported in figures below the outputs of hourly analysis for the hour between 5.00 am and 6.00 (figure 70) am and between 22.00 pm and 23.00 pm (figure 71) to show the huge difference in terms of points of consumption.

Analysing figure 71 it is possible to notice the importance of utilising confidence intervals. In the figure is reported the results of the analysis with all 365 values for the year considered. However, when there is a substantial difference between period with SH and without SH, as happens in this particular case, the interpolation of the temperature dependent part could be less precise. As emerged from the plot and also by value of R^2, regression line is not good representation of what happens in reality. Value of R^2 is 0.37, extremely low value compared with other values from different hours. In general, they stay between 0.5 and 0.71. Lower values are always caused by wrong interpolation due to the high influence of points with zero demand. In fact, even if they are relatively low in terms of number (less than 20)
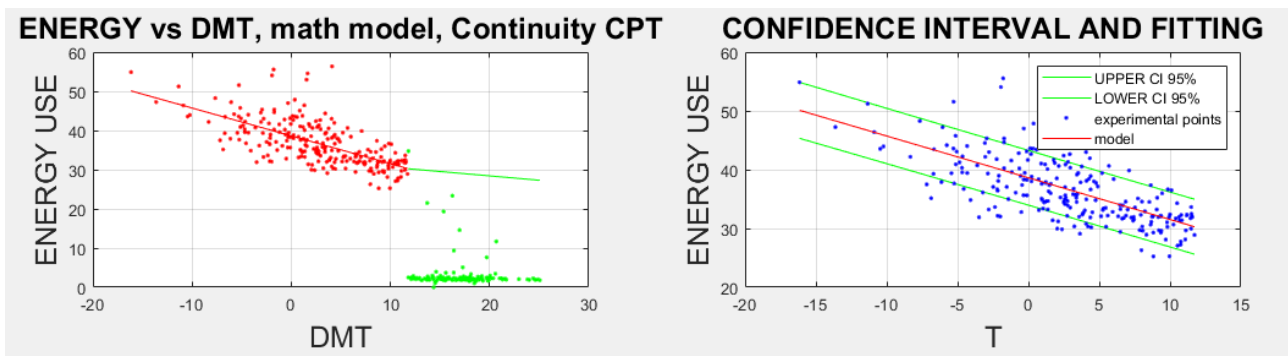
Figure 72 Results of interpolation and confidence interval after remotion of points far from confidence line in building 10686

they have high weight when interpolation is done because of their low values of demand. For this reason, it was decided to perform, just for checking purpose, an analysis by excluding these points in the interpolation of the temperature dependent part of the curve. The result is reported in figure 72. In this way, R^2 increases to 0.45, indicating an improvement in accuracy of the fitting. However, improvement is not so high.

Talking about figure 70, about hour between 5.00 am and 6.00 am, the behaviour is different compared to the others seen until this point. Values of heat load consumption are much lower, closer and it is difficult to individuate a straight line. This fact emerges also from the analysis of β plot, where oscillations are much more evident and present. The interesting part of the line, where the decrease should be evident because of the passage from a temperature independent zone to a temperature dependent presents a descent slower than usual behaviour. Also, the magnitude of β is much different from other hours of the day. In general, it is always lower than -0.8 for all hour of the day, except for this specific hour, where the maximum value is higher than -0.4. This highlight the fact that the demand in this hour is much less dependent to outdoor temperature than other hours, suggesting that SH is not so relevant, or it is limited to a certain value.

Talking about this last point, it emerged also from analysis of Lumbreras M. and Martinez-Garay R. that this building presents a so-called night setback. It consists of the limitation of

energy provided to the building caused by a limit imposed by the manager of the district heating network. This is a usual operation to save some energy by limiting heat provided to the building. It is done in hours when there is small demand, especially during the night, when users are sleeping. It is possible only for small amount of time because a prolonged application could cause discomfort. It is done in this hour because thermal inertia of the building can guarantee an adequate indoor temperature, until the reactivation of the SH plant for the morning heat supply at 6.00 am (data emerged from the analysis of other hours' profiles).

Regarding order of magnitude of heat demand, maximum global heat consumption here is about 55 kWh. It means this is a building with a larger size than 10051 but much smaller than 10922.

With highly dispersed data it is difficult to provide an accurate value for the global heat losses coefficient Q_{tot}. Furthermore, by having significant differences between hours of the day, Q_{tot} values also reflect these discrepancies. In figure 73 it was reported the table with hourly values for Q_{tot} and in figure 74 there is the plot where its trend is drawn. The same axis proportion as the previous similar plot (figures 65 and 52) are maintained, to have the possibility to make a comparison.

N hour	Hours of the day	Qtot
1	0.00 am - 1.00 am	-1.22
2	1.00 am - 2.00 am	-1.26
3	2.00 am - 3.00 am	-1.22
4	3.00 am - 4.00 am	-1.28
5	4.00 am - 5.00 am	-1.18
6	5.00 am - 6.00 am	-0.82
7	6.00 am - 7.00 am	-0.74
8	7.00 am - 8.00 am	-1.16
9	8.00 am - 9.00 am	-1.07
10	9.00 am - 10.00 am	-1.1
11	10.00 am - 11.00 am	-1.12
12	11.00 am - 12.00 am	-0.98
13	12.00 am - 13.00 pm	-1.14
14	13.00 pm - 14.00 pm	-1.12
15	14.00 pm - 15.00 pm	-1.14
16	15.00 pm - 16.00 pm	-1.1
17	16.00 pm - 17.00 pm	-1.12
18	17.00 pm - 18.00 pm	-1.13
19	18.00 pm - 19.00 pm	-1.16
20	19.00 pm - 20.00 pm	-1.19
21	20.00 pm - 21.00 pm	-1.08
22	21.00 pm - 22.00 pm	-1.12
23	22.00 pm - 23.00 pm	-1.47
24	23.00 pm - 0.00 am	-1.29
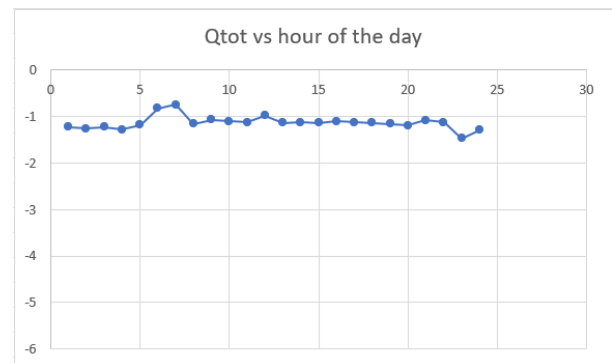
Figure 39 Plot with Q_{tot} hourly values vs hour of the day for building 10686 in Tartu.

Figure 73 Table with Q_{tot} hourly values for building 10686 in Tartu.

It is interesting that there are not big differences in values of Q_{tot}. As said before, biggest discrepancies in these values can be notice in correspondence of hours with more particular behaviour. Hours of major interest are the one between 5.00 am and 7.00 am and between 22.00 pm and 23.00 pm, as highlighted in previous lines. It can be seen that in the morning, when the night setback occurs and the energy provided is presumably limited by the DHN

manager, Q_{tot} is higher (lower in absolute terms), underlining the fact that energy consumption has less dependence to outdoor temperature. Vice versa, during the peak hour at night, Q_{tot} assumes the lowest value (the highest in absolute term) highlighting a high dependence on outdoor temperature. However, these are just particular hours of operation during the day (only 2 hours among 24), so it would be not fair to consider them as a representation of the building's behaviour. It can be seen that the other hours present almost the same value for Q_{tot}, much more similar to the one obtained from 8760-elements vector analysis (an entire year). In this case it emerged that $Q_{tot} = -1.19\ kWh/°C$.

Even in this case, Q_{tot} depends on the dimension of the building, so it is impossible to directly compare this value with other Q_{tot} from different buildings. However, it was calculated even in this case the specific global heat losses coefficient by dividing it by the maximum load. After that, it was compared the line created using specific Q_{tot} and specific demand at 0°C to have an idea of the losses' behaviour of different buildings. The expression used for the calculation of specific Q_{tot} and the specific regression line are written in equations [7.10], [7.11] and [7.12].

$$Q_{tot,spec} = \frac{Q_{tot}}{\max(energy)} \quad [7.10]$$

$$\alpha_{spec} = \frac{\alpha_{tot}}{\max(energy)} \quad [7.11]$$

$$E_{model} = Q_{tot,spec} * T_{out} + \alpha_{spec} \quad [7.12]$$

In the previous equations $Q_{tot,spec}$ is the specific global heat losses coefficient, Q_{tot} is the global heat losses coefficient, $\max(energy)$ is the peak of experimental values of heat consumption, α_{spec} is the specific interception of y-axis (demand when T_{out} is 0°C), α_{tot} is the interception of y-axis (demand when T_{out} is 0°C), E_{model} is the value of demand obtained from the regression model, T_{out} is the outdoor temperature. For the following comparison between different buildings, it was always considered Q_{tot} and α_{tot} from the results of the 8760-elements vector (the entire yearly demand) analysis.
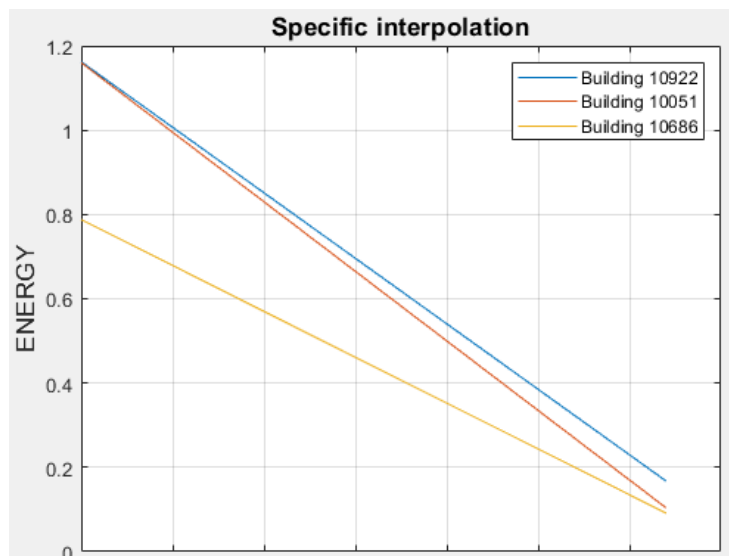
Figure 75 Comparison between specific interpolation lines of the temperature dependent part for buildings 10922, 10051 and 10686 in Tartu.

Results of this analysis are reported in figure 75. It is clear the fact that, using this criterion, building 10922 and 10051 has a similar behaviour, as said in the previous paragraph, while consumption of building 10686 has less dependence on outdoor temperature. This is an indicator of a probable better thermal insulation of building 10686. In case of retrofitting possibility, it would be better to focus the attention before on buildings 10922 and 10051, and after on 10686.

Continuing talking about results, in figure 76 it was reported the output of the 8760-elments vector. Here there is the analysis and the splitting between SH and DHW performed for the whole demand of the year. It is evident the high irregularity of the demand profile. Oscillations are much more evident and with higher intensity than previously analysed buildings. However, the proposed method gives consistent results even in this case. From the last plots, in fact, it is possible to notice that sudden peaks of demand are covered by DHW heat load. This is the most likely interpretation of abrupt growth in demand. Especially
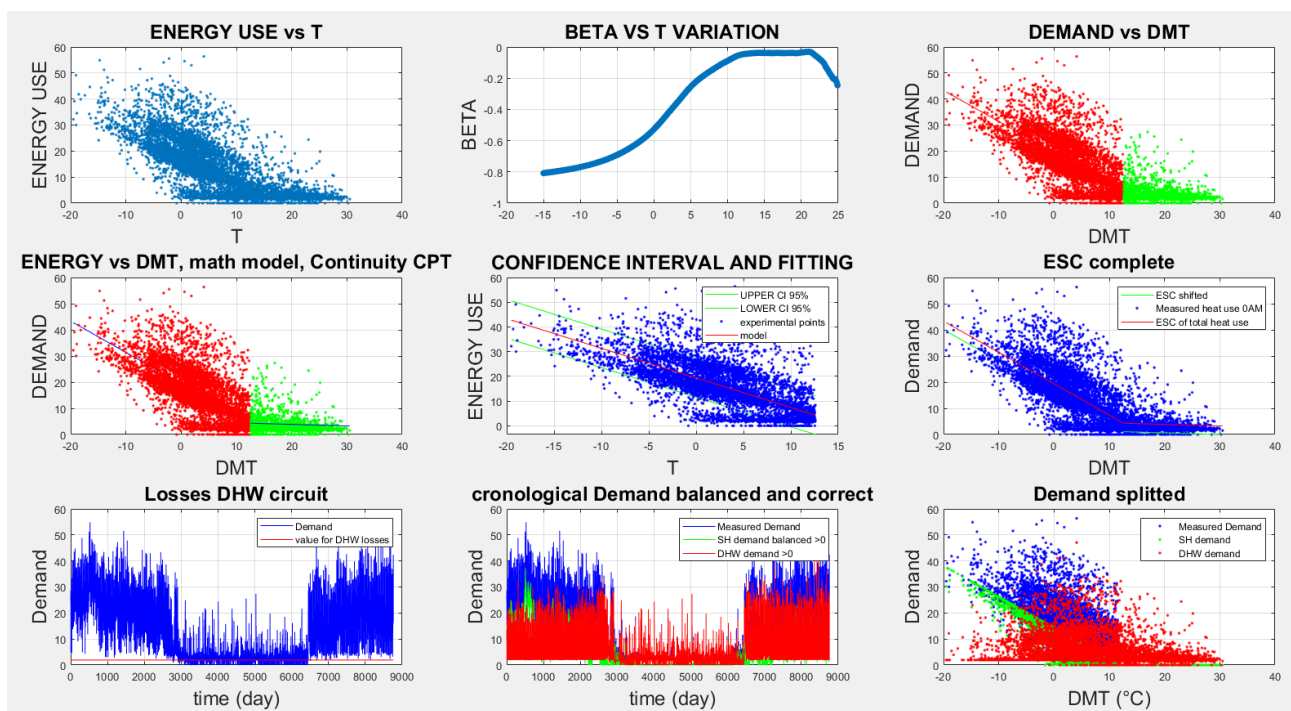
Figure 76 Results of the analysis of the 8760-elements vector of demand for building 10686 in Tartu.

for summer consumption, it would be strange to have peak of space heating consumption during summer days, where outdoor temperature is over 20°C. Summer peaks are clearly caused by sudden DHW heat request.

Talking about confidence intervals, it emerges that there are many points outside the confidence zone. By analysing single hour vector almost, as described before, it is evident that almost every value in the upper zone of the plot belongs to the period between 22.00 pm and 23.00 pm, the peak hour, while a major part of points below the lower confidence line happens between 5.00 am and 7.00 am, the period with minimum consumption. So, by considering points inside the confidence zone it is possible to obtain a fair interpretation of the building's behaviour during most of the time of its operation.

Because there is summer consumption it is possible the estimation of number of people living in the building. Assuming that it is a residential building, hypothesis very probable
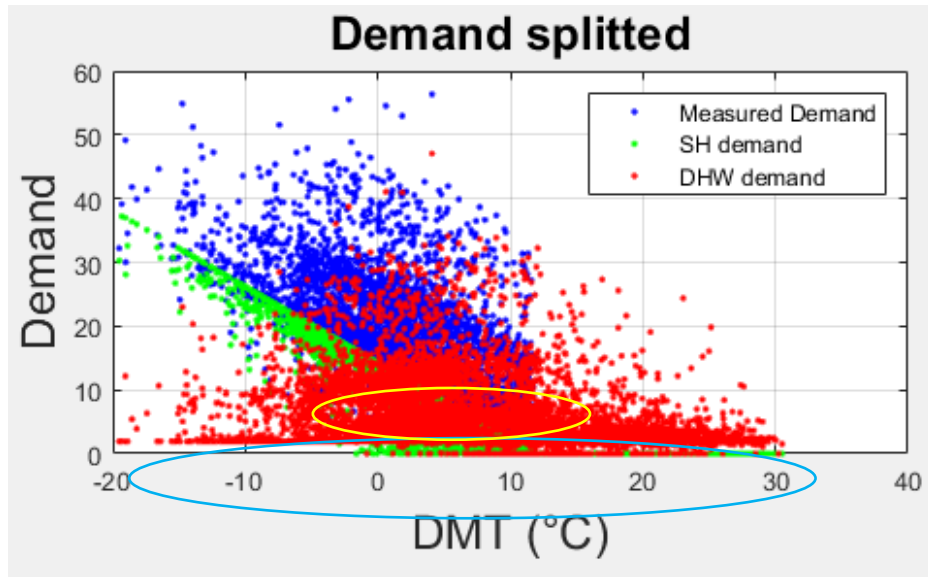
Figure 77 Demand splitted vs T out of 8760-elements demand analysis in building 10686 in Tartu.

because of conclusions emerged from the analysis of the following part of the outputs, presented in the next lines, it can be used the method illustrated in paragraph 7.1 to have an idea of the number of people. By using equation [7.9] and with the DHW heat consumption obtained from this first part of the analysis, there is an estimation of 63 people living there. By assuming an average value of 3 people per apartment, there would be 21 single dwellings. This data is consistent with the ones found previously and it is reasonable with the dimensions of the heat demand. Even in this case after smoothing operation number of estimated people decreases because annual DHW heat load decreases.
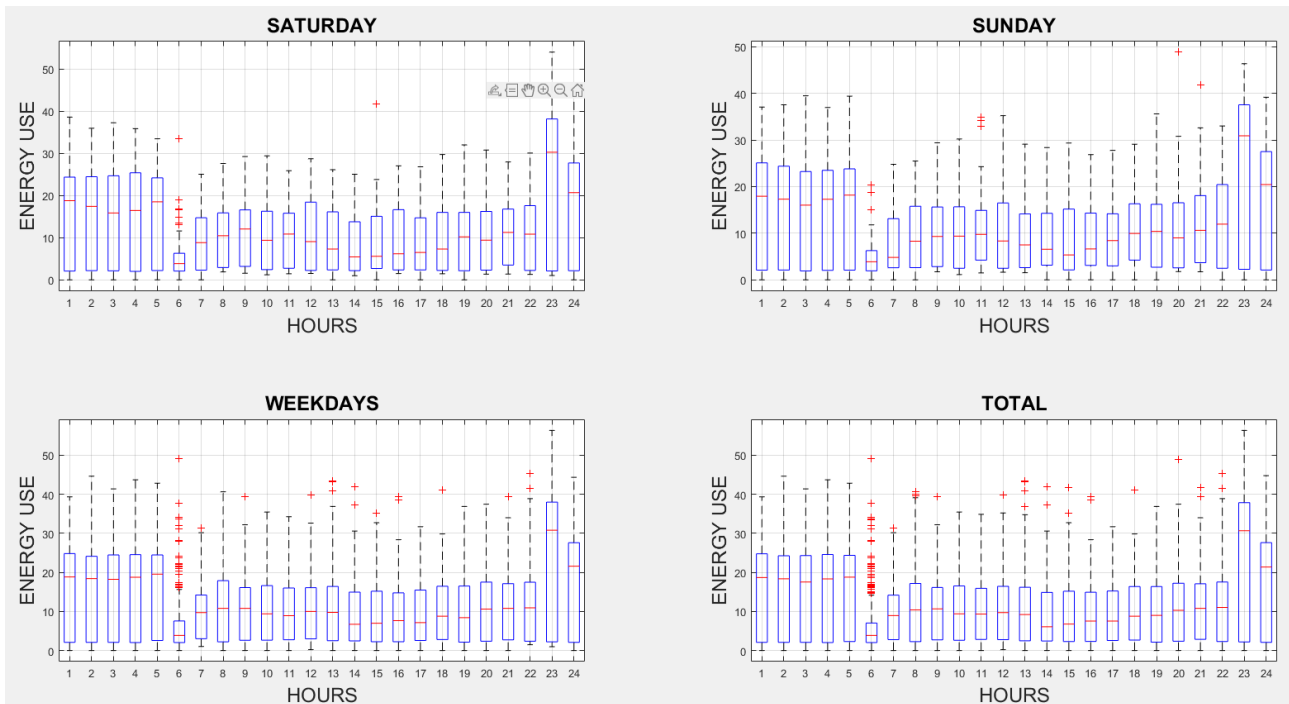
Figure 78 Boxplots of yearly energy consumption divided into Saturdays, Sundays, workdays and total of building 10686 in Tartu.

It is clear by observing the last plot in figure 76, in fact, that DHW load is quite accurate, but between summer and winter there is a global increase in this type of load. it is possible to distinguish a sort of base load both in summer and in winter, highlighted in figure 77 with a blue circle. In addition to those points and the extraordinary high values of DHW heat consumption (peaks), there is a zone, highlighted with yellow circle, of intermediate consumption. It is higher than baseload, but lower than peaks and it is concentrated in temperatures close to the CPT. This zone is created by the model used for estimating SH and the fact that near CPT space heating demand is lower. This means that higher points of demand in temperature near CPT are caused by DHW.

After the hourly vector analysis, in the following lines it will be provided some comments about boxplots of global heat load and about summer profile. Starting from the first one, in figure 78 there are the plots divided in Saturdays, Sundays, workdays and the complete week. In these plots it is evident the presence of outliers, highlighted with red cross (+). Hours where they are much present is the minimum consumption period, between 5.00 am and 7.00 am. This period represents the minimum consumption moment for every day, no distinctions are made between weekends or workdays. It is possible to notice that there is a higher number of outliers during the workdays, but this is just because there are more days considered to draw this plot (for every week there are only 1 Saturday, 1 Sunday but 5 workdays). The maximum consumption hour is between 22.00 pm and 23.00 pm, as said in the previous analysis. Even in this case, no differences are present between workdays and
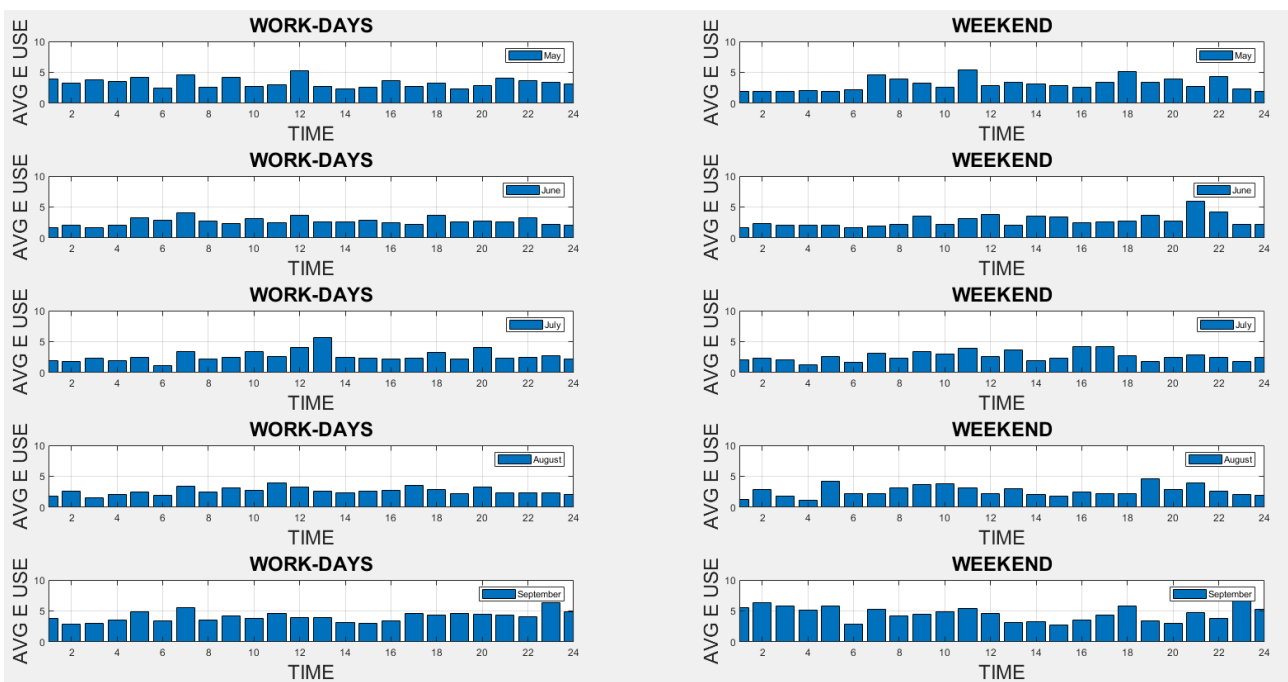
Figure 79 Global consumption in summer months (from May to September) divided in workdays and weekends for building 10686.

weekends.

Regarding the other hours of the day, during the night there is a similar consumption between workdays and weekends. The only notable difference occurs between 2.00 am and 3.00 am, where during weekends consumption is lower. Because this building is subjected to night setback, as emerged also in the already mentioned analysis of Lumbreras M., Garay-Martinez R. et al., it could be possible that this reduction of heat provided during

weekends is caused by DHN management. Another difference can be noticed in correspondence of hour between 12.00 am and 13.00 pm, where load in weekends is lower than workdays. Regarding the rest of the hours, it was not found any other interesting points.

Talking about summer profiles' analysis, there are some peculiarities in this building compared to the ones analysed before. In this case there are less evident differences between May, September, and the rest of the summer. Demand in September is globally higher than other months, but the difference is not as evident as happened for building 10922. This could be explained by the fact that building 10686 is smaller than 10922, so the space heating demand, in absolute term, is lower. Talking about May, the average demand during this month is quite like June and other summer months. The only difference in terms of magnitude of consumption can be found during night hours. It is important to highlight that, because this is a smaller building, influence of behaviour of a single user on the profile of DHW heat load is much more evident. This is the case, for example, of the consumption between 12.00 am and 14.00 pm in July, where demand is higher than the same period in June and August. Considering these differences, the period between June and August was used for the analysis of the DHW profile in summer period. Because there is the presence of differences even among these months, it is expected that this analysis could be less accurate than the previous buildings.

In the figure below, number 95, it was reported results for the analysis of correlation between TMA and energy consumption. The meaning of this exam is reported in paragraph 4.5. for this building it is possible to say that the thermal inertia is quite low compared to the other two buildings analysed. From the result reported in figure 95 it is evident that the maximum correlation between energy consumption and splitted outdoor temperature (using TMA) occurs at 7 hours. However, it is important to underline that, as happened in previous cases, the coefficient correspondent to this hour does not differ too much from the β calculated with the interpolation between outdoor temperature and hourly energy demand, calculated in previous passages.
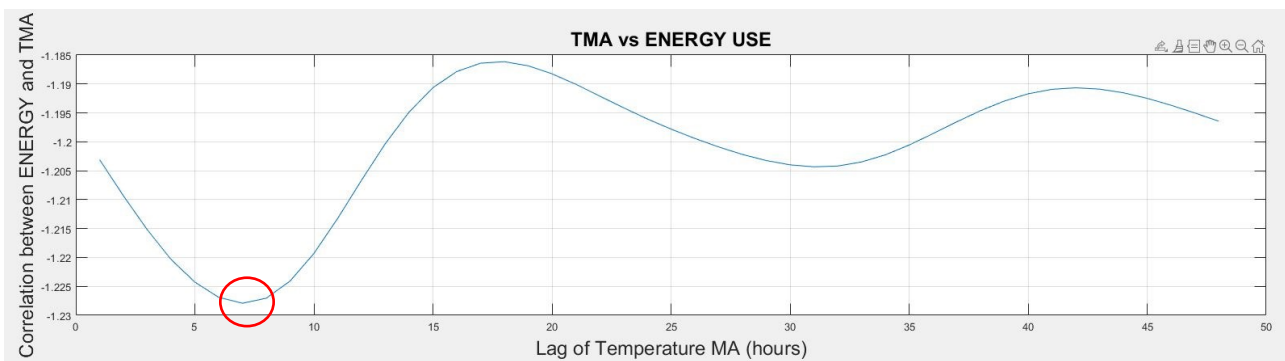
Figure 95 Effect of temperature lag on energy use of building 10686, Tartu.

From the analysis of the match matrix, it emerged that Sunday and Saturday can be considered as similar day. The other days present low similarities with Student's t-test and Fischer's criterion, but to keep analysis simple and because difference in profiles is not so evident in absolute terms, it is possible to maintain even in this case 2 group of days. The first one is composed by Saturdays and Sundays (group 1) and the second one by the other

days of the week (group 2). In figure 80 results of the group separation and the individuation of minimum, medium and peak hours have been reported.
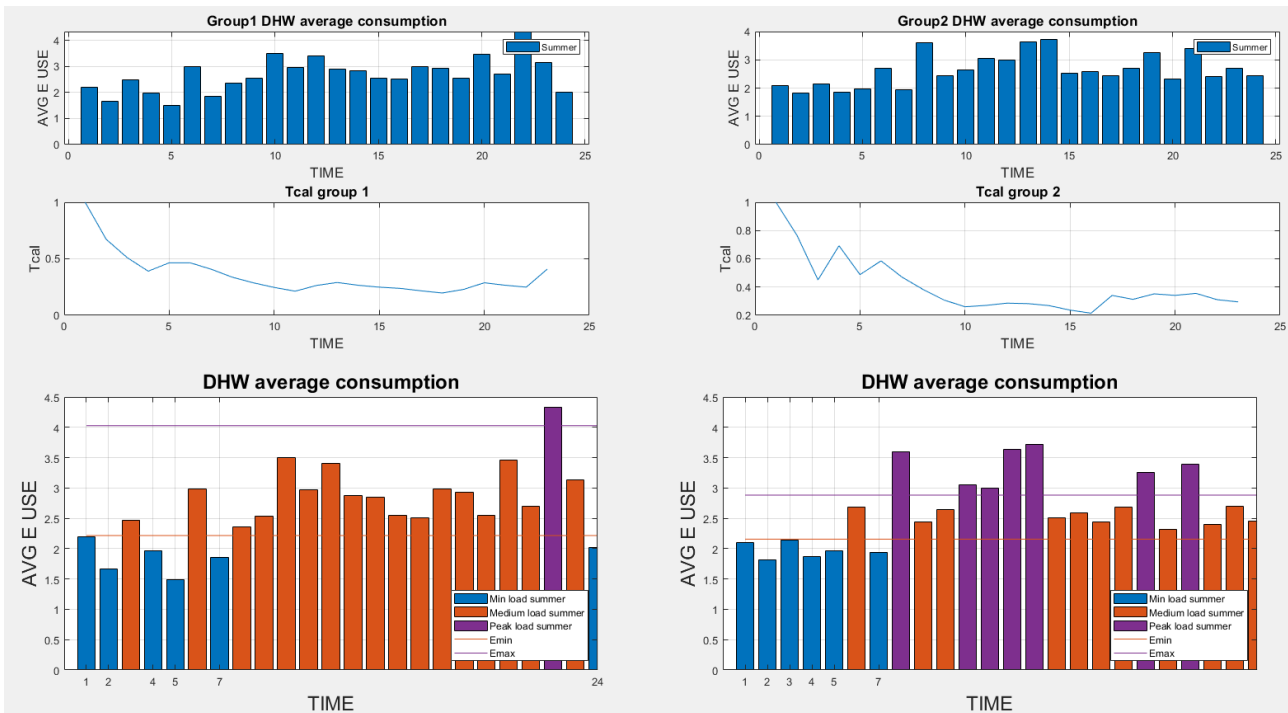
Figure 80 Results from the splitting of hours into minimum, medium and peak periods in building 10686, Tartu

Even in these plots there are information previously found such as the minimum consumption between 5.00 am and 6.00 am and the night peak between 22.00 pm and 23.00 pm. in this building profiles are more variable because of a smaller number of users, so the smoothing effect of DHW heat demand created by large amount of people is less present. It is important to underline that for the workdays, in group 2, hours highlighted as peak hours present a consumption similar to some hours of medium consumption during weekends. This happens because the division between minimum, medium and peak hours take into account the distribution of load in the analysed day and does not consider other days. So, the right interpretation is to consider purple hours in group 2 plot as peak hours for workdays, but not with an absolute meaning. They are calculated, in fact, to have an idea of values for minimum, medium and peak consumption for the considered days. In the next part of the analysis, in fact, to each day of the week is applied the profile of the corresponding group.

Even in this case moments where consumption is higher, which represent points of highest DHW heat demand, are in the morning, during lunch time and during the evening for dinner and after dinner. Talking about boundary values, E_{min} is 2.21 kWh for the group 1 and 2.16 kWh for the group 2, while E_{max} is 4.03 kWh for the group 1 and 2.88 kWh for the group 2. It emerged a higher consumption of DHW between 22.00 pm and 23.00 pm during weekends. To interpretate in a fair way these data it is important to keep in mind that this analysis is conducted only on summer demand. There is a high probability that some users are on holiday in this period, with a consequent modification of the building's behaviour in terms of demand and energy consumed. When there are a lot of people this effect is less relevant, but following previous estimation, in structure 10686 there are about 20 apartments, so even

an extraordinary behaviour of one of them could cause a big variation on the entire demand profile.

In figure 81 it was reported standard profiles in winter and in summer for DHW demand. The difference in terms of mean value is about 40% even in this case, as happened for building 10922. This is mainly caused by the model used for estimating SH consumption and the extraordinary points of consumption in the CPT zone (yellow circle in figure 77).

Figure 81 Standard profiles in winter and in summer for days of group 1 (left) and group 2(right) in building 10686, Tartu.

With the correction performed with standard profiles, situation becomes much better in terms of consistency of DHW profiles between summer and winter. In figure 82 there are results of the correction and filtering analysis. Looking at the last two plots, it emerged that DHW profile in winter is much more regular and like the demand in summer. SH, on the other hand, changes for compensating modifications of DHW profile, but it continues to follow a linear behaviour. It is just more randomize and it does not have any more the interpolation line as upper limit, as happened in the previous analysis. The discrepancies between old profiles and corrected profiles depends on tolerance and decreasing factors. Unlucky there are no possibilities to determine in an accurate way these parameters for the reasons described in paragraph 7.1, so only an estimation based on experience is possible. Even for this building it was decided to consider 0.3 for tolerance and 0.25 for decrease, to have a corrective operation comparable with building 10922.

With new DHW heat consumption, number of people becomes about 52. Number of apartments is 17. There is always a reduction in these values because with corrective operation DHW profile is always lower. The goal is, in fact, to reduce oscillations in winter demand and make DHW consumption in this period more like the one present in summer.
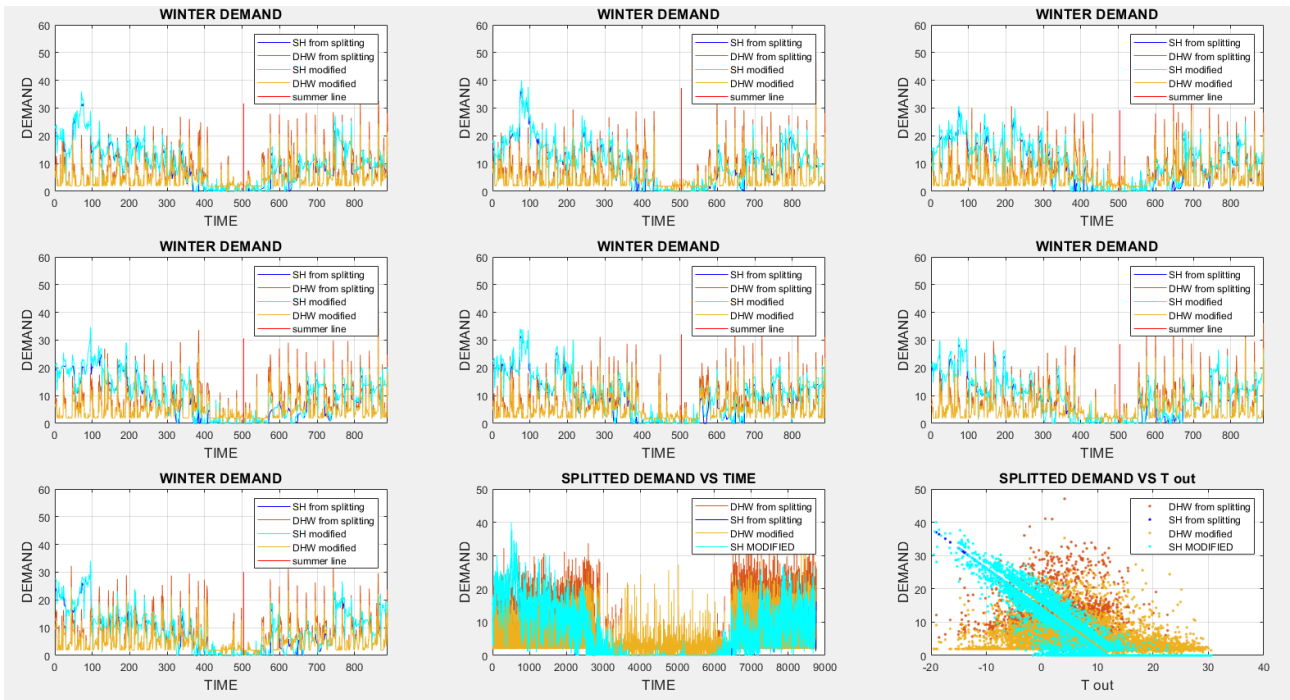
Figure 82 Representations of DHW and SH before and after correction of DHW profiles. From top left there are: Mondays, Tuesdays, Wednesdays, Thursdays, Fridays, Saturdays, and Sundays. Last two plots are about 8760-values hourly vector of year consumption. Building 10686 in Tartu.

In figure 83 it was reported last plot of figure 82 highlighting the different zone of DHW consumption, as done in figure 77. This operation was done to highlight the fact that with corrective operation the high density of DHW points in the yellow circle zone is much lower. This is a good sign: it means that DHW with temperatures a little higher or a little lower than CPT does not change much, as it is reasonable to think.
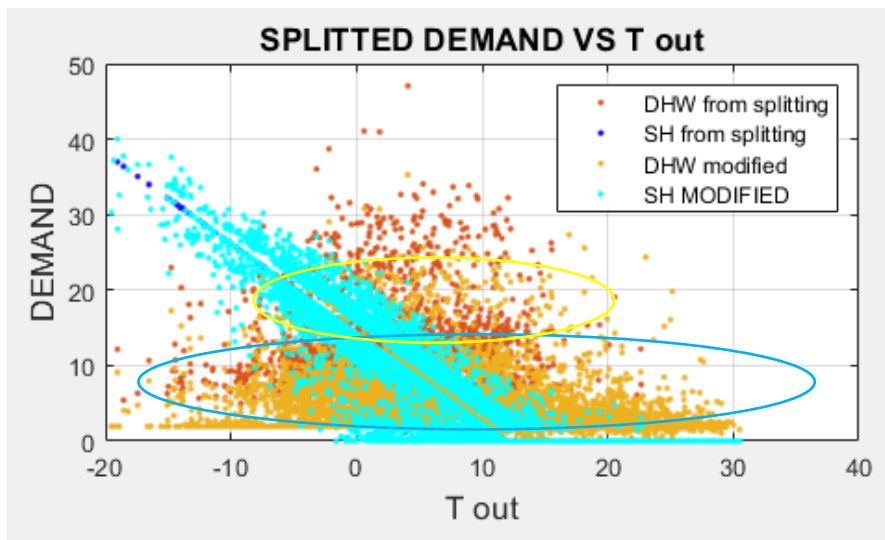
Figure 83 Splitted demand versus Tout of building 10686 in Tartu. In different tones of blue there are SH demands before and after correction and in different tones of orange DHW loads before and after correction. Circles in yellow and in light blue are the same present in figure 77.

It was decided to show this building to demonstrate that the method proposed, and the tools used can give useful and consistent information also with dataset with high irregularities. However, more regular is the dataset, more accurate will be final information, but it is important to have an algorithm that can work well with diverse types of building. With this building it finishes the presentation of Tartu's dataset results. Among the enormous number of buildings analysed it was decided to show you a standard example (10922) and two particularly challenging cases: the first one without summer consumption (10051) and the second one with high dispersion of experimental values (10686). From the results it emerged that the method proposed works effectively in all cases.

7.4 Building "Condominium 10" in Verona.

From this paragraph it begins the analysis of Verona's dataset. The peculiarity, as said many times during the thesis, is that all buildings present a partial database, with only winter global heat consumption data available and, in general, more information about buildings are available. Even in this case it will be presented a quite standard case, the condominium 10, and then it will be shown some particular buildings, such as a school and a train station to highlight some peculiarities. For this analysis and the following ones, it was used methodology described in chapter 5. Most considerations and outputs are like the ones presented in paragraph 7.1, so it is not spent so much time in their description. There are presented only most interesting issues about every building.

Some general considerations valid for every building in this dataset are necessary. First, Verona has a quite different climate than Tartu. As highlighted during the explanation of methodology, in the Italian city outdoor temperatures are higher both in winter and in summer. In winter minimum temperature registered for the considered year (12 months, from November 2019 to November 2020), is -3.4°C, while in Tartu there were temperatures below -20°C. In summer maximum temperature is 35.5°C, much higher than Tartu, but this data is less relevant because no experimental data are available for this season. For this reason, as a general expectation global heat demand is lower in Verona. Space heating has less importance in the global heat demand than what happened in Tartu because thermal losses are less due to higher outdoor temperatures.

Another thing that all building has in common is that the results have meaning just for the part of the year where experimental points are available. Because heat consumption in summer is unknown, it would be meaningless try to estimate it. Too many variables enter in the game and the high dependence of DHW load consumption to users' number and behaviour (information unknown for most of the building) makes impossible an accurate estimation for summer period. So, all temperature independent part of the ESC drawn in summer period, for temperatures higher than CPT, are to be considered just tools for implementing the analysis methodology and not fair estimation of real building's consumption during this period.

Another common thing is about municipality's rules about space heating. Due to higher temperatures during winter season, there is no necessity to keep SH plant in operation for all hours of the day. So, as mentioned in paragraph 5.4, Verona's municipality established that there is a maximum of 14 hours of operation for the SH pant during the day and its operation must happen between 5.00 am and 23.00 pm. These rules contribute to contain

pollution and keep decent quality of the air. Especially in historic centre of the city, due to the high density of houses, the age of buildings (and, as a consequence, of their heating systems) and the traffic, air pollution is a relevant problem in Verona. Due to these issues, it was born the necessity of interventions to improve air quality. The regulation of space heating is a part of these rules. Consequentially, by analysing plot of global hourly demand versus outdoor temperature, it is possible to notice two distinct zones. An upper area for the moments when both SH and DHW are provided to the building and a lower zone where SH is absent.

Furthermore, even in this dataset data of buildings' demand come from the measurements conducted in substations of DHN. For this reason, it is not measured the real demand of the building, but only the energy provided by the network. By happing this, the fact described in paragraph 7.1 is true. Even in this case, in fact, solar heat gain and internal gains are not considered in the available data. However, this is less notable because of the shutdown of SH during night-time and the less energy required for heating up the building.
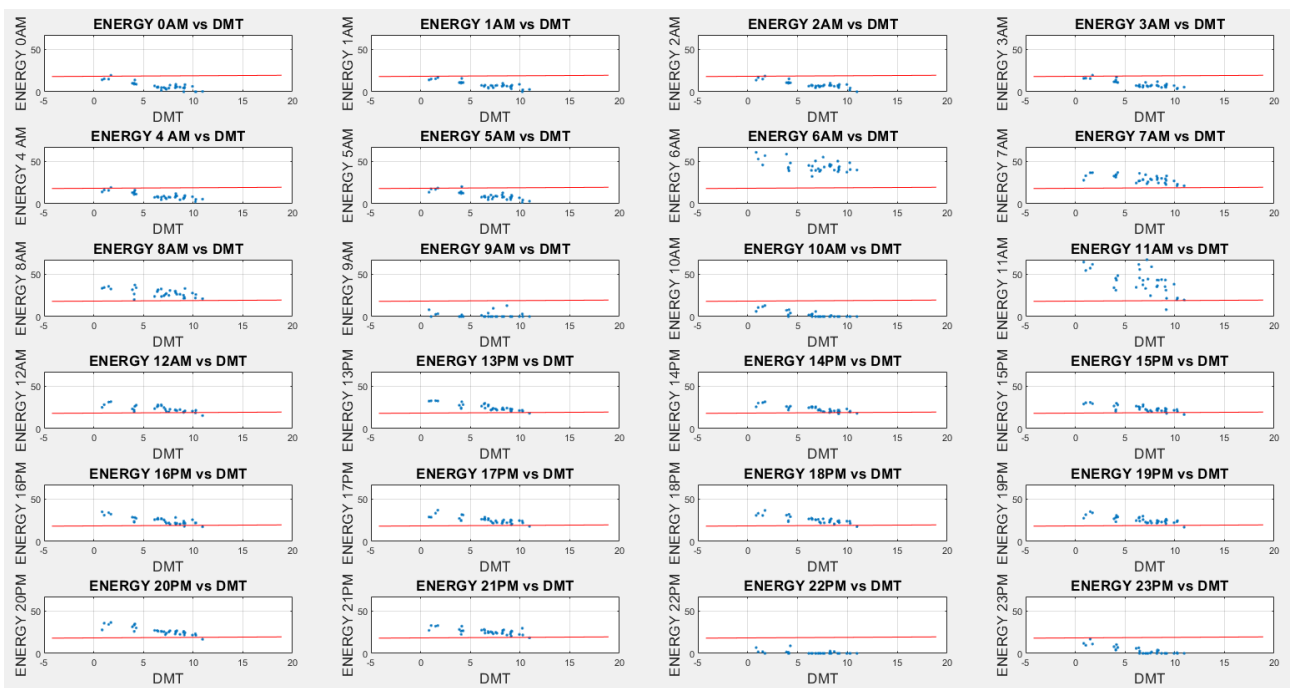
Figure 84 Hourly energy demands representations versus Tout. In red the boundary line between zone with SH and without SH. condominium 10 in Verona.

The first output of the code is reported in figure 84. Here there is the representation of all 24 hourly demand vectors. In red it is represented a line drawn between the two zone of the global demand profile, the one with SH in operation and the other one without SH. It is evident what are the hours with and without SH. It can be noticed that between 22.00 pm and 5.00 am and from 9.00 am to 11.00 am measured demand is below the red line, so no SH is in operation during this period. By summing all these hours, it results that SH stay turned off for 10 hours during the day. This means that there is a total of 14 hours of space heating plant operation, in full compliance with the municipality's rules. Also, limitations regarding hours of operation are respected because SH is turned off between 23.00 pm and 5.00 am.

In addition to that it is possible to obtain some information about the behaviour of the space heating plant. It is clear, in fact, that between 5.00 am and 6.00 am there is a peak in demand. This is reasonable because building must be heated up after 8 hours without space heating. So, a large amount of energy is necessary to reach setpoint temperature in indoor environment. A similar thing happens between 11.00 am and 12.00 am, where it is necessary to provide heat after 2 hours without space heating. Thanks to the small amount of time without heating system in operation (so building's thermal inertia can contribute to keep high indoor temperature) and to the presence of solar heat gain, on average energy required during this hour is lower than the demand between 5.00 am and 6.00 am. However, also in this case demand is quite high, but it must be considered that between 11.00 am and 12.00 am there could be DHW consumption. Consumptions during other hours follows a quite similar pattern, and no other peaks are evident.

It is evident that there are no relevant differences between weekends and workdays. This is perfectly in accordance with the fact that the analysed building is a residential structure. In paragraph 7.5, where it is presented a school, the difference between workdays and weekends is extremely important and by not considering this fact the analysis could lead to very inaccurate, not to say completely wrong, results.
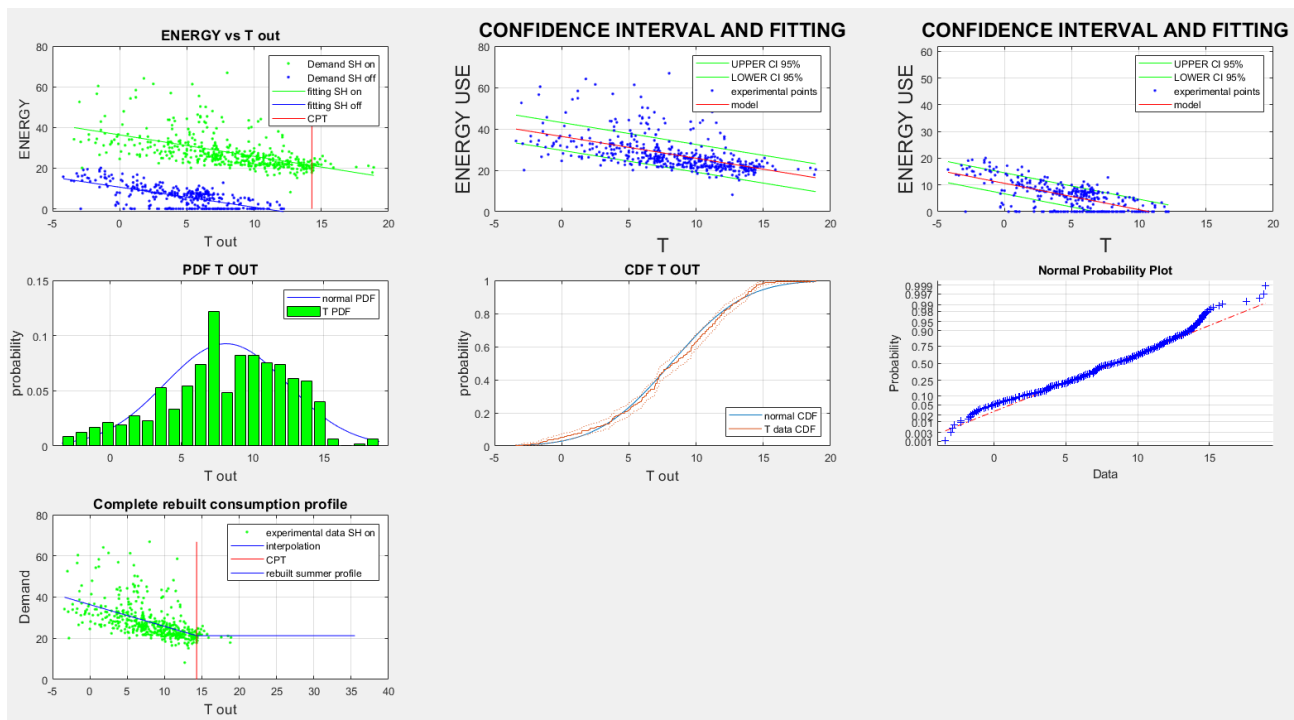
Figure 85 Hours with and without SH separation. Application of confidence intervals, check for normal distribution of temperature and ESC reconstruction in condominium 10, Verona.

In figure 85 it is reported the second image of the analysis' output. Starting from the upper left part of the figure the first plot shows the hourly energy demand versus outdoor temperature, splitted into two different regions. The upper part represents moments when SH is in operation, the lower part when there is no space heating. Vertical red line indicates the CPT. Proceeding on the right there are two plots about confidence intervals. They are referred to SH on period and SH off period. Experimental demand is quite regular, in fact many values are inside the confidence area. Higher values of demand are representation of peaks that happen between 5.00 am and 6.00 am and between 11.00 am and 12.00 am, as described in the previous lines.

Going further, in the central line of the figure there are 3 plots that show results of the distribution type control of outdoor temperatures. As it is clear from the graphs, outdoor temperature follows a normal distribution. These plots are also useful to apply a filter to these values. In fact, as explained in the methodology, it was considered only the 95% right tail for the estimation of CPT. It was found that for this building the change point temperature is 14.3 °C. it is a reasonable result, also in line with the previously found ones in Tartu's database. In fact, despite lower temperatures in winter and a difference in climate, outdoor temperature at which space heating plant is turned off must be quite similar.

In the last plot there is the representation of the ESC, with reconstruction of summer demand, just to methodology purpose, and the temperature dependent part interpolated by a straight line. The same interpretation given in 7.1 can be applied to this part. For this reason, it is possible to say that Q_{tot} is the global heat losses coefficient. In condominium 10, $Q_{tot} = -1.05 \ kWh/°C$. In analogous way to what it was done for Tartu, it is possible to create a specific Q_{tot} to compare buildings in the same location. In the following paragraphs it will be done this operation.

Unfortunately, without summer consumption data it is impossible to estimate total DHW consumption during the year, so it cannot be estimated the number of people living inside the condominium with equation [7.9]. it is, also, impossible to perform a detailed analysis of the summer DHW demand profile and the correction operation on DHW profile.

Figure 86 Third image from the output. ESC shifted, chronological demand representation, normal distribution control for Tout at night, losses evaluation and representation of splitted demand. Condominium 10, Verona.

In figure 86 the third image from the code's output is presented. Starting from the upper left part the first plot shows the energy consumption versus outdoor temperature, highlighting the ESC curve shifting down. The shifting coefficient is the demand in correspondence of change point temperature and for this building it is equal to 21.26 kWh. After that there is a chronological representation of the demand when SH is turned on.

After that, a control on outdoor temperature at night, when SH is not in operation. This is performed because for losses evaluation it is necessary to consider only a part of the lower point of demand, and to select that part it was decided to apply a statistical approach based on normal distribution, as described in paragraph 5.7.1. This method works if outdoor temperatures analysed follow a normal distribution, so these graphs are important to verify this.

In the next plot there is the representation of the experimental values without SH and in red it is highlighted the CPT, while in green there is the estimation of losses into the DHW circuit. Talking about numbers, in this building value of losses for DHW circuit is 2.42 kWh. Comparing this value with peak load it emerged that losses are about 3.5% of the maximum peak load (4.1% if instead of maximum peak it was considered an average of peak points). in Tartu's analysis it was found that losses in DHW circuit are about 4.5% - 5% of maximum load, so the estimation did in this analysis is consistent with that calculation.

The following plots are another representation of global demand versus outdoor temperature or versus time and in the last two plots there is the splitted demand both versus time and versus outdoor temperature. From the analysis of this last part, it is evident that SH is not so important and follows a quite regular behaviour. DHW consumption is much more important, and it is the responsible of peaks.

However, it is important to combine this information with the one obtained from the analysis of figure 84. In fact, due to the linear model used to estimate SH heat load it is not possible to consider sudden peak in this type of demand. So, this model gives reliable results for the hours with average consumption but could give some inaccuracies in peak hours. During these moments, in fact, most of demand is imputed to DHW. It could be possible that there is a peak in demand of hot tap water during these hours but for the reasons explained before, it is more likely that the best interpretation is that most of the peak is caused by the starting of operation of SH system.

Because these peaks in consumption after hours where SH plant was not in operation are common in buildings of this dataset, it was decided to find a way to obtain a more precise model for the building. The attention is focused on the most common profile of demand of every building, i.e., hours with and without SH but of minimum and medium consumption (without peaks). For this reason, it was applied the same analysis but considering only demand points inside confidence interval. In this way it was removed almost every peak of demand that happen between 5.00 am and 6.00 am and between 11.00 am and 12.00 am. In this way, the model created considers only hours with most common profile of consumption during the days. The outputs are remarkably similar to the previous one because code applied is the same.

In figure 87 there are the same plot as figure 85, but without values outside from confidence interval. In the first plot it was highlighted in green points considered for this analysis and in blue excluded values. In this figure only experimental values with SH turned on are represented. Plots whit experimental data and confidence intervals (plots 2 and 3 of figure 85) are not reported here because it would be useless, as CI was already applied. Even in this site there is the control of normal distribution for outdoor temperature. The last plot represents the reconstructed ESC, as happened before.

In numerical terms, CPT is almost the same as the previous one, it is equals to 14.4°C (without CI filter was 14.3°C). Also change point demand (CPD) remains almost unchanged. In this analysis it is equal to 20.61 kWh, while before it was 21.26 kWh. In terms of coefficient of global heat losses, it emerged an interesting consideration. With remotion of values outside the confidence interval, $Q_{tot} = -0.89 \, kWh/°C$. With all values considered, $Q_{tot} = -1.05 \, kWh/°C$. By removing extraordinary points, most of them are peak values, the dependence of energy consumption on outdoor temperature is reduced. This is reasonable: without points outside the confidence interval, peaks of consumption are missing, so there are values more regular to interpolate and, also, lower in terms of energy demand. However, this value is a better representation of the building's behaviour during most of the hours of the day.
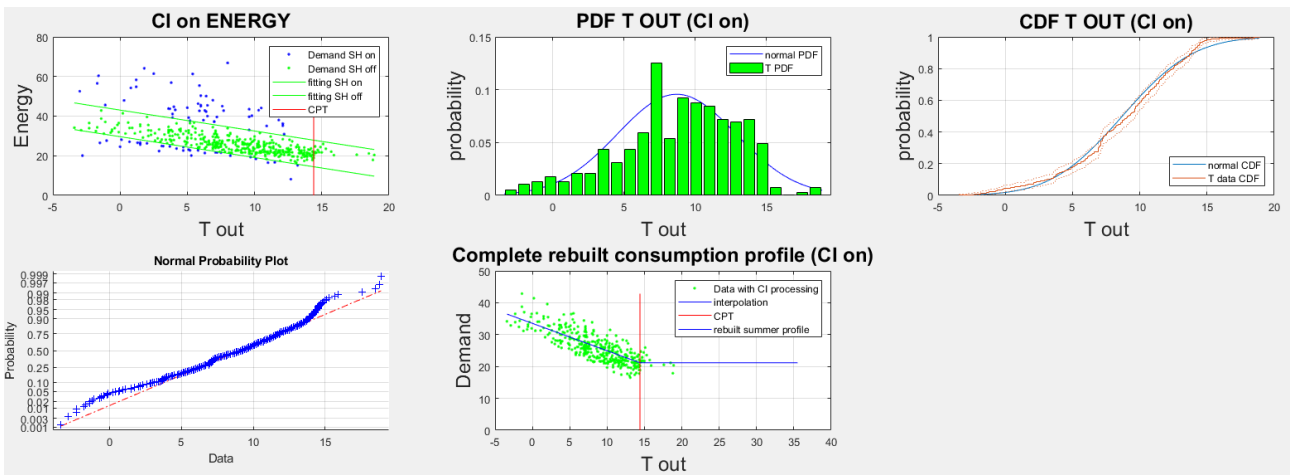
Figure 87 Output of analysis without values outside the confidence interval. from top left: representation of demand values with SH turned on and CI, control for normal distribution of Tout, reconstruction of new ESC. Condominium 10, Verona.
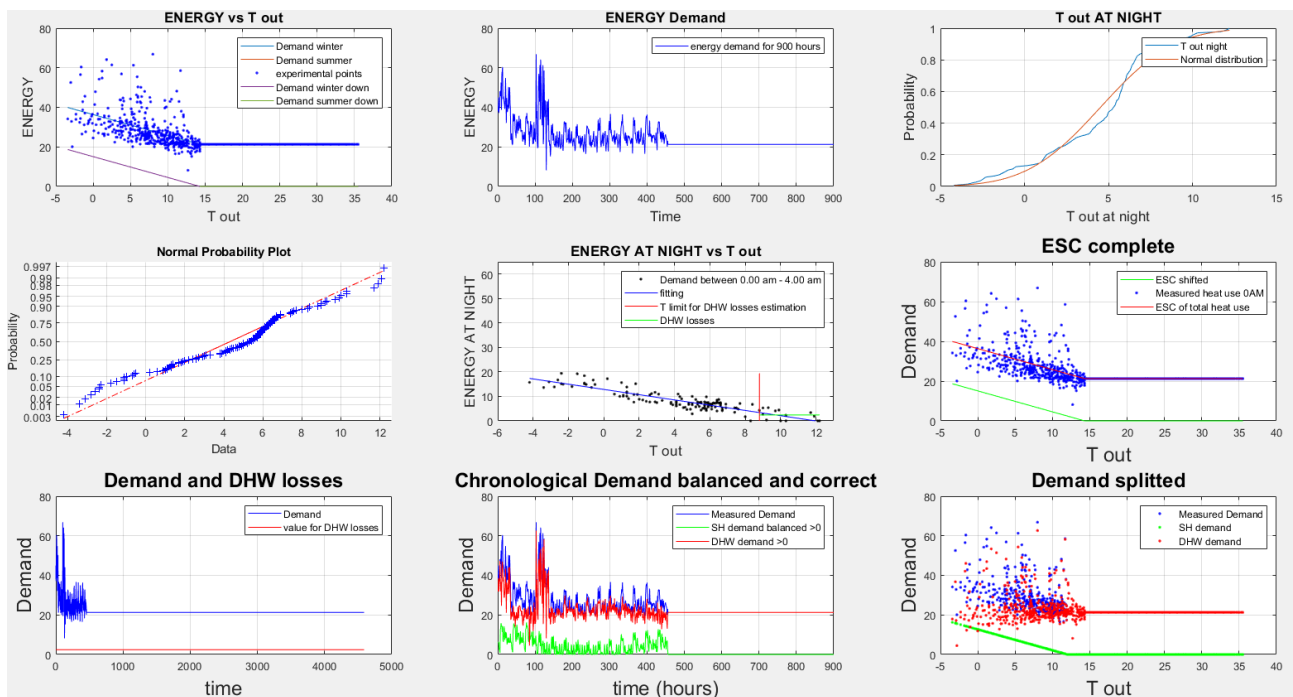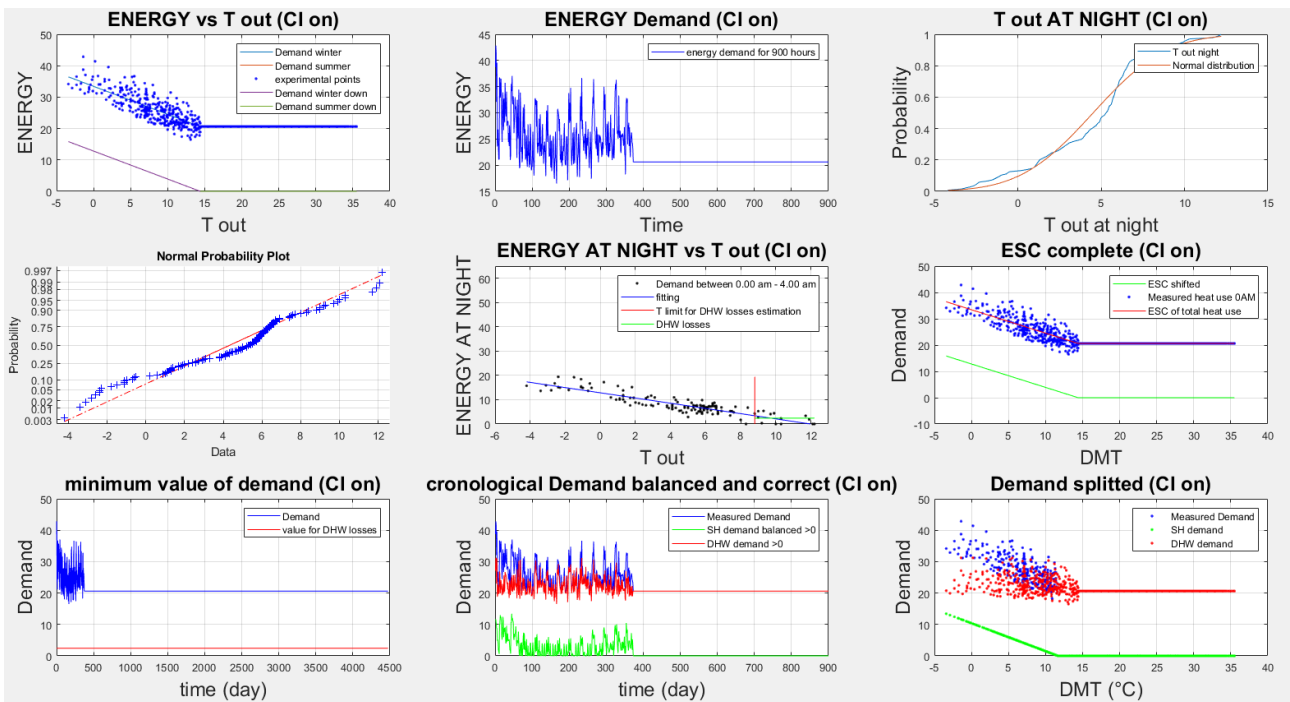

Figure 88 5th image from the output. ESC shifted, chronological demand representation, normal distribution control for Tout at night, losses evaluation and representation of splitted demand. CI filter. Condominium 10, Verona.

The last output of the code in exam is presented in figure 88. There are the same plots as figure 86 but with the filter of confidence interval. No particular differences can be noticed in the first plots. The value for losses in DHW circuit remains the same because in this part of the analysis CI filter is not applied. The most significant difference can be found in the last two plots. Here, compared to figure 86, DHW demand is more regular. No high peaks are identified in demand, of course.

To conclude, analysis with CI is better for modelling most common behaviour of the building but, of course, it does not take into account of sudden peaks in demand caused by big DHW request or the starting of operation of SH plant after hours of stop.

As last output, in figure 96 reported below, we can find the analysis of correlation between TMA and energy consumption. In this building the pattern of β coefficient is less regular than previous buildings, but it is recognizable a local absolute maximum value in correspondence of hour 5 (and even 28). It is reasonable to consider a sort of similarity between the 5th previous hour and the 28th previous one because they are almost 24 hours apart, so it means it is more or less the same moment of the day. In case of necessity of a more detailed model it would be it would be good to consider both hours.
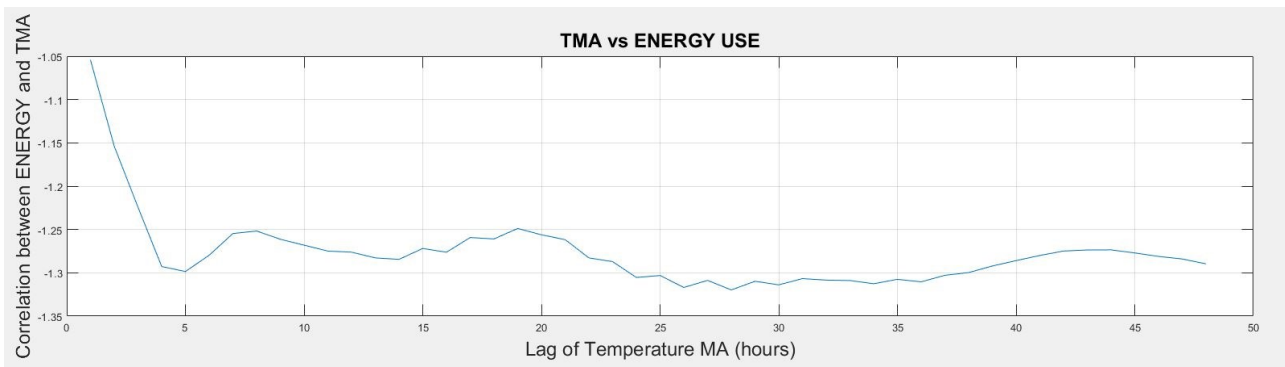
Figure 96 Effect of temperature lag on energy use of condominium 10, Verona.

7.5 Building "School 2" in Verona.

In this example it was analysed a school. It is an interesting case because school is much different than a residential building in terms of energy consumption. First, there is high influence of day of the week. Of course, during Saturday and Sunday school is closed, no people are inside. For this reason, demand during these days is null. In addition to that, students are at school during the morning and, sometimes, the first hours of the afternoon. During evening and night-time no people are present in the structure so there is no necessity for space heating or hot water.

Another important thing regards the distribution of SH and DHW load. In fact, in a school there are not showers, dishwashers, washing machines or kitchens. For this reason, DHW load in a school is way less than in a residential building. According to standard values (values taken from school output specification document from the Department of Education of UK), the average daily hot water consumption is 2.3 litres per person for primary school and 4 litres per person for secondary school with showers. Assuming that there are no

showers in Verona's school, the average consumption is about 2 litres per person per day, compared to the 40 litres per person per day of a residential building.
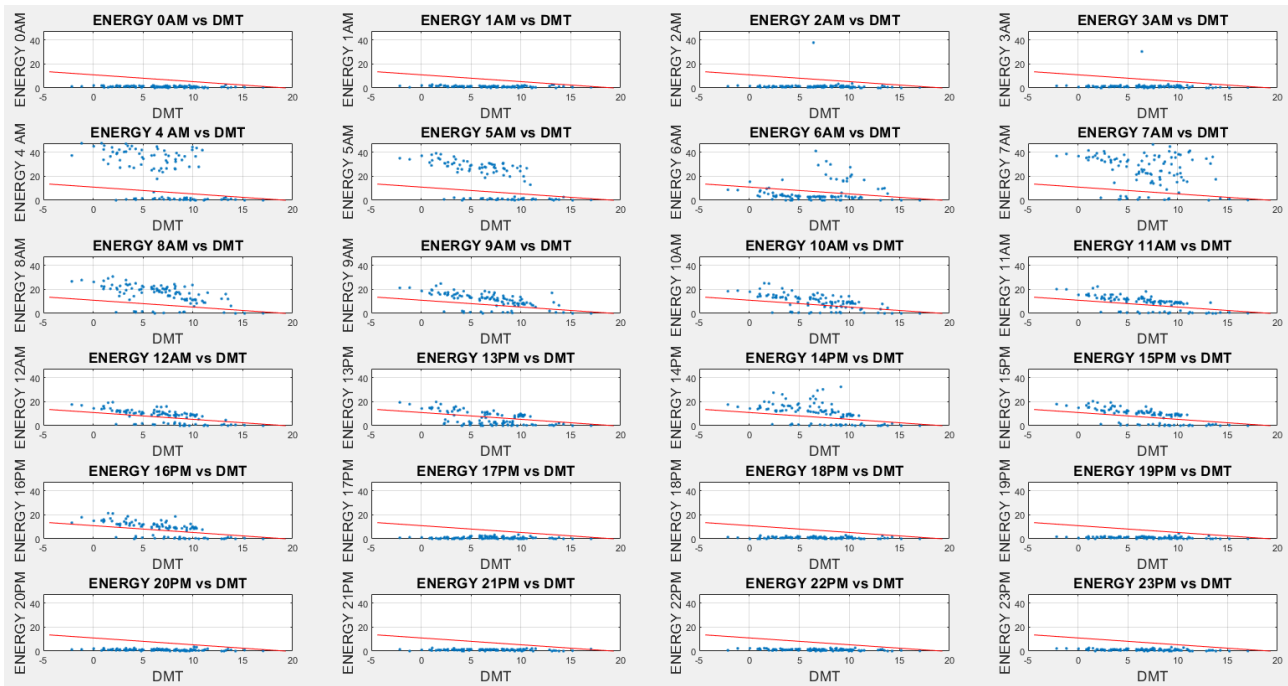
Figure 86 Hourly energy demands representations versus Tout. In red the boundary line between zone with SH and without SH. School 2 in Verona.

According to these considerations, it was implemented a little modification to adapt the general model presented in chapter 5 to the case of a school analysis. It will be explained in the following lines in what consists of this edit.

From figure 86 a lot of information about operation of the building can be obtained. In the figure, and in the analysis, it was considered only workdays, excluding weekends. In fact, demand was almost zero for the entire day and by considering them there would be a lot of points at null demand. They can produce some inaccuracies in the model's construction because they do not represent point of operation, but simply moments when the plant is completely turned off. For this reason, weekends are not considered in the analysis.

Another important fact to underline is that even if it would be unknown the intended use of the building, it would be easy to notice that it is not a residential building. This fact can be understood by observing period when there is consumption. In this building from 17.00 pm to 4.00 am there is neither DHW nor SH heat load, as demand is always null. A residential building, especially in winter, cannot stay 11 consecutive hours without heat provided keeping good comfort conditions. After this period without heating a big peak in demand is measured between 4.00 am and 8.00 am. This is due to the fact that after eleven hours without heat provided is necessary a big amount of energy to increase indoor temperature up to setpoint value. After that there is a period of average consumption, until people leave the building, around 17.00 pm. There is no information about the type of school. So, it is probable that students are present also during afternoon if it is a primary school or a kindergarten.

Another crucial difference between this building and a residential one, for example condominium 10 analysed before, regards moments without SH. In the previous case, a small amount of demand was still present, due to DHW consumption or losses into DHW circuit. Here consumption during these hours is almost null for a major part of the time. Of course, this can be imputed to the absence of people from the building.

However, even in hours of heat consumption there are points of null demand. By analysing dataset can be found that they are caused by days of extraordinary operation. In fact, it was cancelled weekends from database, but during wintertime there are holidays, where students are not at school, and some days (Christmas, boxing day, January 1st) where no people are present for sure. These days created null demand values even in hours where, in normal operation, there is SH and DHW heat load.

The next output of the code is presented in figure 87. Meaning of every single plot is explained in the previous paragraph about condominium 10. In this site it will be highlighted peculiarities of this specific building. Starting from the first plot in the upper left zone, it is possible to notice that blue points are much more regular than before. They represent hours without SH consumption and, presumably without DHW too. Green points are more disperse than before: the influence of high peak hours is especially important. High peaks are even more present than the previous case due to the larger number of hours without SH (when the plant starts indoor temperature will be much lower than setpoint value).
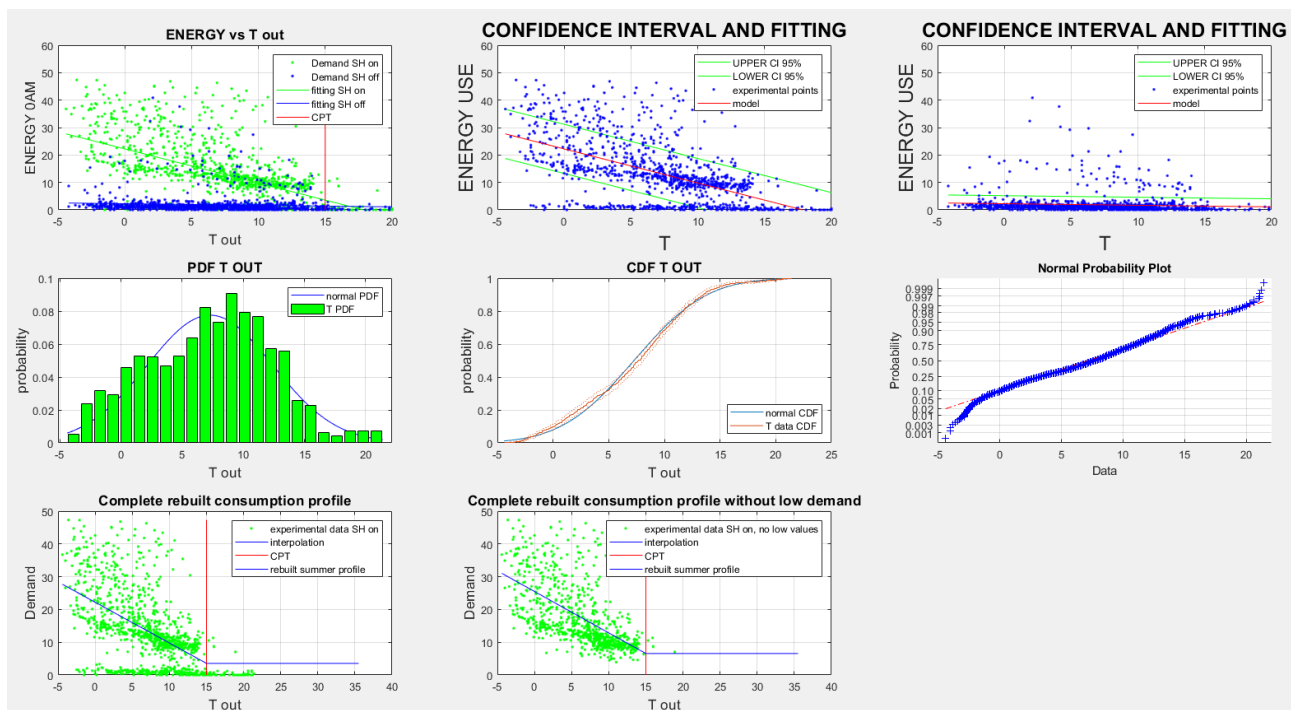
Figure 87 Hours with and without SH separation. Application of confidence intervals, check for normal distribution of temperature and ESC reconstruction in school 2, Verona.

From confidence interval's analysis it can be noticed that points in the upper zone are generated by peak moments, while values in the lower zone are caused by previously mentioned days of extraordinary operation. Even in this case outdoor temperature follows a normal distribution, but this happens for every building analysed in this zone because temperature values are the same for every structure, because the area of interest is the historic centre of Verona.

In the last plot (number 8 in figure 87) there is the representation of one edit that it was done for this case. By analysing the reconstructed ESC represent in plot 7, in fact, it is clear that blue line, the energy signature curve, is influenced by the null demand values. However, these points consider moment when heating plant is turned off an no people are in the building, so it is not correct to consider them part of the regular behaviour of the building. For this reason, it was decided to not consider values that are below a certain level of demand (as a boundary it was used CPD, represented by the blue line in the temperature independent part of the plot number 7).

In simpler words, it was decided to divide temperature dependent part and temperature independent part not with a temperature value but with a demand value. While in previous cases there be not so many differences between using CPD or CPT, in this case using a change point demand gives the possibility to consider only point where the heating system is in operation for the construction of the temperature dependent part of the ESC. In this way there is a better consideration of thermal losses and the real behaviour of the building when it is in operation.

Figure 88 Third image from the output. ESC shifted, chronological demand representation, normal distribution control for Tout at night, losses evaluation and representation of splitted demand. School 2, Verona.

To build the new temperature independent part of ESC it was recalculated demand when outdoor temperature is equal to CPT, and it was assumed that it is constant for every moment. To resume the steps of this new part of the analysis implemented here, once the plot number 7 is obtained it was used CPD value to exclude all experimental points < CPD. After that it was calculated a new interpolation line considering only new set of measured value. After that it was recalculated CPD using regression line equation and assuming Tout=CPT. However, there is not a substantial change in Q_{tot} value. By considering all points of demand there is $Q_{tot} = -1.24 \, kWh/°C$, while excluding null demand points the result is $Q_{tot} = -1.26 \, kWh/°C$. There is a minor difference, so this phase is optional. This is a good result because it means this model is not so sensitive to null demand points.

Analysing the next figure (number 88) there is a peculiarity respect to the previous building. The meaning of each plot is explained in the previous paragraph, so it will not be dwelled on these descriptions. Something interesting can be said regarding the evaluation of losses into DHW circuit. It was used values of consumption when SH was not in operation, as done for the previous analysis. However, in this case these points of demand were particularly low, and losses are accordingly low. In numerical terms, losses in DHW circuit are equal to 0.85 kWh. Compared to the peak load, they are about 2.5%, so slightly low, in relative terms, then previous analysed buildings.

Another relevant aspect is about demand splitted. As can be observed from the last two plots, there is an inconsistency in DHW profile. Because it is a school, in fact, it is unrealistic that DHW heat load increases when outdoor temperatures are close to CPT. As explained before, in fact, in a school DHW demand is much lower than in a residential building, so too high values for it are unlikely. For a better vision of this phenomenon, it was reported in figure 89 the last plot (number 9) of figure 88, and with a yellow circle it was underlined the problematic zone of DHW profile.

To solve this issue the idea was to modify the implementation of the model. In particular, one possibility was to do not shift down the ESC during the creation of the SH model. However, this idea could not be correct. By doing this thing, there would be no more dependency of demand on outdoor temperature. Better, the linear dependence is kept, but all values are shifted up by a coefficient that has no physical meaning.

So, the zone on the yellow circle in figure 89 can be explained by the demand itself and it is a demonstration that the method is consistent. It is possible to notice, in fact, that there are points with the same value of DHW demand also in the left part of the plot, with lower temperatures. It seems to be lower because there are less measured values in this zone. In addition to that, by doing some calculations it is possible to check if the value obtained for DHW consumption is likely or not.
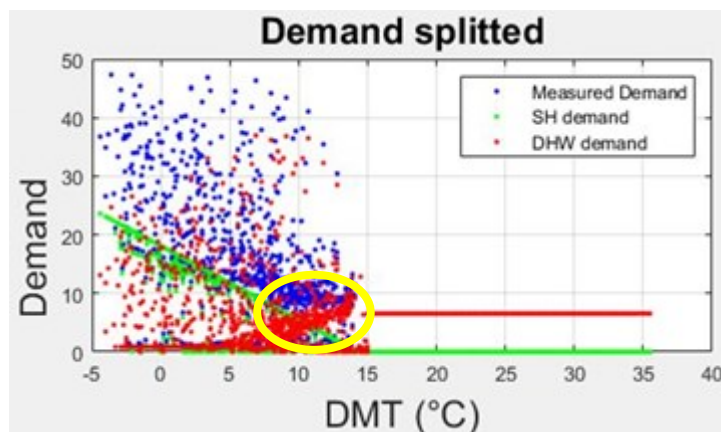
Figure 89 Splitted demand versus T out, School 2, Verona.

Assuming that there is an average consumption of 3 litres per person per day, that people are at school for 8 hours, and that the difference between temperature of water available at users and temperature of water from the aqueduct is about 30°C, it is possible to determine a sort of average standard energy consumed for DHW purpose. This value is about 0.013 kWh per person every hour. Considering an average hourly value of DHW calculated from the splitting model of 6.21 kWh, there would be about 480 people at school, which is a reasonable number for a school in historic centre of a big city like Verona. It is important to specify that these data are just to have a general idea, without detailed information it is impossible to obtain more accurate results. To conclude, the points highlighted with a yellow circle in figure 89 are consistent with the results of the analysis and the intended use of the building.
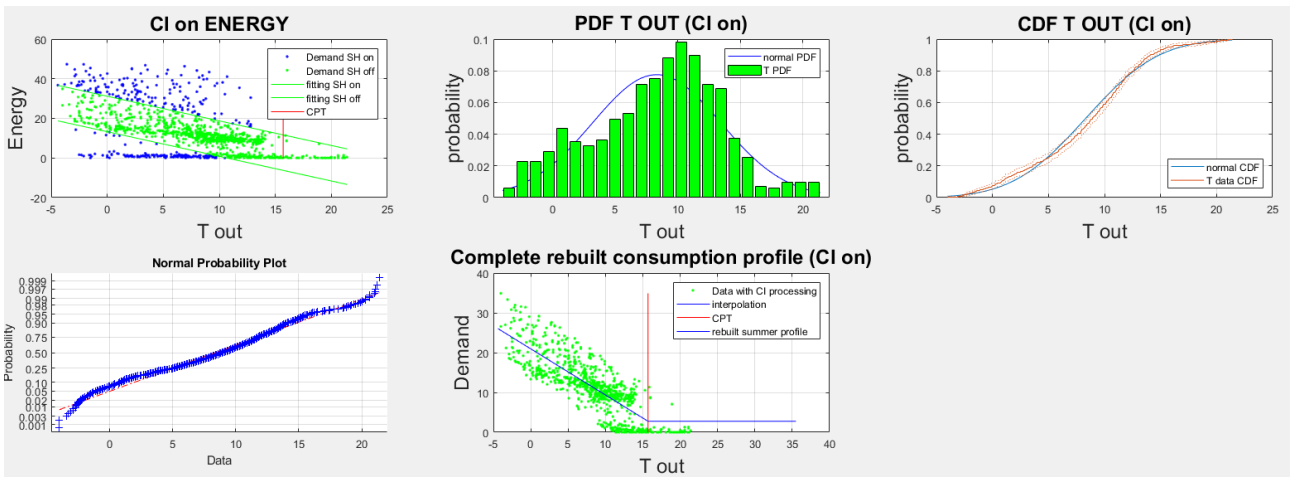
Figure 90 Output of analysis without values outside the confidence interval. from top left: representation of demand values with SH turned on and CI, control for normal distribution of Tout, reconstruction of new ESC. School 2, Verona.

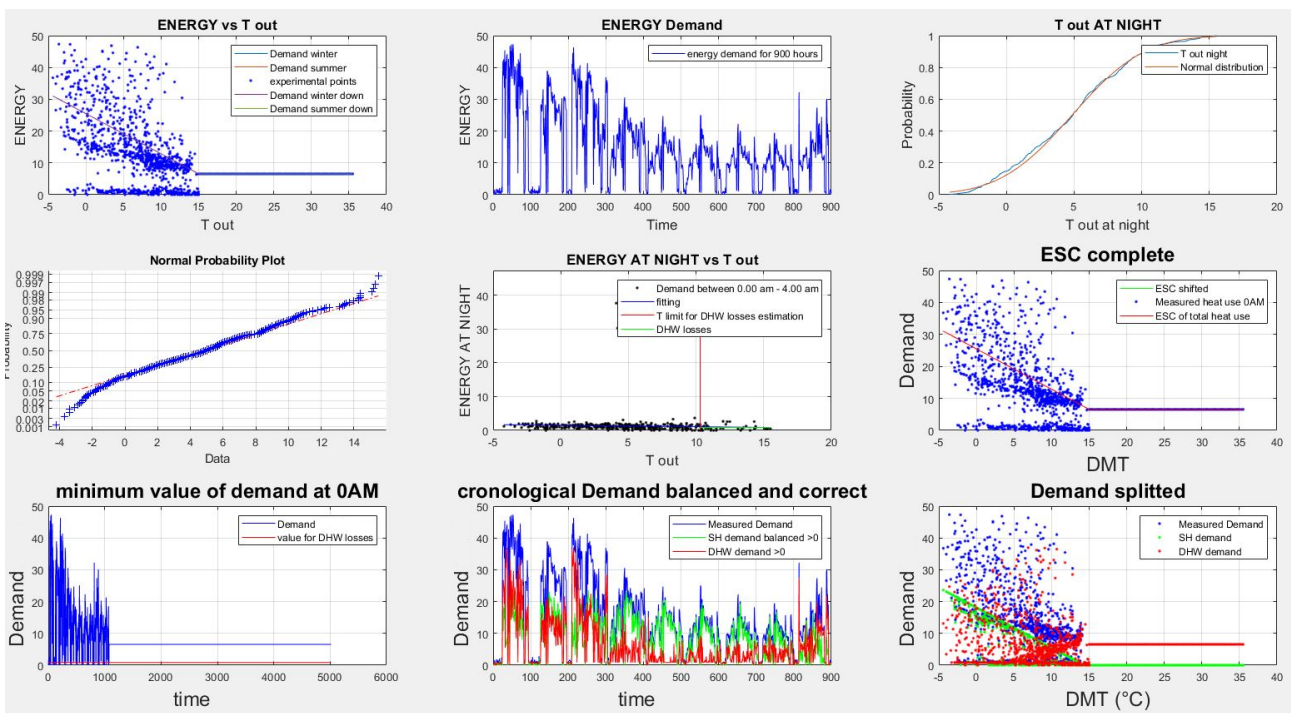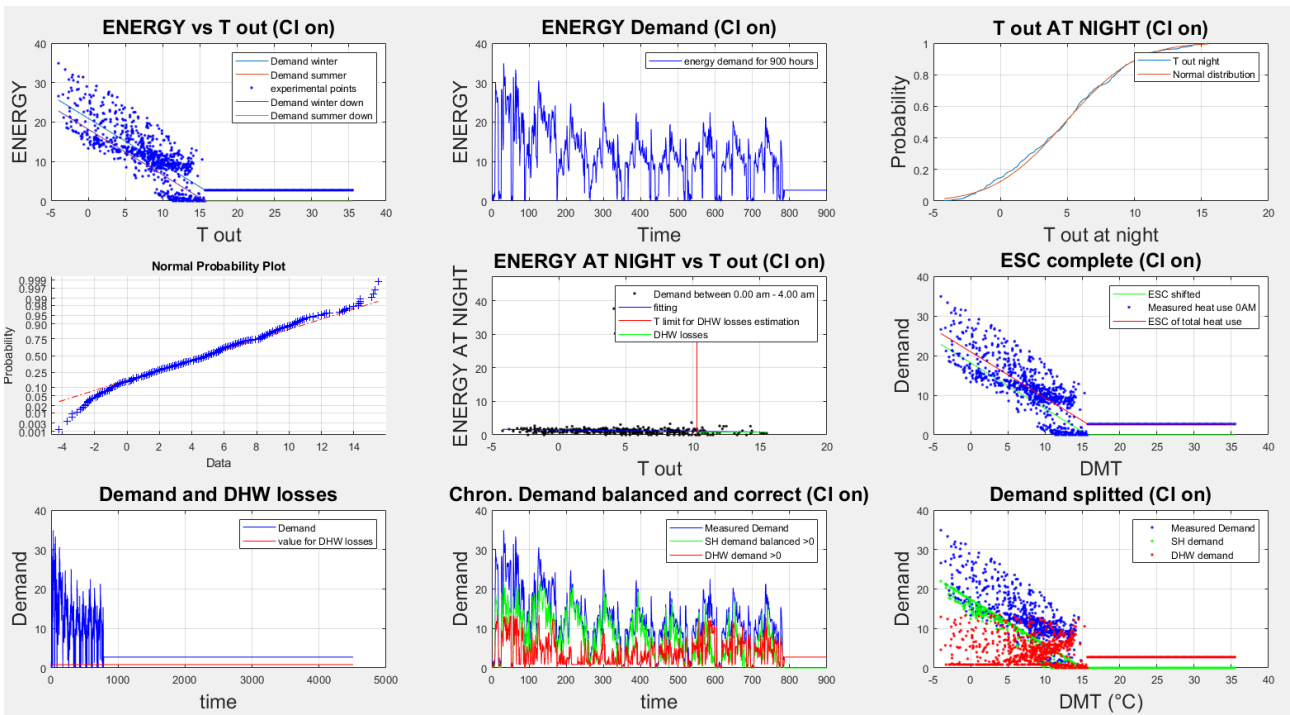*Figure 91 5th image from the output. ESC shifted, chronological demand representation, normal distribution control for Tout at night, losses evaluation and representation of splitted demand. CI filter. School 2, Verona.*

To conclude, in figures 90 and 91 it was provided results of the analysis with confidence interval filters. In this way, point of peak of demand and most of the values with null consumption are not considered and the profile of both SH and DHW follows more regular shape. Even in this case are true considerations done for in the previous building (paragraph 7.4) about the same issue. Mean values of DHW is almost the same as the previous one, but by considering only points inside confidence interval there are no more sudden peak or extraordinary high values for DHW.

With the school analysis it was also discussed the applicability of the model in the analysis of a building with a very irregular demand and with a profile hardly different than a residential building. Most of the other buildings in the dataset are condominiums, with a behaviour similar to the one discussed in paragraph 7.4. There are, also, schools and offices, which have similar demand profile in terms of behaviour of users, both are empty during weekends and during nights, both have relatively low DHW consumption per user, so from analysis it emerged that also offices, where the analysis was possible, present profiles and characteristics like the school discussed in 7.5. It is said "where the analysis was possible" because some building suffers of lack of data. Too much reading errors or missing data are present, for example, in hotel 1.
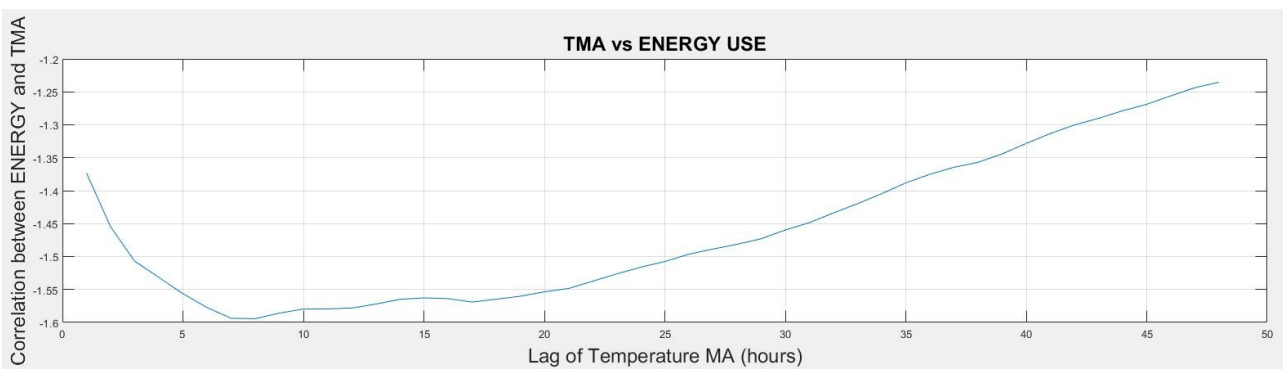
Figure 97 Effect of temperature lag on energy use of school 2, Verona.

Last output showed here is the analysis of correlation between energy consumption and shifted outdoor temperature, represented by different values of TMA. The graph is reported in figure 97. For this school it is evident that the hours with the highest correlation are the 7th and the 8th before the one considered. The thermal inertia of the building is quite high, but it is necessary to keep in mind the particular behaviour of the school in terms of operation of space heating plant.

Regarding other types of structures, for example rehab centre or train station, the profile of experimental data makes the analysis impossible. As can be observed in figure 92, the measured data present no recognizable trend. The linear interpolation does not present consistent results. Observing that plot, it would be possible that the DHN provides only DHW to the building, because the profile of experimental values is clear temperature independent, but this is just a hypothesis.
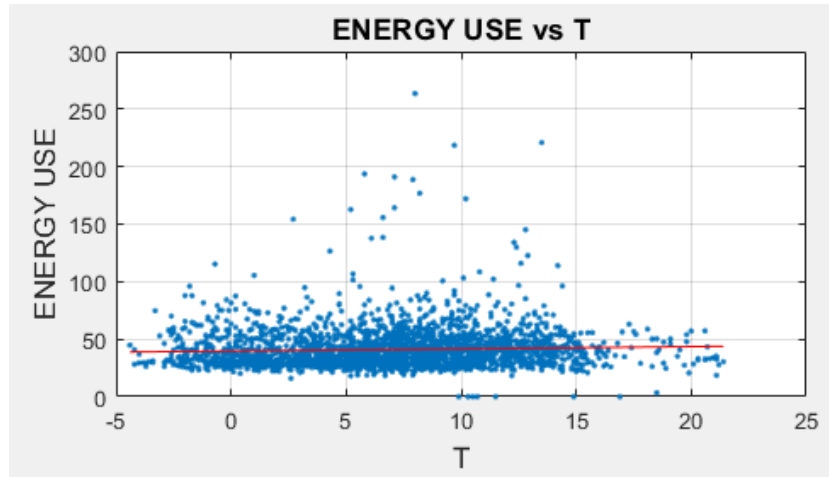
Figure 92 Measured values of demand versus Tout for the rehab centre in Verona.

Chapter 8: Future work and possibilities.

Thanks to the large field of applicability of the developed method and the high number of information obtainable from it, the future possibilities are limitless. It would be impossible to present all possible scenarios of future work, so it was decided to discuss just a few things, giving priority to the more feasible in the next period.

The first necessary operation is to apply this model in a dataset where DHW and SH heat consumptions are measured in a separate way. It would be extremely useful to have the possibility to check the accuracy of the model in terms of splitting operation. It would be possible to determine exactly the level of accuracy of the implemented tools, both in case of complete database and partial data. It would be enough to have just few weeks of separate measurements if a complete year is unfeasible. The best situation would be to have some weeks in summer and some weeks in winter. In this way the detailed analysis of DHW in summer could become more accurate and this information can be used to correct the DHW in the rest of the year, as explained in chapter 4.

Furthermore, thanks to the information about DHW consumption in winter, it would be possible to determine precisely values of tolerance and decreasing mentioned in paragraph 4.9.5 in an iterative way. In this way the method developed for correcting and flattening DHW profile in winter becomes more precise. The idea at the basis of this work is to try, starting from 0 and increasing by 0.01 at each iteration, different values of tolerance and decreasing numbers. After that, they could be used to correct winter profile and the error between corrected DHW profile and measured DHW consumption can be calculated hour per hour. Values of tolerance and decreasing numbers which minimize error are the most adapt to describe the behaviour of the building. In this way with the knowledge of few weeks of separate consumption in winter it is possible to create a model for obtaining DHW and SH in a more accurate way.

Talking about other possibilities, data that come from the application of the model have a large potential and can be used to describe well the energy performance of buildings. An advantage of this method is that no physical or geometrical information are necessary, so one of the most interesting possible applications is the study of historic buildings. This approach yields more comprehensive insights compared to relying solely on specific energy use in kWh/m^2, the most common data used in energy analysis of buildings. By providing information on the building's total loss term (Q_{tot}), change point temperature (CPT), losses in DHW circuit (E_{loss}) and consumption for SH and DHW heat loads, this model provides a more complete picture of the building's behaviour.

These data can be used in future works, for example, for the examination of renovation effects, such as changes in Q_{tot} and CPT before and after a renovation project. Also, these tools can be applied to check SH and DHW consumption before and after retrofitting operations. About this issue, these data can be used to decide which building has the highest necessity of retrofitting by identifying the one with worst thermal performances. They can be also used to see what the dominant consumption in the building is, in order to focus the attention of possible improvement and interventions on the most relevant type of heat load.

The use of this method has another advantage: it can effectively filter out the variations that result from different occupants' energy usage behaviours in the building. As a result, it

becomes possible to design effective packages of measures based on the technical performance of the buildings, rather than solely on how the current occupants use them. Thanks to this characteristic of the implemented tools, further works can use information obtained from this analysis to support housing companies' decision-making processes and provide valuable input for the national renovation strategy. Furthermore, from a housing company's perspective, this is a convenient way to identify malfunctions such as control problems and obtain an overview of the building stock.

Moreover, analysis' outputs could be also implemented in a DHN management software. It could be possible to obtain a detailed model of consumption for every building by starting only from information about global heat delivered in the corresponding substation. Thanks to the automaticity and the low computing time of the model, it is possible to easily implement this analysis in every building connected to the DHN. In this way it would be known for every hour the amount of DHW and SH necessary for every building, improving quality of service or making easier operations such as night setback for the manager of DHN. Because this analysis is conducted in a year and the demand depends on Tout, it is possible to use these data also to make forecasting. In this way, from manager perspective, it is possible to know, basing on day of the year and outdoor temperature, what is the DHW and SH load in every building connected to the grid.

As said in the first lines of the chapter, possibilities are endless. For users there is the possibility to know what the operation of the building in terms of DHW and SH is, to check losses in the circuit, to know where to act for possible interventions etc. For manager of the DHN there is the possibility to use these tools to build a detailed model of users and to know hour per hour SH and DHW heat load that must be provided. For housing companies, data available for energy audit are more precise and higher quality, without using more complex meters.

Chapter 9: Conclusions.

This thesis proposed a series of tools useful for the study of energy consumption in building. The strength of the proposed methodology is that it uses only global energy consumption and outdoor temperature, both with hourly resolution, as input for the entire analysis. There are data easily to obtain with the implementation of a simple meter in the building for energy consumption data and outdoor temperature values can be obtained from the closest weather station. The goal of the thesis was to obtain the highest possible amount of information regarding the building by knowing only these two inputs. The guideline that it was followed for the development of this methodology was the automaticity of the code and the possibility to apply it to a large variety of different building.

The outputs of the analysis include a detailed analysis hour per hour of global consumption of the building, with individuation of peak moments, extraordinary values, and eventually null demand hours. In addition to that the energy signature curve of the building, including information about the global heat losses coefficient Q_{tot}, change point temperature CPT, losses in DHW circuit. Furthermore, splitting hour per hour of global demand into SH and DHW heat load components. Then, a detailed analysis on consumption in workdays, Saturdays, and Sundays. In addition to that, a detailed analysis of DHW summer profile with the individuation of similar days (in terms of consumption) and of hours with minimum, medium, and peak heat load. Moreover, the development of a method to flatten DHW profile in winter to make it more like the summer one and, so, more convincing. It was also built a similar model to applicate in case of building with partial database (only some months during winter).

The tools were tested on two different case studies: Tartu and Verona. Both of them present data form buildings connected to a district heating network. However, differences are quite large. Tartu has a colder climate than Verona which means lower outdoor temperatures and higher heat required. Level of information about buildings was almost the same, with the only difference that in Verona intended use was known. Other types of geometrical, physical or occupants' information was unknown in both cases. The other significant difference is that every building in Verona presents a partial database, with only some month during winter season available. For this reason, in Tartu the complete method of analysis was applied, while in Verona it must be implemented only a partial part, developed for incomplete databases.

A detailed analysis of the outputs for the most relevant buildings is discussed in chapter 7. To resume, the method proposed gives consistent results in most of analysed building. It works very well in provides accurate information especially in buildings with quite regular behaviour in terms of heat load (residential buildings among others, where SH follows regular pattern and also DHW). In buildings with different intended use, for example schools or offices, it gives consistent results, but some trick is necessary. The most important to keep in mind is that, especially for the splitting operation between SH and DHW, the model provides consistent results for the most common profile of the building, but it struggles when there are moments of extraordinary peak operation. However, with a good interpretation of results it is possible to obtain consistent information even in this case.

Unlucky, it was not possible to check the performance of the method by applying it in a dataset where SH and DHW were measured in a separate way. However, thanks to the

method developed for the correction procedure of DHW profile, it is easy to adapt it to every building and to the specific operation of the examined structure.

One step of the future work is about the check of model's performances with this type of dataset. Furthermore, a different model for DHW and SH splitting could be introduced, to check differences between these approaches and see what is the most suitable one. In particular, one of the crucial issues is that there is a model for SH consumption, while DHW is calculated by subtracting SH from the measured hourly global demand. Implementing a model for estimating DHW consumption hour per hour could be one possibility to improve performances of the model.

Appendix.

A1 Complete database (Tartu case study)

"CI.m"

```matlab
%function per il CI
function [y_CI_up,y_CI_down]=CI(coeff_DEF_SH_0AM,T_SH_0AM,Demand_SH_0AM,i)

%valutazione punti con il modello lineare
SH_value_0AM=polyval(coeff_DEF_SH_0AM,T_SH_0AM);
n_lib_deg=length(SH_value_0AM)-2;

%caclolo parametri necessari
Se_SH_0AM=sqrt((n_lib_deg+1)*((std(Demand_SH_0AM)^2)-
(coeff_DEF_SH_0AM(1)^2)*std(T_SH_0AM))/(n_lib_deg));
%usare il valore di porbabilità cumulata (cumulative density function)
%nella t-student per includere tutti i valori all'interno della probabilità
%scelt, il 95% solitamente basta. Avendo un alto numero di gradi di libertà
%coincide con la distribuzione gaussiana.
t_stud_value=tcdf(0.95,n_lib_deg);
prod_factor=sqrt(1+(1/length(SH_value_0AM))+(((T_SH_0AM-
mean(T_SH_0AM)).^2)/(length(SH_value_0AM)*(std(T_SH_0AM)^2))));

%calcolo y degli intervalli di confidenza
y_CI_up=SH_value_0AM+Se_SH_0AM.*t_stud_value.*prod_factor;
y_CI_down=SH_value_0AM-Se_SH_0AM.*t_stud_value.*prod_factor;

fontSize=18;

%grafico
figure(i)
subplot(3,3,5)
% pos=[0.65 0.2 0.3 0.3];
% subplot('Position',pos)
plot(T_SH_0AM,y_CI_up,'g')
grid on
hold on
plot(T_SH_0AM,y_CI_down,'g')
plot(T_SH_0AM,Demand_SH_0AM,'.b')
plot(T_SH_0AM,SH_value_0AM,'r')
hold off
xlabel('T', 'FontSize', fontSize);
ylabel('ENERGY USE', 'FontSize', fontSize);
title('CONFIDENCE INTERVAL AND FITTING', 'FontSize', fontSize);
legend('UPPER CI 95%','LOWER CI 95%','experimental points','model')

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

end
```

"CI_year.m"

```matlab
%function per il CI year
function [y_CI_up,y_CI_down]=CI_year(coeff_DEF_SH_0AM,T_SH_0AM,Demand_SH_0AM,i)

%valutazione punti con il modello lineare
```

```matlab
SH_value_0AM=polyval(coeff_DEF_SH_0AM,T_SH_0AM);
n_lib_deg=length(SH_value_0AM)-2;

%caclolo parametri necessari
Se_SH_0AM=sqrt((n_lib_deg+1)*((std(Demand_SH_0AM)^2)-
(coeff_DEF_SH_0AM(1)^2)*std(T_SH_0AM))/(n_lib_deg));
%usare il valore di porbabilità cumulata (cumulative density function)
%nella t-student per includere tutti i valori all'interno della probabilità
%scelt, il 95% solitamente basta. Avendo un alto numero di gradi di libertà
%coincide con la distribuzione gaussiana.
t_stud_value=tcdf(0.95,n_lib_deg);
prod_factor=sqrt(1+(1/length(SH_value_0AM))+(((T_SH_0AM-
mean(T_SH_0AM)).^2)/(length(SH_value_0AM)*(std(T_SH_0AM)^2))));

%calcolo y degli intervalli di confidenza
y_CI_up=SH_value_0AM+Se_SH_0AM.*t_stud_value.*prod_factor;
y_CI_down=SH_value_0AM-Se_SH_0AM.*t_stud_value.*prod_factor;

fontSize=18;

%grafico
figure(i)
% pos=[0.65 0.2 0.3 0.3];
% subplot('Position',pos)
subplot(3,3,5)
plot(T_SH_0AM,y_CI_up,'g')
grid on
hold on
plot(T_SH_0AM,y_CI_down,'g')
plot(T_SH_0AM,Demand_SH_0AM,'.b')
plot(T_SH_0AM,SH_value_0AM,'r')
hold off
xlabel('T', 'FontSize', fontSize);
ylabel('ENERGY USE', 'FontSize', fontSize);
title('CONFIDENCE INTERVAL AND FITTING', 'FontSize', fontSize);
legend('UPPER CI 95%','LOWER CI 95%','experimental points','model')

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

end
```

"ESC_10922.m"

```matlab
%ESC DEFINITIVO
close all
clear all
clc

%IMPORTARE I DATI DAL FILE EXCEL DESIDERATO
ENERGY=xlsread('10922 no doppi.xlsx','O2:O8760');
%importare dati non doppi, processare in excel prima
HOURS=xlsread('10922 no doppi.xlsx','P2:P8760');

%% CREAZIONE VETTORI DIVISI PER ORA

%ENERGY 0AM
[ENERGY_0AM]=HOUR_split(ENERGY,1);                          %0AM
è la ora 1 del giorno, var_0AM=1
```

```
%ENERGY 1AM
[ENERGY_1AM]=HOUR_split(ENERGY,2);                                          %1AM
è la ora 2 del giorno, var_0AM=2


%ENERGY 2AM
[ENERGY_2AM]=HOUR_split(ENERGY,3);                                          %1AM
è la ora 3 del giorno, var_0AM=3


%ENERGY 3AM
[ENERGY_3AM]=HOUR_split(ENERGY,4);                                          %1AM
è la ora 4 del giorno, var_0AM=4


%ENERGY 4AM
[ENERGY_4AM]=HOUR_split(ENERGY,5);                                          %1AM
è la ora 5 del giorno, var_0AM=5


%ENERGY 5AM
[ENERGY_5AM]=HOUR_split(ENERGY,6);                                          %1AM
è la ora 6 del giorno, var_0AM=6


%ENERGY 6AM
[ENERGY_6AM]=HOUR_split(ENERGY,7);                                          %1AM
è la ora 7 del giorno, var_0AM=7


%ENERGY 7AM
[ENERGY_7AM]=HOUR_split(ENERGY,8);                                          %1AM
è la ora 8 del giorno, var_0AM=8


%ENERGY 8AM
[ENERGY_8AM]=HOUR_split(ENERGY,9);                                          %1AM
è la ora 9 del giorno, var_0AM=9


%ENERGY 9AM
[ENERGY_9AM]=HOUR_split(ENERGY,10);                                         %1AM
è la ora 10 del giorno, var_0AM=10


%ENERGY 10AM
[ENERGY_10AM]=HOUR_split(ENERGY,11);                                        %1AM
è la ora 11 del giorno, var_0AM=11


%ENERGY 11AM
[ENERGY_11AM]=HOUR_split(ENERGY,12);                                        %1AM
è la ora 12 del giorno, var_0AM=12


%ENERGY 12AM
[ENERGY_12AM]=HOUR_split(ENERGY,13);                                        %1AM
è la ora 13 del giorno, var_0AM=13


%ENERGY 13PM
[ENERGY_13PM]=HOUR_split(ENERGY,14);                                        %1AM
è la ora 14 del giorno, var_0AM=14


%ENERGY 14PM
[ENERGY_14PM]=HOUR_split(ENERGY,15);                                        %1AM
è la ora 15 del giorno, var_0AM=15


%ENERGY 15PM
```

```matlab
[ENERGY_15PM]=HOUR_split(ENERGY,16);                                    %1AM
è la ora 16 del giorno, var_0AM=16


%ENERGY 16PM
[ENERGY_16PM]=HOUR_split(ENERGY,17);                                    %1AM
è la ora 17 del giorno, var_0AM=17


%ENERGY 17PM
[ENERGY_17PM]=HOUR_split(ENERGY,18);                                    %1AM
è la ora 18 del giorno, var_0AM=18


%ENERGY 18PM
[ENERGY_18PM]=HOUR_split(ENERGY,19);                                    %1AM
è la ora 19 del giorno, var_0AM=19


%ENERGY 19PM
[ENERGY_19PM]=HOUR_split(ENERGY,20);                                    %1AM
è la ora 20 del giorno, var_0AM=20


%ENERGY 20PM
[ENERGY_20PM]=HOUR_split(ENERGY,21);                                    %1AM
è la ora 21 del giorno, var_0AM=21


%ENERGY 21PM
[ENERGY_21PM]=HOUR_split(ENERGY,22);                                    %1AM
è la ora 22 del giorno, var_0AM=22


%ENERGY 22PM
[ENERGY_22PM]=HOUR_split(ENERGY,23);                                    %1AM
è la ora 23 del giorno, var_0AM=23


%ENERGY 23PM
[ENERGY_23PM]=HOUR_split_11PM(ENERGY,24);                               %1AM
è la ora 24 del giorno, var_0AM=24
ENERGY_23PM(365,1)=ENERGY_0AM(1,1);

%% calcolo della mean daily temperature
T=xlsread('Clima_Hour_2019 ex.xlsx','I2:I8760');
DMT=[];
scr=1;

for j=1:364

dailyMT=(T(scr,1)+T(scr+1,1)+T(scr+2,1)+T(scr+3,1)+T(scr+4,1)+T(scr+5,1)+T(scr+6
,1)+T(scr+7,1)+T(scr+8,1)+T(scr+9,1)+
T(scr+10,1)+T(scr+11,1)+T(scr+12,1)+T(scr+13,1)+T(scr+14,1)+T(scr+15,1)+T(scr+16
,1)+T(scr+17,1)+T(scr+18,1)+T(scr+19,1)+T(scr+20,1)+T(scr+21,1)+T(scr+22,1)+T(sc
r+23,1))/24;
    DMT(j,1)=dailyMT;


    scr=scr+24;
end


j=365;

dailyMT=(T(scr,1)+T(scr+1,1)+T(scr+2,1)+T(scr+3,1)+T(scr+4,1)+T(scr+5,1)+T(scr+6
,1)+T(scr+7,1)+T(scr+8,1)+T(scr+9,1)+
T(scr+10,1)+T(scr+11,1)+T(scr+12,1)+T(scr+13,1)+T(scr+14,1)+T(scr+15,1)+T(scr+16
,1)+T(scr+17,1)+T(scr+18,1)+T(scr+19,1)+T(scr+20,1)+T(scr+21,1)+T(scr+22,1))/23;
 DMT(j,1)=dailyMT;
```

```matlab
%% plot Tdailymean (DMT) vs ENERGY_0AM
%SE OK POSSIBILE IMPLEMENTARE LA CREAZIONE DI 24 FIGURE, UNA PER OGNI ORA
%DEL GIORNO

fontSize = 18;
markerSize = 20;

figure (1)
subplot(3,3,1)
plot(DMT,ENERGY_0AM,'*')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 0AM', 'FontSize', fontSize);
title('ENERGY 0AM vs DMT', 'FontSize', fontSize);

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (2)
subplot(3,3,1)
plot(DMT,ENERGY_1AM,'*')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 1AM', 'FontSize', fontSize);
title('ENERGY 1AM vs DMT', 'FontSize', fontSize);

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (3)
subplot(3,3,1)
plot(DMT,ENERGY_2AM,'*')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 2AM', 'FontSize', fontSize);
title('ENERGY 2AM vs DMT', 'FontSize', fontSize);

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (4)
subplot(3,3,1)
plot(DMT,ENERGY_3AM,'*')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 3AM', 'FontSize', fontSize);
title('ENERGY 3AM vs DMT', 'FontSize', fontSize);

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (5)
subplot(3,3,1)
plot(DMT,ENERGY_4AM,'*')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 4 AM', 'FontSize', fontSize);
title('ENERGY 4 AM vs DMT', 'FontSize', fontSize);
```

```matlab
% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (6)
subplot(3,3,1)
plot(DMT,ENERGY_5AM,'*')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 5AM', 'FontSize', fontSize);
title('ENERGY 5AM vs DMT', 'FontSize', fontSize);

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (7)
subplot(3,3,1)
plot(DMT,ENERGY_6AM,'*')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 6AM', 'FontSize', fontSize);
title('ENERGY 6AM vs DMT', 'FontSize', fontSize);

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (8)
subplot(3,3,1)
plot(DMT,ENERGY_7AM,'*')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 7AM', 'FontSize', fontSize);
title('ENERGY 7AM vs DMT', 'FontSize', fontSize);

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (9)
subplot(3,3,1)
plot(DMT,ENERGY_8AM,'*')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 8AM', 'FontSize', fontSize);
title('ENERGY 8AM vs DMT', 'FontSize', fontSize);

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (10)
subplot(3,3,1)
plot(DMT,ENERGY_9AM,'*')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 9AM', 'FontSize', fontSize);
title('ENERGY 9AM vs DMT', 'FontSize', fontSize);

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (11)
subplot(3,3,1)
```

```matlab
plot(DMT,ENERGY_10AM,'*')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 10AM', 'FontSize', fontSize);
title('ENERGY 10AM vs DMT', 'FontSize', fontSize);

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (12)
subplot(3,3,1)
plot(DMT,ENERGY_11AM,'*')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 11AM', 'FontSize', fontSize);
title('ENERGY 11AM vs DMT', 'FontSize', fontSize);

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (13)
subplot(3,3,1)
plot(DMT,ENERGY_12AM,'*')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 12AM', 'FontSize', fontSize);
title('ENERGY 12AM vs DMT', 'FontSize', fontSize);

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (14)
subplot(3,3,1)
plot(DMT,ENERGY_13PM,'*')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 13PM', 'FontSize', fontSize);
title('ENERGY 13PM vs DMT', 'FontSize', fontSize);

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (15)
subplot(3,3,1)
plot(DMT,ENERGY_14PM,'*')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 14PM', 'FontSize', fontSize);
title('ENERGY 14PM vs DMT', 'FontSize', fontSize);

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (16)
subplot(3,3,1)
plot(DMT,ENERGY_15PM,'*')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 15PM', 'FontSize', fontSize);
title('ENERGY 15PM vs DMT', 'FontSize', fontSize);
```

```matlab
% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (17)
subplot(3,3,1)
plot(DMT,ENERGY_16PM,'*')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 16PM', 'FontSize', fontSize);
title('ENERGY 16PM vs DMT', 'FontSize', fontSize);

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (18)
subplot(3,3,1)
plot(DMT,ENERGY_17PM,'*')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 17PM', 'FontSize', fontSize);
title('ENERGY 17PM vs DMT', 'FontSize', fontSize);

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (19)
subplot(3,3,1)
plot(DMT,ENERGY_18PM,'*')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 18PM', 'FontSize', fontSize);
title('ENERGY 18PM vs DMT', 'FontSize', fontSize);

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (20)
subplot(3,3,1)
plot(DMT,ENERGY_19PM,'*')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 19PM', 'FontSize', fontSize);
title('ENERGY 19PM vs DMT', 'FontSize', fontSize);

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (21)
subplot(3,3,1)
plot(DMT,ENERGY_20PM,'*')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 20PM', 'FontSize', fontSize);
title('ENERGY 20PM vs DMT', 'FontSize', fontSize);

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (22)
```

```matlab
subplot(3,3,1)
plot(DMT,ENERGY_21PM,'*')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 21PM', 'FontSize', fontSize);
title('ENERGY 21PM vs DMT', 'FontSize', fontSize);

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (23)
subplot(3,3,1)
plot(DMT,ENERGY_22PM,'*')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 22PM', 'FontSize', fontSize);
title('ENERGY 22PM vs DMT', 'FontSize', fontSize);

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (24)
subplot(3,3,1)
plot(DMT,ENERGY_23PM,'*')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 23PM', 'FontSize', fontSize);
title('ENERGY 23PM vs DMT', 'FontSize', fontSize);

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

%% CALCOLO ESC (energy signature curve)
%nella function ESC_BETA ci sono già i comandi per la creazione del
%grafico.

% ENERGY_0AM
[BETA_0AM,TEMP_0AM]=ESC_BETA(ENERGY_0AM,DMT,1);

%ENERGY_1AM
[BETA_1AM,TEMP_1AM]=ESC_BETA(ENERGY_1AM,DMT,2);

%ENERGY_2AM
[BETA_2AM,TEMP_2AM]=ESC_BETA(ENERGY_2AM,DMT,3);

%ENERGY_3AM
[BETA_3AM,TEMP_3AM]=ESC_BETA(ENERGY_3AM,DMT,4);

%ENERGY_4AM
[BETA_4AM,TEMP_4AM]=ESC_BETA(ENERGY_4AM,DMT,5);

%ENERGY_5AM
[BETA_5AM,TEMP_5AM]=ESC_BETA(ENERGY_5AM,DMT,6);

%ENERGY_6AM
[BETA_6AM,TEMP_6AM]=ESC_BETA(ENERGY_6AM,DMT,7);

%ENERGY_7AM
[BETA_7AM,TEMP_7AM]=ESC_BETA(ENERGY_7AM,DMT,8);
```

```matlab
%ENERGY_8AM
[BETA_8AM,TEMP_8AM]=ESC_BETA(ENERGY_8AM,DMT,9);

%ENERGY_9AM
[BETA_9AM,TEMP_9AM]=ESC_BETA(ENERGY_9AM,DMT,10);

%ENERGY_10AM
[BETA_10AM,TEMP_10AM]=ESC_BETA(ENERGY_10AM,DMT,11);

%ENERGY_11AM
[BETA_11AM,TEMP_11AM]=ESC_BETA(ENERGY_11AM,DMT,12);

%ENERGY_12AM
[BETA_12AM,TEMP_12AM]=ESC_BETA(ENERGY_12AM,DMT,13);

%ENERGY_13PM
[BETA_13PM,TEMP_13PM]=ESC_BETA(ENERGY_13PM,DMT,14);

%ENERGY_14PM
[BETA_14PM,TEMP_14PM]=ESC_BETA(ENERGY_14PM,DMT,15);

%ENERGY_15PM
[BETA_15PM,TEMP_15PM]=ESC_BETA(ENERGY_15PM,DMT,16);

%ENERGY_16PM
[BETA_16PM,TEMP_16PM]=ESC_BETA(ENERGY_16PM,DMT,17);

%ENERGY_17PM
[BETA_17PM,TEMP_17PM]=ESC_BETA(ENERGY_17PM,DMT,18);

%ENERGY_18PM
[BETA_18PM,TEMP_18PM]=ESC_BETA(ENERGY_18PM,DMT,19);

%ENERGY_19PM
[BETA_19PM,TEMP_19PM]=ESC_BETA(ENERGY_19PM,DMT,20);

%ENERGY_20PM
[BETA_20PM,TEMP_20PM]=ESC_BETA(ENERGY_20PM,DMT,21);

%ENERGY_21PM
[BETA_21PM,TEMP_21PM]=ESC_BETA(ENERGY_21PM,DMT,22);

%ENERGY_22PM
[BETA_22PM,TEMP_22PM]=ESC_BETA(ENERGY_22PM,DMT,23);

%ENERGY_23PM
[BETA_23PM,TEMP_23PM]=ESC_BETA(ENERGY_23PM,DMT,24);

%% CPT definition
%analize beta plot to find it. Where beta starts decreasing
CPT_0AM=11.5;
CPT_1AM=11.6;
CPT_2AM=11.5;
CPT_3AM=11.7;
CPT_4AM=11.5;
CPT_5AM=11.7;
CPT_6AM=11.9;
CPT_7AM=11.8;
```

```matlab
 CPT_8AM=11.3;
 CPT_9AM=11.3;
 CPT_10AM=10.9;
 CPT_11AM=11.3;
 CPT_12AM=11.2;
 CPT_13PM=11.3;
 CPT_14PM=11.3;
 CPT_15PM=11.7;
 CPT_16PM=11.6;
 CPT_17PM=11.4;
 CPT_18PM=11.6;
 CPT_19PM=11.5;
 CPT_20PM=11.5;
 CPT_21PM=11.5;
 CPT_22PM=11.8;
 CPT_23PM=11.7;
%% SPLITTING DHW E SH+DHW


%ENERGY 0AM
[T_SH_0AM,Demand_SH_0AM,T_DHW_0AM,Demand_DHW_0AM,coeff_DEF_SH_0AM,coeff_DEF_DHW_
0AM]=ESC_DIVISION(CPT_0AM,ENERGY_0AM,DMT,1);


%ENERGY 1AM
[T_SH_1AM,Demand_SH_1AM,T_DHW_1AM,Demand_DHW_1AM,coeff_DEF_SH_1AM,coeff_DEF_DHW_
1AM]=ESC_DIVISION(CPT_1AM,ENERGY_1AM,DMT,2);


%ENERGY 2AM
[T_SH_2AM,Demand_SH_2AM,T_DHW_2AM,Demand_DHW_2AM,coeff_DEF_SH_2AM,coeff_DEF_DHW_
2AM]=ESC_DIVISION(CPT_2AM,ENERGY_2AM,DMT,3);


%ENERGY 3AM
[T_SH_3AM,Demand_SH_3AM,T_DHW_3AM,Demand_DHW_3AM,coeff_DEF_SH_3AM,coeff_DEF_DHW_
3AM]=ESC_DIVISION(CPT_3AM,ENERGY_3AM,DMT,4);


%ENERGY 4AM
[T_SH_4AM,Demand_SH_4AM,T_DHW_4AM,Demand_DHW_4AM,coeff_DEF_SH_4AM,coeff_DEF_DHW_
4AM]=ESC_DIVISION(CPT_4AM,ENERGY_4AM,DMT,5);


%ENERGY 5AM
[T_SH_5AM,Demand_SH_5AM,T_DHW_5AM,Demand_DHW_5AM,coeff_DEF_SH_5AM,coeff_DEF_DHW_
5AM]=ESC_DIVISION(CPT_5AM,ENERGY_5AM,DMT,6);


%ENERGY 6AM
[T_SH_6AM,Demand_SH_6AM,T_DHW_6AM,Demand_DHW_6AM,coeff_DEF_SH_6AM,coeff_DEF_DHW_
6AM]=ESC_DIVISION(CPT_6AM,ENERGY_6AM,DMT,7);


%ENERGY 7AM
[T_SH_7AM,Demand_SH_7AM,T_DHW_7AM,Demand_DHW_7AM,coeff_DEF_SH_7AM,coeff_DEF_DHW_
7AM]=ESC_DIVISION(CPT_7AM,ENERGY_7AM,DMT,8);


%ENERGY 8AM
[T_SH_8AM,Demand_SH_8AM,T_DHW_8AM,Demand_DHW_8AM,coeff_DEF_SH_8AM,coeff_DEF_DHW_
8AM]=ESC_DIVISION(CPT_8AM,ENERGY_8AM,DMT,9);


%ENERGY 9AM
[T_SH_9AM,Demand_SH_9AM,T_DHW_9AM,Demand_DHW_9AM,coeff_DEF_SH_9AM,coeff_DEF_DHW_
9AM]=ESC_DIVISION(CPT_9AM,ENERGY_9AM,DMT,10);


%ENERGY 10AM
```

```matlab
[T_SH_10AM,Demand_SH_10AM,T_DHW_10AM,Demand_DHW_10AM,coeff_DEF_SH_10AM,coeff_DEF
_DHW_10AM]=ESC_DIVISION(CPT_10AM,ENERGY_10AM,DMT,11);

%ENERGY 11AM
[T_SH_11AM,Demand_SH_11AM,T_DHW_11AM,Demand_DHW_11AM,coeff_DEF_SH_11AM,coeff_DEF
_DHW_11AM]=ESC_DIVISION(CPT_11AM,ENERGY_11AM,DMT,12);

%ENERGY 12AM
[T_SH_12AM,Demand_SH_12AM,T_DHW_12AM,Demand_DHW_12AM,coeff_DEF_SH_12AM,coeff_DEF
_DHW_12AM]=ESC_DIVISION(CPT_12AM,ENERGY_12AM,DMT,13);

%ENERGY 13PM
[T_SH_13PM,Demand_SH_13PM,T_DHW_13PM,Demand_DHW_13PM,coeff_DEF_SH_13PM,coeff_DEF
_DHW_13PM]=ESC_DIVISION(CPT_13PM,ENERGY_13PM,DMT,14);

%ENERGY 14PM
[T_SH_14PM,Demand_SH_14PM,T_DHW_14PM,Demand_DHW_14PM,coeff_DEF_SH_14PM,coeff_DEF
_DHW_14PM]=ESC_DIVISION(CPT_14PM,ENERGY_14PM,DMT,15);

%ENERGY 15PM
[T_SH_15PM,Demand_SH_15PM,T_DHW_15PM,Demand_DHW_15PM,coeff_DEF_SH_15PM,coeff_DEF
_DHW_15PM]=ESC_DIVISION(CPT_15PM,ENERGY_15PM,DMT,16);

%ENERGY 16PM
[T_SH_16PM,Demand_SH_16PM,T_DHW_16PM,Demand_DHW_16PM,coeff_DEF_SH_16PM,coeff_DEF
_DHW_16PM]=ESC_DIVISION(CPT_16PM,ENERGY_16PM,DMT,17);

%ENERGY 17PM
[T_SH_17PM,Demand_SH_17PM,T_DHW_17PM,Demand_DHW_17PM,coeff_DEF_SH_17PM,coeff_DEF
_DHW_17PM]=ESC_DIVISION(CPT_17PM,ENERGY_17PM,DMT,18);

%ENERGY 18PM
[T_SH_18PM,Demand_SH_18PM,T_DHW_18PM,Demand_DHW_18PM,coeff_DEF_SH_18PM,coeff_DEF
_DHW_18PM]=ESC_DIVISION(CPT_18PM,ENERGY_18PM,DMT,19);

%ENERGY 19PM
[T_SH_19PM,Demand_SH_19PM,T_DHW_19PM,Demand_DHW_19PM,coeff_DEF_SH_19PM,coeff_DEF
_DHW_19PM]=ESC_DIVISION(CPT_19PM,ENERGY_19PM,DMT,20);

%ENERGY 20PM
[T_SH_20PM,Demand_SH_20PM,T_DHW_20PM,Demand_DHW_20PM,coeff_DEF_SH_20PM,coeff_DEF
_DHW_20PM]=ESC_DIVISION(CPT_20PM,ENERGY_20PM,DMT,21);

%ENERGY 21PM
[T_SH_21PM,Demand_SH_21PM,T_DHW_21PM,Demand_DHW_21PM,coeff_DEF_SH_21PM,coeff_DEF
_DHW_21PM]=ESC_DIVISION(CPT_21PM,ENERGY_21PM,DMT,22);

%ENERGY 22PM
[T_SH_22PM,Demand_SH_22PM,T_DHW_22PM,Demand_DHW_22PM,coeff_DEF_SH_22PM,coeff_DEF
_DHW_22PM]=ESC_DIVISION(CPT_22PM,ENERGY_22PM,DMT,23);

%ENERGY 23PM
[T_SH_23PM,Demand_SH_23PM,T_DHW_23PM,Demand_DHW_23PM,coeff_DEF_SH_23PM,coeff_DEF
_DHW_23PM]=ESC_DIVISION(CPT_23PM,ENERGY_23PM,DMT,24);
%% Equation CONTINUE IN X=CPT, from mathematical model with imposed continuity

%ENERGY 0AM
```

```
[Demand_NOR_SH_0AM,Demand_NOR_DHW_0AM,R2_nor_SH_0AM,R2_nor_DHW_0AM]=ESC_CONT_CPT
(coeff_DEF_SH_0AM,coeff_DEF_DHW_0AM,T_SH_0AM,T_DHW_0AM,CPT_0AM,Demand_SH_0AM,Dem
and_DHW_0AM,1);
```

%ENERGY 1AM
```
[Demand_NOR_SH_1AM,Demand_NOR_DHW_1AM,R2_nor_SH_1AM,R2_nor_DHW_1AM]=ESC_CONT_CPT
(coeff_DEF_SH_1AM,coeff_DEF_DHW_1AM,T_SH_1AM,T_DHW_1AM,CPT_1AM,Demand_SH_1AM,Dem
and_DHW_1AM,2);
```

%ENERGY 2AM
```
[Demand_NOR_SH_2AM,Demand_NOR_DHW_2AM,R2_nor_SH_2AM,R2_nor_DHW_2AM]=ESC_CONT_CPT
(coeff_DEF_SH_2AM,coeff_DEF_DHW_2AM,T_SH_2AM,T_DHW_2AM,CPT_2AM,Demand_SH_2AM,Dem
and_DHW_2AM,3);
```

%ENERGY 3AM
```
[Demand_NOR_SH_3AM,Demand_NOR_DHW_3AM,R2_nor_SH_3AM,R2_nor_DHW_3AM]=ESC_CONT_CPT
(coeff_DEF_SH_3AM,coeff_DEF_DHW_3AM,T_SH_3AM,T_DHW_3AM,CPT_3AM,Demand_SH_3AM,Dem
and_DHW_3AM,4);
```

%ENERGY 4AM
```
[Demand_NOR_SH_4AM,Demand_NOR_DHW_4AM,R2_nor_SH_4AM,R2_nor_DHW_4AM]=ESC_CONT_CPT
(coeff_DEF_SH_4AM,coeff_DEF_DHW_4AM,T_SH_4AM,T_DHW_4AM,CPT_4AM,Demand_SH_4AM,Dem
and_DHW_4AM,5);
```

%ENERGY 5AM
```
[Demand_NOR_SH_5AM,Demand_NOR_DHW_5AM,R2_nor_SH_5AM,R2_nor_DHW_5AM]=ESC_CONT_CPT
(coeff_DEF_SH_5AM,coeff_DEF_DHW_5AM,T_SH_5AM,T_DHW_5AM,CPT_5AM,Demand_SH_5AM,Dem
and_DHW_5AM,6);
```

%ENERGY 6AM
```
[Demand_NOR_SH_6AM,Demand_NOR_DHW_6AM,R2_nor_SH_6AM,R2_nor_DHW_6AM]=ESC_CONT_CPT
(coeff_DEF_SH_6AM,coeff_DEF_DHW_6AM,T_SH_6AM,T_DHW_6AM,CPT_6AM,Demand_SH_6AM,Dem
and_DHW_6AM,7);
```

%ENERGY 7AM
```
[Demand_NOR_SH_7AM,Demand_NOR_DHW_7AM,R2_nor_SH_7AM,R2_nor_DHW_7AM]=ESC_CONT_CPT
(coeff_DEF_SH_7AM,coeff_DEF_DHW_7AM,T_SH_7AM,T_DHW_7AM,CPT_7AM,Demand_SH_7AM,Dem
and_DHW_7AM,8);
```

%ENERGY 8AM
```
[Demand_NOR_SH_8AM,Demand_NOR_DHW_8AM,R2_nor_SH_8AM,R2_nor_DHW_8AM]=ESC_CONT_CPT
(coeff_DEF_SH_8AM,coeff_DEF_DHW_8AM,T_SH_8AM,T_DHW_8AM,CPT_8AM,Demand_SH_8AM,Dem
and_DHW_8AM,9);
```

%ENERGY 9AM
```
[Demand_NOR_SH_9AM,Demand_NOR_DHW_9AM,R2_nor_SH_9AM,R2_nor_DHW_9AM]=ESC_CONT_CPT
(coeff_DEF_SH_9AM,coeff_DEF_DHW_9AM,T_SH_9AM,T_DHW_9AM,CPT_9AM,Demand_SH_9AM,Dem
and_DHW_9AM,10);
```

%ENERGY 10AM
```
[Demand_NOR_SH_10AM,Demand_NOR_DHW_10AM,R2_nor_SH_10AM,R2_nor_DHW_10AM]=ESC_CONT
_CPT(coeff_DEF_SH_10AM,coeff_DEF_DHW_10AM,T_SH_10AM,T_DHW_10AM,CPT_10AM,Demand_S
H_10AM,Demand_DHW_10AM,11);
```

%ENERGY 11AM
```
[Demand_NOR_SH_11AM,Demand_NOR_DHW_11AM,R2_nor_SH_11AM,R2_nor_DHW_11AM]=ESC_CONT
_CPT(coeff_DEF_SH_11AM,coeff_DEF_DHW_11AM,T_SH_11AM,T_DHW_11AM,CPT_11AM,Demand_S
H_11AM,Demand_DHW_11AM,12);
```

%ENERGY 12AM

```
[Demand_NOR_SH_12AM,Demand_NOR_DHW_12AM,R2_nor_SH_12AM,R2_nor_DHW_12AM]=ESC_CONT
_CPT(coeff_DEF_SH_12AM,coeff_DEF_DHW_12AM,T_SH_12AM,T_DHW_12AM,CPT_12AM,Demand_S
H_12AM,Demand_DHW_12AM,13);

%ENERGY 13PM
[Demand_NOR_SH_13PM,Demand_NOR_DHW_13PM,R2_nor_SH_13PM,R2_nor_DHW_13PM]=ESC_CONT
_CPT(coeff_DEF_SH_13PM,coeff_DEF_DHW_13PM,T_SH_13PM,T_DHW_13PM,CPT_13PM,Demand_S
H_13PM,Demand_DHW_13PM,14);

%ENERGY 14PM
[Demand_NOR_SH_14PM,Demand_NOR_DHW_14PM,R2_nor_SH_14PM,R2_nor_DHW_14PM]=ESC_CONT
_CPT(coeff_DEF_SH_14PM,coeff_DEF_DHW_14PM,T_SH_14PM,T_DHW_14PM,CPT_14PM,Demand_S
H_14PM,Demand_DHW_14PM,15);

%ENERGY 15PM
[Demand_NOR_SH_15PM,Demand_NOR_DHW_15PM,R2_nor_SH_15PM,R2_nor_DHW_15PM]=ESC_CONT
_CPT(coeff_DEF_SH_15PM,coeff_DEF_DHW_15PM,T_SH_15PM,T_DHW_15PM,CPT_15PM,Demand_S
H_15PM,Demand_DHW_15PM,16);

%ENERGY 16PM
[Demand_NOR_SH_16PM,Demand_NOR_DHW_16PM,R2_nor_SH_16PM,R2_nor_DHW_16PM]=ESC_CONT
_CPT(coeff_DEF_SH_16PM,coeff_DEF_DHW_16PM,T_SH_16PM,T_DHW_16PM,CPT_16PM,Demand_S
H_16PM,Demand_DHW_16PM,17);

%ENERGY 17PM
[Demand_NOR_SH_17PM,Demand_NOR_DHW_17PM,R2_nor_SH_17PM,R2_nor_DHW_17PM]=ESC_CONT
_CPT(coeff_DEF_SH_17PM,coeff_DEF_DHW_17PM,T_SH_17PM,T_DHW_17PM,CPT_17PM,Demand_S
H_17PM,Demand_DHW_17PM,18);

%ENERGY 18PM
[Demand_NOR_SH_18PM,Demand_NOR_DHW_18PM,R2_nor_SH_18PM,R2_nor_DHW_18PM]=ESC_CONT
_CPT(coeff_DEF_SH_18PM,coeff_DEF_DHW_18PM,T_SH_18PM,T_DHW_18PM,CPT_18PM,Demand_S
H_18PM,Demand_DHW_18PM,19);

%ENERGY 19PM
[Demand_NOR_SH_19PM,Demand_NOR_DHW_19PM,R2_nor_SH_19PM,R2_nor_DHW_19PM]=ESC_CONT
_CPT(coeff_DEF_SH_19PM,coeff_DEF_DHW_19PM,T_SH_19PM,T_DHW_19PM,CPT_19PM,Demand_S
H_19PM,Demand_DHW_19PM,20);

%ENERGY 20PM
[Demand_NOR_SH_20PM,Demand_NOR_DHW_20PM,R2_nor_SH_20PM,R2_nor_DHW_20PM]=ESC_CONT
_CPT(coeff_DEF_SH_20PM,coeff_DEF_DHW_20PM,T_SH_20PM,T_DHW_20PM,CPT_20PM,Demand_S
H_20PM,Demand_DHW_20PM,21);

%ENERGY 21PM
[Demand_NOR_SH_21PM,Demand_NOR_DHW_21PM,R2_nor_SH_21PM,R2_nor_DHW_21PM]=ESC_CONT
_CPT(coeff_DEF_SH_21PM,coeff_DEF_DHW_21PM,T_SH_21PM,T_DHW_21PM,CPT_21PM,Demand_S
H_21PM,Demand_DHW_21PM,22);

%ENERGY 22PM
[Demand_NOR_SH_22PM,Demand_NOR_DHW_22PM,R2_nor_SH_22PM,R2_nor_DHW_22PM]=ESC_CONT
_CPT(coeff_DEF_SH_22PM,coeff_DEF_DHW_22PM,T_SH_22PM,T_DHW_22PM,CPT_22PM,Demand_S
H_22PM,Demand_DHW_22PM,23);

%ENERGY 23PM
[Demand_NOR_SH_23PM,Demand_NOR_DHW_23PM,R2_nor_SH_23PM,R2_nor_DHW_23PM]=ESC_CONT
_CPT(coeff_DEF_SH_23PM,coeff_DEF_DHW_23PM,T_SH_23PM,T_DHW_23PM,CPT_23PM,Demand_S
H_23PM,Demand_DHW_23PM,24);

%% calcolo di R2 per la valutazione delle performance dei modelli
```

```matlab
%ENERGY 0AM
R2_0AM=ones(1,2);
R2_0AM(1,1)=R2_nor_SH_0AM;
R2_0AM(1,2)=R2_nor_DHW_0AM;

%ENERGY 1AM
R2_1AM=ones(1,2);
R2_1AM(1,1)=R2_nor_SH_1AM;
R2_1AM(1,2)=R2_nor_DHW_1AM;

%ENERGY 2AM
R2_2AM=ones(1,2);
R2_2AM(1,1)=R2_nor_SH_2AM;
R2_2AM(1,2)=R2_nor_DHW_2AM;

%ENERGY 3AM
R2_3AM=ones(1,2);
R2_3AM(1,1)=R2_nor_SH_3AM;
R2_3AM(1,2)=R2_nor_DHW_3AM;

%ENERGY 4AM
R2_4AM=ones(1,2);
R2_4AM(1,1)=R2_nor_SH_4AM;
R2_4AM(1,2)=R2_nor_DHW_4AM;

%ENERGY 5AM
R2_5AM=ones(1,2);
R2_5AM(1,1)=R2_nor_SH_5AM;
R2_5AM(1,2)=R2_nor_DHW_5AM;

%ENERGY 6AM
R2_6AM=ones(1,2);
R2_6AM(1,1)=R2_nor_SH_6AM;
R2_6AM(1,2)=R2_nor_DHW_6AM;

%ENERGY 7AM
R2_7AM=ones(1,2);
R2_7AM(1,1)=R2_nor_SH_7AM;
R2_7AM(1,2)=R2_nor_DHW_7AM;

%ENERGY 8AM
R2_8AM=ones(1,2);
R2_8AM(1,1)=R2_nor_SH_8AM;
R2_8AM(1,2)=R2_nor_DHW_8AM;

%ENERGY 9AM
R2_9AM=ones(1,2);
R2_9AM(1,1)=R2_nor_SH_9AM;
R2_9AM(1,2)=R2_nor_DHW_9AM;

%ENERGY 10AM
R2_10AM=ones(1,2);
R2_10AM(1,1)=R2_nor_SH_10AM;
R2_10AM(1,2)=R2_nor_DHW_10AM;

%ENERGY 11AM
R2_11AM=ones(1,2);
R2_11AM(1,1)=R2_nor_SH_11AM;
R2_11AM(1,2)=R2_nor_DHW_11AM;
```

```matlab
%ENERGY 12AM
R2_12AM=ones(1,2);
R2_12AM(1,1)=R2_nor_SH_12AM;
R2_12AM(1,2)=R2_nor_DHW_12AM;


%ENERGY 13PM
R2_13PM=ones(1,2);
R2_13PM(1,1)=R2_nor_SH_13PM;
R2_13PM(1,2)=R2_nor_DHW_13PM;


%ENERGY 14PM
R2_14PM=ones(1,2);
R2_14PM(1,1)=R2_nor_SH_14PM;
R2_14PM(1,2)=R2_nor_DHW_14PM;


%ENERGY 15PM
R2_15PM=ones(1,2);
R2_15PM(1,1)=R2_nor_SH_15PM;
R2_15PM(1,2)=R2_nor_DHW_15PM;


%ENERGY 16PM
R2_16PM=ones(1,2);
R2_16PM(1,1)=R2_nor_SH_16PM;
R2_16PM(1,2)=R2_nor_DHW_16PM;


%ENERGY 17PM
R2_17PM=ones(1,2);
R2_17PM(1,1)=R2_nor_SH_17PM;
R2_17PM(1,2)=R2_nor_DHW_17PM;


%ENERGY 18PM
R2_18PM=ones(1,2);
R2_18PM(1,1)=R2_nor_SH_18PM;
R2_18PM(1,2)=R2_nor_DHW_18PM;


%ENERGY 19PM
R2_19PM=ones(1,2);
R2_19PM(1,1)=R2_nor_SH_19PM;
R2_19PM(1,2)=R2_nor_DHW_19PM;


%ENERGY 20PM
R2_20PM=ones(1,2);
R2_20PM(1,1)=R2_nor_SH_20PM;
R2_20PM(1,2)=R2_nor_DHW_20PM;


%ENERGY 21PM
R2_21PM=ones(1,2);
R2_21PM(1,1)=R2_nor_SH_21PM;
R2_21PM(1,2)=R2_nor_DHW_21PM;


%ENERGY 22PM
R2_22PM=ones(1,2);
R2_22PM(1,1)=R2_nor_SH_22PM;
R2_22PM(1,2)=R2_nor_DHW_22PM;


%ENERGY 23PM
R2_23PM=ones(1,2);
R2_23PM(1,1)=R2_nor_SH_23PM;
R2_23PM(1,2)=R2_nor_DHW_23PM;
```

```matlab
%% boxplot analysis
%con la funzione boxplot è possibile creare un grafico a scatole. Linea
%centrale è la mediana, la linea superiore e inferiore sono
%rispetttivamente il 75 e il 25 percentile mentre gli outliers sono
%indicati con dei '+'.

%creazione della matrice contenenti tutti i weekeend dell'anno divisa per
%ore (ogni colonna è un'ora) e quella con tutti i weekedaysù

%non sappiamo per cosa sia utilizzato questo edificio, conviene dividere
%anche sabato e domenica.

%il primo gennaio 2019, quando sono stati raccolti i dati, era di martedì.
%5 dato per sabati
%6 dato per le domeniche
%tutto il resto per i restanti weekdays.

%inizializzazione delle matrici
SUNDAY=ones(52,24);
SATURDAY=ones(52,24);
WEEKDAYS=[];

%correzione per ENERGY 23 PM (manca l'ultimo valore, l'ho considerato
%uguale al gionro precedente. L'ultimo gionro è un martedì, quindi come
%carico sarà simile al lunedì in quanto entrambi giorni lavorativi e
%successivi.

ENERGY_23PM(365,1)=ENERGY_23PM(364,1);

%creazione matrice contenente tutte e 24 le ore (colonne) ognuna con il
%consumo orario di un giorno dell'anno (ongi riga corrisponde ad un giorno
%del 2019

ENERGY_GLO=[ENERGY_0AM,ENERGY_1AM,ENERGY_2AM,ENERGY_3AM,ENERGY_4AM,ENERGY_5AM,EN
ERGY_6AM,ENERGY_7AM,ENERGY_8AM,ENERGY_9AM,ENERGY_10AM,ENERGY_11AM,ENERGY_12AM,EN
ERGY_13PM,ENERGY_14PM,ENERGY_15PM,ENERGY_16PM,ENERGY_17PM,ENERGY_18PM,ENERGY_19P
M,ENERGY_20PM,ENERGY_21PM,ENERGY_22PM,ENERGY_23PM];


%SABATO
for k=1:24
    l=1;
    for i=5:7:365
        SATURDAY(l,k)=ENERGY_GLO(i,k);
        l=l+1;
    end
end

%DOMENICA
for k=1:24
    l=1;
    for i=6:7:365
        SUNDAY(l,k)=ENERGY_GLO(i,k);
        l=l+1;
    end

end

%WEEK-DAYS
```

```matlab
%CI SONO PRIMA TUTTI I MARTEDì, POI TUTTI I MERCOLEDì ECC.. TANTO PER LA
%CREAZIONE DEL BOXPLOT NON è IMPORTANTE L'ORDINE DEI GIONRI, BASTA AVERE I
%DATI CORRETTAMENTE POSIZIONATI NELLE RIGHE E CORRETTAMENTE DIVISI PER ORE
for k=1:24
    l=1;
    for i=1:7:365
        WEEKDAYS(l,k)=ENERGY_GLO(i,k);
        l=l+1;
    end

    for i=2:7:365
        WEEKDAYS(l,k)=ENERGY_GLO(i,k);
        l=l+1;
    end

    for i=3:7:365
        WEEKDAYS(l,k)=ENERGY_GLO(i,k);
        l=l+1;
    end

    for i=4:7:365
        WEEKDAYS(l,k)=ENERGY_GLO(i,k);
        l=l+1;
    end

    for i=7:7:365
        WEEKDAYS(l,k)=ENERGY_GLO(i,k);
        l=l+1;
    end

end




figure (25)
subplot(2,2,1)
boxplot(SATURDAY,'PlotStyle','traditional')
grid on
xlabel('HOURS', 'FontSize', fontSize);
ylabel('ENERGY USE', 'FontSize', fontSize);
title('SATURDAY', 'FontSize', fontSize);

subplot(2,2,2)
boxplot(SUNDAY,'PlotStyle','traditional')
grid on
xlabel('HOURS', 'FontSize', fontSize);
ylabel('ENERGY USE', 'FontSize', fontSize);
title('SUNDAY', 'FontSize', fontSize);

subplot(2,2,3)
boxplot(WEEKDAYS,'PlotStyle','traditional')
grid on
xlabel('HOURS', 'FontSize', fontSize);
ylabel('ENERGY USE', 'FontSize', fontSize);
title('WEEKDAYS', 'FontSize', fontSize);

subplot(2,2,4)
boxplot(ENERGY_GLO,'PlotStyle','traditional')
grid on
```

```matlab
xlabel('HOURS', 'FontSize', fontSize);
ylabel('ENERGY USE', 'FontSize', fontSize);
title('TOTAL', 'FontSize', fontSize);

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

%% uso di confidence interval (CI) per la rimozione di outliers.
%dal grafico a scatole  emerso che in questo edificio non ci sono
%particolari punti outliers. implementare comunque per altri edifici.

%ENERGY 0AM
[y_CI_up_0AM,y_CI_down_0AM]=CI(coeff_DEF_SH_0AM,T_SH_0AM,Demand_SH_0AM,1);

%ENERGY 1AM
[y_CI_up_1AM,y_CI_down_1AM]=CI(coeff_DEF_SH_1AM,T_SH_1AM,Demand_SH_1AM,2);

%ENERGY 2AM
[y_CI_up_2AM,y_CI_down_2AM]=CI(coeff_DEF_SH_2AM,T_SH_2AM,Demand_SH_2AM,3);

%ENERGY 3AM
[y_CI_up_3AM,y_CI_down_3AM]=CI(coeff_DEF_SH_3AM,T_SH_3AM,Demand_SH_3AM,4);

%ENERGY 4AM
[y_CI_up_4AM,y_CI_down_4AM]=CI(coeff_DEF_SH_4AM,T_SH_4AM,Demand_SH_4AM,5);

%ENERGY 5AM
[y_CI_up_5AM,y_CI_down_5AM]=CI(coeff_DEF_SH_5AM,T_SH_5AM,Demand_SH_5AM,6);

%ENERGY 6AM
[y_CI_up_6AM,y_CI_down_6AM]=CI(coeff_DEF_SH_6AM,T_SH_6AM,Demand_SH_6AM,7);

%ENERGY 7AM
[y_CI_up_7AM,y_CI_down_7AM]=CI(coeff_DEF_SH_7AM,T_SH_7AM,Demand_SH_7AM,8);

%ENERGY 8AM
[y_CI_up_8AM,y_CI_down_8AM]=CI(coeff_DEF_SH_8AM,T_SH_8AM,Demand_SH_8AM,9);

%ENERGY 9AM
[y_CI_up_9AM,y_CI_down_9AM]=CI(coeff_DEF_SH_9AM,T_SH_9AM,Demand_SH_9AM,10);

%ENERGY 10AM
[y_CI_up_10AM,y_CI_down_10AM]=CI(coeff_DEF_SH_10AM,T_SH_10AM,Demand_SH_10AM,11);

%ENERGY 11AM
[y_CI_up_11AM,y_CI_down_11AM]=CI(coeff_DEF_SH_11AM,T_SH_11AM,Demand_SH_11AM,12);

%ENERGY 12AM
[y_CI_up_12AM,y_CI_down_12AM]=CI(coeff_DEF_SH_12AM,T_SH_12AM,Demand_SH_12AM,13);

%ENERGY 13PM
[y_CI_up_13PM,y_CI_down_13PM]=CI(coeff_DEF_SH_13PM,T_SH_13PM,Demand_SH_13PM,14);

%ENERGY 14PM
[y_CI_up_14PM,y_CI_down_14PM]=CI(coeff_DEF_SH_14PM,T_SH_14PM,Demand_SH_14PM,15);

%ENERGY 15PM
[y_CI_up_15PM,y_CI_down_15PM]=CI(coeff_DEF_SH_15PM,T_SH_15PM,Demand_SH_15PM,16);
```

```matlab
%ENERGY 16PM
[y_CI_up_16PM,y_CI_down_16PM]=CI(coeff_DEF_SH_16PM,T_SH_16PM,Demand_SH_16PM,17);

%ENERGY 17PM
[y_CI_up_17PM,y_CI_down_17PM]=CI(coeff_DEF_SH_17PM,T_SH_17PM,Demand_SH_17PM,18);

%ENERGY 18PM
[y_CI_up_18PM,y_CI_down_18PM]=CI(coeff_DEF_SH_18PM,T_SH_18PM,Demand_SH_18PM,19);

%ENERGY 19PM
[y_CI_up_19PM,y_CI_down_19PM]=CI(coeff_DEF_SH_19PM,T_SH_19PM,Demand_SH_19PM,20);

%ENERGY 20PM
[y_CI_up_20PM,y_CI_down_20PM]=CI(coeff_DEF_SH_20PM,T_SH_20PM,Demand_SH_20PM,21);

%ENERGY 21PM
[y_CI_up_21PM,y_CI_down_21PM]=CI(coeff_DEF_SH_21PM,T_SH_21PM,Demand_SH_21PM,22);

%ENERGY 22PM
[y_CI_up_22PM,y_CI_down_22PM]=CI(coeff_DEF_SH_22PM,T_SH_22PM,Demand_SH_22PM,23);

%ENERGY 23PM
[y_CI_up_23PM,y_CI_down_23PM]=CI(coeff_DEF_SH_23PM,T_SH_23PM,Demand_SH_23PM,24);
%% calcolo di ESC totale, non solo quella divisa in ore
%nella ESC totale applicare anche il calcolo del CI

figure (27)
%grafico punti sperimentali
subplot(3,3,1)
plot(T,ENERGY,'.')
grid on
xlabel('T', 'FontSize', fontSize);
ylabel('ENERGY USE', 'FontSize', fontSize);
title('ENERGY USE vs T', 'FontSize', fontSize);

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

%RICREARE LE FUNCTION PERCHè SONO SCRITTE METTENDO COME LIMITE 365
%calcolo BETA per ESC
[BETA_year,TEMP_year]=ESC_BETA_year(ENERGY,T,27);

%vedere il valore di CPT dal grafico
CPT_year=14.1;        %variazione di circa il 3% dal valore precedente

%divisione punti
[T_SH_year,Demand_SH_year,T_DHW_year,Demand_DHW_year,coeff_DEF_SH_year,coeff_DEF
_DHW_year]=ESC_DIVISION_year(CPT_year,ENERGY,T,27);

%imposizione continuità del modello in CPT
[Demand_NOR_SH_year,Demand_NOR_DHW_year,R2_nor_SH_year,R2_nor_DHW_year]=ESC_CONT
_CPT_year(coeff_DEF_SH_year,coeff_DEF_DHW_year,T_SH_year,T_DHW_year,CPT_year,Dem
and_SH_year,Demand_DHW_year,27);

%calcolo di CI per individuazione outliers
%va benissimo, coincide con lo studio. CI serve per l'individuazione di
%eventuali ouliers nella coda del profilo di domanda, non nei punti vicino
%a CPT dove sicuramente ci saranno punti al dì fuori perchè rappresentano
```

157

```matlab
%il consumo di DHW quando SH è spento nei mesi invernali.
[y_CI_up_year,y_CI_down_year]=CI_year(coeff_DEF_SH_year,T_SH_year,Demand_SH_year
,27);
%% TMA dependency control, corrected
%TMA determination from 1 to 48 hours before. only with SH part of demand
BETA=[];
i=1;
for k=1:48
    TMA_hour=movmean(T_SH_year,[0 k]);
    pol_hour=polyfit(TMA_hour,Demand_SH_year,1);
    control_beta=pol_hour(1,1);
    BETA(i)=control_beta;
    i=i+1;
end

guida_1=1:length(BETA);

figure (26)
subplot(2,1,1)
plot(T,ENERGY,'.')
hold on
plot(T_SH_year,Demand_NOR_SH_year,'.')
grid on
xlabel('T OUTDOOR', 'FontSize', fontSize);
ylabel('ENERGY USE', 'FontSize', fontSize);
title('T OUTDOOR vs ENERGY USE', 'FontSize', fontSize);

subplot(2,1,2)
plot(guida_1,BETA)
grid on
xlabel('Lag of Temperature MA (hours)', 'FontSize', fontSize);
ylabel('Correlation between ENERGY and TMA', 'FontSize', fontSize);
title('TMA vs ENERGY USE', 'FontSize', fontSize);

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);
%% DHW mean profiles analysis
%non dipende dalla Tout ma solamente dall'utilizzo dell'edificio e dal
%numero di utenti al suo interno. Non è possibile usare la regressione
%lineare con la T per ottenere un modello accettabile. Più utile usare dei
%profili tipo

%FATTA PER MAGGIO-SETTEMBRE ma a maggio e settembre ci sono già influenze
%di SH, come si può vedere dall'aumento del consumo e dal cambio di profilo
%rispetto a giugno, luglio e agosto

%Maggio

%weekend
for k=1:24
    l=1;
    for i=124:7:151
        MAY_DHW_END(l,k)=ENERGY_GLO(i,k);
        l=l+1;
    end
     for i=125:7:151
        MAY_DHW_END(l,k)=ENERGY_GLO(i,k);
        l=l+1;
    end
end
```

```matlab
%work days
for k=1:24
    l=1;
    for i=121:7:151
        MAY_DHW_WORK(l,k)=ENERGY_GLO(i,k);
        l=l+1;
    end
     for i=122:7:151
        MAY_DHW_WORK(l,k)=ENERGY_GLO(i,k);
        l=l+1;
     end

      for i=123:7:151
        MAY_DHW_WORK(l,k)=ENERGY_GLO(i,k);
        l=l+1;
      end

       for i=126:7:151
        MAY_DHW_WORK(l,k)=ENERGY_GLO(i,k);
        l=l+1;
       end

       for i=127:7:151
        MAY_DHW_WORK(l,k)=ENERGY_GLO(i,k);
        l=l+1;
    end
end
MAY_MEAN_END=mean(MAY_DHW_END);
MAY_MEAN_WORK=mean(MAY_DHW_WORK);
time=[1:24];
figure (28)
subplot(5,2,1)
bar(time,MAY_MEAN_WORK)
grid on
axis([1 24 0 50]);
xlabel('TIME', 'FontSize', fontSize);
ylabel('AVG E USE', 'FontSize', fontSize);
title('WORK-DAYS', 'FontSize', fontSize);
legend('May')

subplot(5,2,2)
bar(time,MAY_MEAN_END)
grid on
axis([1 24 0 50]);
xlabel('TIME', 'FontSize', fontSize);
ylabel('AVG E USE', 'FontSize', fontSize);
title('WEEKEND', 'FontSize', fontSize);
legend('May')
% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

%giugno

%weekend
for k=1:24
    l=1;
    for i=152:7:181
        JUNE_DHW_END(l,k)=ENERGY_GLO(i,k);
        l=l+1;
    end
```

```matlab
    for i=153:7:181
        JUNE_DHW_END(l,k)=ENERGY_GLO(i,k);
        l=l+1;
    end
end

%work days
for k=1:24
    l=1;
    for i=154:7:181
        JUNE_DHW_WORK(l,k)=ENERGY_GLO(i,k);
        l=l+1;
    end
     for i=155:7:181
        JUNE_DHW_WORK(l,k)=ENERGY_GLO(i,k);
        l=l+1;
     end

     for i=156:7:181
        JUNE_DHW_WORK(l,k)=ENERGY_GLO(i,k);
        l=l+1;
     end

      for i=157:7:181
        JUNE_DHW_WORK(l,k)=ENERGY_GLO(i,k);
        l=l+1;
      end

      for i=158:7:181
        JUNE_DHW_WORK(l,k)=ENERGY_GLO(i,k);
        l=l+1;
    end
end
JUNE_MEAN_END=mean(JUNE_DHW_END);
JUNE_MEAN_WORK=mean(JUNE_DHW_WORK);
time=[1:24];
figure (28)
subplot(5,2,3)
bar(time,JUNE_MEAN_WORK)
grid on
axis([1 24 0 50]);
xlabel('TIME', 'FontSize', fontSize);
ylabel('AVG E USE', 'FontSize', fontSize);
title('WORK-DAYS', 'FontSize', fontSize);
legend('June')

subplot(5,2,4)
bar(time,JUNE_MEAN_END)
grid on
axis([1 24 0 50]);
xlabel('TIME', 'FontSize', fontSize);
ylabel('AVG E USE', 'FontSize', fontSize);
title('WEEKEND', 'FontSize', fontSize);
legend('June')

%LUGLIO
%WEEKEND
for k=1:24
    l=1;
    for i=187:7:212
        JULY_DHW_END(l,k)=ENERGY_GLO(i,k);
```

```matlab
            l=l+1;
        end
         for i=188:7:212
            JULY_DHW_END(l,k)=ENERGY_GLO(i,k);
            l=l+1;
        end
end

%WORK-DAYS
for k=1:24
    l=1;
    for i=182:7:212
        JULY_DHW_WORK(l,k)=ENERGY_GLO(i,k);
        l=l+1;
    end
     for i=183:7:212
        JULY_DHW_WORK(l,k)=ENERGY_GLO(i,k);
        l=l+1;
     end


      for i=184:7:212
        JULY_DHW_WORK(l,k)=ENERGY_GLO(i,k);
        l=l+1;
       end


       for i=185:7:212
        JULY_DHW_WORK(l,k)=ENERGY_GLO(i,k);
        l=l+1;
        end


    for i=186:7:212
        JULY_DHW_WORK(l,k)=ENERGY_GLO(i,k);
        l=l+1;
    end
end
JULY_MEAN_END=mean(JULY_DHW_END);
JULY_MEAN_WORK=mean(JULY_DHW_WORK);
time=[1:24];
figure (28)
subplot(5,2,5)
bar(time,JULY_MEAN_WORK)
grid on
axis([1 24 0 50]);
xlabel('TIME', 'FontSize', fontSize);
ylabel('AVG E USE', 'FontSize', fontSize);
title('WORK-DAYS', 'FontSize', fontSize);
legend('July')

subplot(5,2,6)
bar(time,JULY_MEAN_END)
grid on
axis([1 24 0 50]);
xlabel('TIME', 'FontSize', fontSize);
ylabel('AVG E USE', 'FontSize', fontSize);
title('WEEKEND', 'FontSize', fontSize);
legend('July')

%AGOSTO
%WEEKEND
for k=1:24
    l=1;
```

```matlab
    for i=215:7:243
        AUGUST_DHW_END(l,k)=ENERGY_GLO(i,k);
        l=l+1;
    end
     for i=216:7:243
        AUGUST_DHW_END(l,k)=ENERGY_GLO(i,k);
        l=l+1;
    end
end

%WORK-DAYS
for k=1:24
    l=1;
    for i=213:7:243
        AUGUST_DHW_WORK(l,k)=ENERGY_GLO(i,k);
        l=l+1;
    end
     for i=214:7:243
        AUGUST_DHW_WORK(l,k)=ENERGY_GLO(i,k);
        l=l+1;
     end


     for i=217:7:243
        AUGUST_DHW_WORK(l,k)=ENERGY_GLO(i,k);
        l=l+1;
     end


      for i=218:7:243
        AUGUST_DHW_WORK(l,k)=ENERGY_GLO(i,k);
        l=l+1;
       end


    for i=219:7:243
        AUGUST_DHW_WORK(l,k)=ENERGY_GLO(i,k);
        l=l+1;
    end
end
AUGUST_MEAN_END=mean(AUGUST_DHW_END);
AUGUST_MEAN_WORK=mean(AUGUST_DHW_WORK);
time=[1:24];
figure (28)
subplot(5,2,7)
bar(time,AUGUST_MEAN_WORK)
grid on
axis([1 24 0 50]);
xlabel('TIME', 'FontSize', fontSize);
ylabel('AVG E USE', 'FontSize', fontSize);
title('WORK-DAYS', 'FontSize', fontSize);
legend('August')

subplot(5,2,8)
bar(time,AUGUST_MEAN_END)
grid on
axis([1 24 0 50]);
xlabel('TIME', 'FontSize', fontSize);
ylabel('AVG E USE', 'FontSize', fontSize);
title('WEEKEND', 'FontSize', fontSize);
legend('August')

%SETTEMBRE
%WEEKEND
```

```matlab
for k=1:24
    l=1;
    for i=244:7:273
        SEPTEMBER_DHW_END(l,k)=ENERGY_GLO(i,k);
        l=l+1;
    end
     for i=250:7:273
        SEPTEMBER_DHW_END(l,k)=ENERGY_GLO(i,k);
        l=l+1;
    end
end

%WORK-DAYS
for k=1:24
    l=1;
    for i=245:7:273
        SEPTEMBER_DHW_WORK(l,k)=ENERGY_GLO(i,k);
        l=l+1;
    end
     for i=246:7:273
        SEPTEMBER_DHW_WORK(l,k)=ENERGY_GLO(i,k);
        l=l+1;
     end


      for i=247:7:273
        SEPTEMBER_DHW_WORK(l,k)=ENERGY_GLO(i,k);
        l=l+1;
      end


       for i=248:7:273
        SEPTEMBER_DHW_WORK(l,k)=ENERGY_GLO(i,k);
        l=l+1;
       end


    for i=249:7:273
        SEPTEMBER_DHW_WORK(l,k)=ENERGY_GLO(i,k);
        l=l+1;
    end
end
SEPTEMBER_MEAN_END=mean(SEPTEMBER_DHW_END);
SEPTEMBER_MEAN_WORK=mean(SEPTEMBER_DHW_WORK);
time=[1:24];
figure (28)
subplot(5,2,9)
bar(time,SEPTEMBER_MEAN_WORK)
grid on
axis([1 24 0 50]);
xlabel('TIME', 'FontSize', fontSize);
ylabel('AVG E USE', 'FontSize', fontSize);
title('WORK-DAYS', 'FontSize', fontSize);
legend('September')

subplot(5,2,10)
bar(time,SEPTEMBER_MEAN_END)
grid on
axis([1 24 0 50]);
xlabel('TIME', 'FontSize', fontSize);
ylabel('AVG E USE', 'FontSize', fontSize);
title('WEEKEND', 'FontSize', fontSize);
legend('September')
```

```matlab
%% Calcolo SH e DHW
%T dependent part dell'ESC è traslata verso l'alto anche del consumo di
%DHW. L'obiettivo ora è di calcolare questo shifting coefficient per
%ottenere solo i valori di SH e DHW separati

%la parte di DHW coincide con il minimo valore di ESC
Dem_mod_SHDHW_0AM=polyval(coeff_DEF_SH_0AM,T_SH_0AM);
Dem_mod_DHW_0AM=polyval(coeff_DEF_DHW_0AM,T_DHW_0AM);
E_SH_model_0AM=Dem_mod_SHDHW_0AM-min(Dem_mod_DHW_0AM);

%ENERGY 0AM
[val_1_0AM,val_2_0AM,E_loss_0AM,E_SH_balanced_0AM,E_DHW_balanced_0AM]=SHDHW_spli
t_function(Demand_NOR_SH_0AM,Demand_NOR_DHW_0AM,T_SH_0AM,T_DHW_0AM,DMT,ENERGY_GL
O,CPT_0AM,ENERGY_0AM,1);

%ENERGY 1AM
[val_1_1AM,val_2_1AM,E_loss_1AM,E_SH_balanced_1AM,E_DHW_balanced_1AM]=SHDHW_spli
t_function(Demand_NOR_SH_1AM,Demand_NOR_DHW_1AM,T_SH_1AM,T_DHW_1AM,DMT,ENERGY_GL
O,CPT_1AM,ENERGY_1AM,2);

%ENERGY 2AM
[val_1_2AM,val_2_2AM,E_loss_2AM,E_SH_balanced_2AM,E_DHW_balanced_2AM]=SHDHW_spli
t_function(Demand_NOR_SH_2AM,Demand_NOR_DHW_2AM,T_SH_2AM,T_DHW_2AM,DMT,ENERGY_GL
O,CPT_2AM,ENERGY_2AM,3);

%ENERGY 3AM
[val_1_3AM,val_2_3AM,E_loss_3AM,E_SH_balanced_3AM,E_DHW_balanced_3AM]=SHDHW_spli
t_function(Demand_NOR_SH_3AM,Demand_NOR_DHW_3AM,T_SH_3AM,T_DHW_3AM,DMT,ENERGY_GL
O,CPT_3AM,ENERGY_3AM,4);

%ENERGY 4AM
[val_1_4AM,val_2_4AM,E_loss_4AM,E_SH_balanced_4AM,E_DHW_balanced_4AM]=SHDHW_spli
t_function(Demand_NOR_SH_4AM,Demand_NOR_DHW_4AM,T_SH_4AM,T_DHW_4AM,DMT,ENERGY_GL
O,CPT_4AM,ENERGY_4AM,5);

%ENERGY 5AM
[val_1_5AM,val_2_5AM,E_loss_5AM,E_SH_balanced_5AM,E_DHW_balanced_5AM]=SHDHW_spli
t_function(Demand_NOR_SH_5AM,Demand_NOR_DHW_5AM,T_SH_5AM,T_DHW_5AM,DMT,ENERGY_GL
O,CPT_5AM,ENERGY_5AM,6);

%ENERGY 6AM
[val_1_6AM,val_2_6AM,E_loss_6AM,E_SH_balanced_6AM,E_DHW_balanced_6AM]=SHDHW_spli
t_function(Demand_NOR_SH_6AM,Demand_NOR_DHW_6AM,T_SH_6AM,T_DHW_6AM,DMT,ENERGY_GL
O,CPT_6AM,ENERGY_6AM,7);

%ENERGY 7AM
[val_1_7AM,val_2_7AM,E_loss_7AM,E_SH_balanced_7AM,E_DHW_balanced_7AM]=SHDHW_spli
t_function(Demand_NOR_SH_7AM,Demand_NOR_DHW_7AM,T_SH_7AM,T_DHW_7AM,DMT,ENERGY_GL
O,CPT_7AM,ENERGY_7AM,8);

%ENERGY 8AM
[val_1_8AM,val_2_8AM,E_loss_8AM,E_SH_balanced_8AM,E_DHW_balanced_8AM]=SHDHW_spli
t_function(Demand_NOR_SH_8AM,Demand_NOR_DHW_8AM,T_SH_8AM,T_DHW_8AM,DMT,ENERGY_GL
O,CPT_8AM,ENERGY_8AM,9);

%ENERGY 9AM
[val_1_9AM,val_2_9AM,E_loss_9AM,E_SH_balanced_9AM,E_DHW_balanced_9AM]=SHDHW_spli
t_function(Demand_NOR_SH_9AM,Demand_NOR_DHW_9AM,T_SH_9AM,T_DHW_9AM,DMT,ENERGY_GL
O,CPT_9AM,ENERGY_9AM,10);
```

```matlab
%ENERGY 10AM
[val_1_10AM,val_2_10AM,E_loss_10AM,E_SH_balanced_10AM,E_DHW_balanced_10AM]=SHDHW
_split_function(Demand_NOR_SH_10AM,Demand_NOR_DHW_10AM,T_SH_10AM,T_DHW_10AM,DMT,
ENERGY_GLO,CPT_10AM,ENERGY_10AM,11);

%ENERGY 11AM
[val_1_11AM,val_2_11AM,E_loss_11AM,E_SH_balanced_11AM,E_DHW_balanced_11AM]=SHDHW
_split_function(Demand_NOR_SH_11AM,Demand_NOR_DHW_11AM,T_SH_11AM,T_DHW_11AM,DMT,
ENERGY_GLO,CPT_11AM,ENERGY_11AM,12);

%ENERGY 12AM
[val_1_12AM,val_2_12AM,E_loss_12AM,E_SH_balanced_12AM,E_DHW_balanced_12AM]=SHDHW
_split_function(Demand_NOR_SH_12AM,Demand_NOR_DHW_12AM,T_SH_12AM,T_DHW_12AM,DMT,
ENERGY_GLO,CPT_12AM,ENERGY_12AM,13);

%ENERGY 13PM
[val_1_13PM,val_2_13PM,E_loss_13PM,E_SH_balanced_13PM,E_DHW_balanced_13PM]=SHDHW
_split_function(Demand_NOR_SH_13PM,Demand_NOR_DHW_13PM,T_SH_13PM,T_DHW_13PM,DMT,
ENERGY_GLO,CPT_13PM,ENERGY_13PM,14);

%ENERGY 14PM
[val_1_14PM,val_2_14PM,E_loss_14PM,E_SH_balanced_14PM,E_DHW_balanced_14PM]=SHDHW
_split_function(Demand_NOR_SH_14PM,Demand_NOR_DHW_14PM,T_SH_14PM,T_DHW_14PM,DMT,
ENERGY_GLO,CPT_14PM,ENERGY_14PM,15);

%ENERGY 15PM
[val_1_15PM,val_2_15PM,E_loss_15PM,E_SH_balanced_15PM,E_DHW_balanced_15PM]=SHDHW
_split_function(Demand_NOR_SH_15PM,Demand_NOR_DHW_15PM,T_SH_15PM,T_DHW_15PM,DMT,
ENERGY_GLO,CPT_15PM,ENERGY_15PM,16);

%ENERGY 16PM
[val_1_16PM,val_2_16PM,E_loss_16PM,E_SH_balanced_16PM,E_DHW_balanced_16PM]=SHDHW
_split_function(Demand_NOR_SH_16PM,Demand_NOR_DHW_16PM,T_SH_16PM,T_DHW_16PM,DMT,
ENERGY_GLO,CPT_16PM,ENERGY_16PM,17);

%ENERGY 17PM
[val_1_17PM,val_2_17PM,E_loss_17PM,E_SH_balanced_17PM,E_DHW_balanced_17PM]=SHDHW
_split_function(Demand_NOR_SH_17PM,Demand_NOR_DHW_17PM,T_SH_17PM,T_DHW_17PM,DMT,
ENERGY_GLO,CPT_17PM,ENERGY_17PM,18);

%ENERGY 18PM
[val_1_18PM,val_2_18PM,E_loss_18PM,E_SH_balanced_18PM,E_DHW_balanced_18PM]=SHDHW
_split_function(Demand_NOR_SH_18PM,Demand_NOR_DHW_18PM,T_SH_18PM,T_DHW_18PM,DMT,
ENERGY_GLO,CPT_18PM,ENERGY_18PM,19);

%ENERGY 19PM
[val_1_19PM,val_2_19PM,E_loss_19PM,E_SH_balanced_19PM,E_DHW_balanced_19PM]=SHDHW
_split_function(Demand_NOR_SH_19PM,Demand_NOR_DHW_19PM,T_SH_19PM,T_DHW_19PM,DMT,
ENERGY_GLO,CPT_19PM,ENERGY_19PM,20);

%ENERGY 20PM
[val_1_20PM,val_2_20PM,E_loss_20PM,E_SH_balanced_20PM,E_DHW_balanced_20PM]=SHDHW
_split_function(Demand_NOR_SH_20PM,Demand_NOR_DHW_20PM,T_SH_20PM,T_DHW_20PM,DMT,
ENERGY_GLO,CPT_20PM,ENERGY_20PM,21);

%ENERGY 21PM
[val_1_21PM,val_2_21PM,E_loss_21PM,E_SH_balanced_21PM,E_DHW_balanced_21PM]=SHDHW
_split_function(Demand_NOR_SH_21PM,Demand_NOR_DHW_21PM,T_SH_21PM,T_DHW_21PM,DMT,
ENERGY_GLO,CPT_21PM,ENERGY_21PM,22);
```

```
%ENERGY 22PM
[val_1_22PM,val_2_22PM,E_loss_22PM,E_SH_balanced_22PM,E_DHW_balanced_22PM]=SHDHW
_split_function(Demand_NOR_SH_22PM,Demand_NOR_DHW_22PM,T_SH_22PM,T_DHW_22PM,DMT,
ENERGY_GLO,CPT_22PM,ENERGY_22PM,23);

%ENERGY 23PM
[val_1_23PM,val_2_23PM,E_loss_23PM,E_SH_balanced_23PM,E_DHW_balanced_23PM]=SHDHW
_split_function(Demand_NOR_SH_23PM,Demand_NOR_DHW_23PM,T_SH_23PM,T_DHW_23PM,DMT,
ENERGY_GLO,CPT_23PM,ENERGY_23PM,24);
%% global splitting
%ENERGY GLOBAL
[val_1_year,val_2_year,E_loss_year,E_SH_balanced_year_corr,E_DHW_year_corr]=SHDH
W_split_function_year(Demand_NOR_SH_year,Demand_NOR_DHW_year,T_SH_year,T_DHW_yea
r,T,ENERGY_GLO,CPT_year,ENERGY,27);

%% control dependence on indoor temperature and influence on Qtot
T_in_day=22;     %from extimation of residential buildings, cold climate
T_in_night=18;

T_in_day_vector=ones(1,length(T_SH_17PM));
T_in_day_vector(:)=T_in_day;

T_in_night_vector=ones(1,length(T_SH_0AM));
T_in_night_vector(:)=T_in_night;

DELTA_T_0AM=T_in_night-T_SH_0AM;
DELTA_T_17PM=T_in_day-T_SH_17PM;

coeff_delta_0AM=polyfit(DELTA_T_0AM,Demand_SH_0AM,1);
coeff_delta_17PM=polyfit(DELTA_T_17PM,Demand_SH_17PM,1);
%no differences between considering delta T or just Tout
%% Total energy for DHW
DHW_year_energy=sum(E_DHW_year_corr);
DHW_year_energy_smooth=sum(E_DHW_Y_NEW_line);
DHW_year_standard=(40*4.186*(46-3)/3600)*365;
people_inside=DHW_year_energy/DHW_year_standard;
people_inside_smooth=DHW_year_energy_smooth/DHW_year_standard;
```

"ESC_BETA.m"

```
%function per costruire e calcolare ESC usando i vetttori di partenza
function [BETA,TEMP]=ESC_BETA(ENERGY_0AM,DMT,h)
 Temp=22;    %messo 22, si perdono i punti più a desra, ma tanto sono sicuro che
il CPT non è lì. in questo modo evito warning (ho sempre più di 2 punti per fare
il polyfit)
 f=1;
 vectorE=[];
 L=1;
 for mod=0.1:0.1:40
     T_int=Temp-mod;
     for j=1:365
         valueE=ENERGY_0AM(j,1);
         valueT=DMT(j,1);
         if valueT>T_int
             vectorE(f,1)=valueE;
             vectorT(f,1)=valueT;
             f=f+1;
         end
     end
     poli=polyfit(vectorT,vectorE,1);
     BETA(L,1)=poli(1,1);
```

```
    TEMP(L,1)=T_int;
    L=L+1;
 end


figure(h)
markerSize = 20;
fontSize = 18;
subplot (3,3,2);
 plot(TEMP,BETA,'.','MarkerSize', markerSize)
grid on;
ylabel('BETA', 'FontSize', fontSize);
title('BETA VS T VARIATION', 'FontSize', fontSize)
end
```

"ESC_BETA_year.m"

```
%function ESC BETA per l'anno completo
function [BETA,TEMP]=ESC_BETA_year(ENERGY,T,h)
 Temp=25;
 f=1;
 vectorE=[];
 L=1;
 for mod=0.1:0.1:40
     T_int=Temp-mod;
     for j=1:8759
         valueE=ENERGY(j,1);
         valueT=T(j,1);
         if valueT>T_int
             vectorE(f,1)=valueE;
             vectorT(f,1)=valueT;
             f=f+1;
         end
     end
    poli=polyfit(vectorT,vectorE,1);
    BETA(L,1)=poli(1,1);
    TEMP(L,1)=T_int;
    L=L+1;
 end


figure(h)
markerSize = 20;
fontSize = 18;
subplot (3,3,2);
plot(TEMP,BETA,'.','MarkerSize', markerSize)
grid on;
ylabel('BETA', 'FontSize', fontSize);
title('BETA VS T VARIATION', 'FontSize', fontSize)
end
```

"ESC_CONT_CPT"

```
%function per la creazione di ESC con continuità in CPT, con coefficienti
%uguali a quelli trovati dal fit nei due tratti separati. confrontandola
%con altri metodo di fitting è la soluzione migliore per la creazione della
%curva voluta.

function
[Demand_NOR_SH_0AM,Demand_NOR_DHW_0AM,R2_nor_SH,R2_nor_DHW]=ESC_CONT_CPT(coeff_D
EF_SH_0AM,coeff_DEF_DHW_0AM,T_SH_0AM,T_DHW_0AM,CPT_0AM,Demand_SH_0AM,Demand_DHW_
0AM,i)
```

```
Demand_NOR_SH_0AM=polyval(coeff_DEF_SH_0AM,T_SH_0AM);
Demand_NOR_DHW_0AM=coeff_DEF_SH_0AM(2)+CPT_0AM*(coeff_DEF_SH_0AM(1)-
coeff_DEF_DHW_0AM(1))+coeff_DEF_DHW_0AM(1).*T_DHW_0AM;
value_DHW=polyval(coeff_DEF_DHW_0AM,T_DHW_0AM);
%calcolo di R2

 SSE_SH=sum((Demand_SH_0AM-Demand_NOR_SH_0AM).^2);
 SSYY_SH=sum((Demand_SH_0AM-mean(Demand_SH_0AM)).^2);
 SSE_DHW=sum((Demand_DHW_0AM-value_DHW).^2);
 SSYY_DHW=sum((Demand_DHW_0AM-mean(Demand_DHW_0AM)).^2);
R2_nor_SH=1-(SSE_SH/SSYY_SH);
R2_nor_DHW=1-(SSE_DHW/SSYY_DHW);

figure(i)
subplot(3,3,4)
% pos1=[0.1 0.1 0.5 0.5];
% subplot('Position',pos1)
fontSize = 18;

plot(T_SH_0AM,Demand_NOR_SH_0AM,'r')
hold on
grid on
%plot(DMT,ENERGY_0AM,'.')
plot(T_SH_0AM,Demand_SH_0AM,'.r')
plot(T_DHW_0AM,Demand_DHW_0AM,'.g')
plot(T_DHW_0AM,Demand_NOR_DHW_0AM,'g')
hold off
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 0AM', 'FontSize', fontSize);
title('ENERGY vs DMT, math model, Continuity CPT', 'FontSize', fontSize);
end
```

"ESC_CONT_CPT_year.m"

```
%continuità in CPT per l'anno completo

function
[Demand_NOR_SH_0AM,Demand_NOR_DHW_0AM,R2_nor_SH,R2_nor_DHW]=ESC_CONT_CPT_year(co
eff_DEF_SH_0AM,coeff_DEF_DHW_0AM,T_SH_0AM,T_DHW_0AM,CPT_0AM,Demand_SH_0AM,Demand
_DHW_0AM,i)

Demand_NOR_SH_0AM=polyval(coeff_DEF_SH_0AM,T_SH_0AM);
Demand_NOR_DHW_0AM=coeff_DEF_SH_0AM(2)+CPT_0AM*(coeff_DEF_SH_0AM(1)-
coeff_DEF_DHW_0AM(1))+coeff_DEF_DHW_0AM(1).*T_DHW_0AM;
value_DHW=polyval(coeff_DEF_DHW_0AM,T_DHW_0AM);
%calcolo di R2

 SSE_SH=sum((Demand_SH_0AM-Demand_NOR_SH_0AM).^2);
 SSYY_SH=sum((Demand_SH_0AM-mean(Demand_SH_0AM)).^2);
 SSE_DHW=sum((Demand_DHW_0AM-value_DHW).^2);
 SSYY_DHW=sum((Demand_DHW_0AM-mean(Demand_DHW_0AM)).^2);
R2_nor_SH=1-(SSE_SH/SSYY_SH);
R2_nor_DHW=1-(SSE_DHW/SSYY_DHW);

figure(i)
% pos1=[0.1 0.1 0.5 0.5];
% subplot('Position',pos1)
subplot(3,3,4)
```

```matlab
fontSize = 18;

plot(T_SH_0AM,Demand_NOR_SH_0AM,'b')
hold on
%plot(DMT,ENERGY_0AM,'.')
plot(T_SH_0AM,Demand_SH_0AM,'.r')
plot(T_DHW_0AM,Demand_DHW_0AM,'.g')
plot(T_DHW_0AM,Demand_NOR_DHW_0AM,'b')
grid on
hold off
xlabel('DMT', 'FontSize', fontSize);
ylabel('DEMAND', 'FontSize', fontSize);
title('ENERGY vs DMT, math model, Continuity CPT', 'FontSize', fontSize);
end
```

"ESC_DIVISION.m"

```matlab
%function per divisione punti

function
[T_SH,Demand_SH,T_DHW,Demand_DHW,coeff_DEF_SH,coeff_DEF_DHW]=ESC_DIVISION(CPT,EN
ERGY_0AM,DMT,k)

scor=1;
 scor2=1;
 T_SH=[];
 Demand_SH=[];
 T_DHW=[];
 Demand_DHW=[];

 for j=1:365
     Tvar=DMT(j,1);
     Dvar=ENERGY_0AM(j,1);

     if Tvar<CPT
         T_SH(1,scor)=Tvar;
         Demand_SH(1,scor)=Dvar;
         scor=scor+1;

     else T_DHW(1,scor2)=Tvar;
         Demand_DHW(1,scor2)=Dvar;
         scor2=scor2+1;
     end
 end
fontSize = 18;
markerSize = 20;

figure(k)
 subplot(3,3,3)
  plot(T_SH,Demand_SH,'r.')
  hold on
  plot(T_DHW,Demand_DHW,'g.')

  grid on


  coeff_DEF_SH=polyfit(T_SH,Demand_SH,1);
  coeff_DEF_DHW=polyfit(T_DHW,Demand_DHW,1);
```

```matlab
    value_SH=polyval(coeff_DEF_SH,T_SH);
    value_DHW=polyval(coeff_DEF_DHW,T_DHW);

    plot(T_SH,value_SH,'r')
    plot(T_DHW,value_DHW,'g')
    xlabel('DMT', 'FontSize', fontSize);
ylabel('DEMAND', 'FontSize', fontSize);
title('DEMAND vs DMT', 'FontSize', fontSize);
    hold off

end
```

"ESC_DIVISION_year.m"

```matlab
%function per dividere SH e DHW con l'intero anno

function
[T_SH,Demand_SH,T_DHW,Demand_DHW,coeff_DEF_SH,coeff_DEF_DHW]=ESC_DIVISION_year(C
PT,ENERGY,T,k)

scor=1;
 scor2=1;
 T_SH=[];
 Demand_SH=[];
 T_DHW=[];
 Demand_DHW=[];

 for j=1:8759
     Tvar=T(j,1);
     Dvar=ENERGY(j,1);

     if Tvar<CPT
         T_SH(1,scor)=Tvar;
         Demand_SH(1,scor)=Dvar;
         scor=scor+1;

     else T_DHW(1,scor2)=Tvar;
         Demand_DHW(1,scor2)=Dvar;
         scor2=scor2+1;
     end
 end
fontSize = 18;

figure(k)
 subplot(3,3,3)
  plot(T_SH,Demand_SH,'r.')
  hold on
  plot(T_DHW,Demand_DHW,'g.')

  grid on


  coeff_DEF_SH=polyfit(T_SH,Demand_SH,1);
  coeff_DEF_DHW=polyfit(T_DHW,Demand_DHW,1);

  value_SH=polyval(coeff_DEF_SH,T_SH);
  value_DHW=polyval(coeff_DEF_DHW,T_DHW);
```

```matlab
  plot(T_SH,value_SH,'r')
  plot(T_DHW,value_DHW,'g')
  xlabel('DMT', 'FontSize', fontSize);
ylabel('DEMAND', 'FontSize', fontSize);
title('DEMAND vs DMT', 'FontSize', fontSize);
  hold off


end
```

"HOUR_split.m"

```matlab
%function per trovare il vettore con 365 valori (1 anno) della
%corrispondente ora del giorno

function [ENERGY_0AM]=HOUR_split(ENERGY,var_0AM)
ENERGY_0AM=[];
%vettore energia ore 0 am
scor=1;
for j=1:365
    ENERGY_0AM(scor,1)=ENERGY(var_0AM);

    scor=scor+1;
    var_0AM=var_0AM+24;
end
end
```

"iterative_merging_process.m"

```matlab
%function for mergiging elements inside the group
function
[DHW_split_it11,it_DHW_split_last_new]=iterative_merging_process(DHW_split_g2)
%% iterative merging process

%cicle for automatic iterative procedure

%the first division is necessary
[DHW_split_it2]=iterative_merging_v3(DHW_split_g2);

%start comparing the difference between 2 succesive iteration
it_DHW_split_last=DHW_split_it2;
it_DHW_split_pre=DHW_split_g2;

it_number=1; %number of iterations

for sc=1:10  %maximum number of iteration possibile. the cicle ends when we
reach 10 iterations or when we reach 2% of tolerance

    %matrix dimensions must agree
    it_size_pre=size(it_DHW_split_pre);
    it_size_last=size(it_DHW_split_last);

    it_col_pre=it_size_pre(1,2);
    it_col_last=it_size_last(1,2);

    if it_col_pre==it_col_last
    it_difference=abs(it_DHW_split_last-it_DHW_split_pre);
    [it_maximum1,it_index1]=max(it_difference);
    [it_maximum2,it_index2]=max(it_maximum1);
```

```matlab
        else
            if it_col_pre>it_col_last
                it_difference_num=it_col_pre-it_col_last;
                it_raw=it_size_pre(1,1);
                it_corr=zeros(it_raw,it_difference_num);
                it_DHW_split_last=[it_DHW_split_last,it_corr];

                it_difference=abs(it_DHW_split_last-it_DHW_split_pre);
                [it_maximum1,it_index1]=max(it_difference);
                [it_maximum2,it_index2]=max(it_maximum1);
            end
        end

    it_raw_max=it_index1(1,it_index2);
    it_max_index=[it_raw_max,it_index2];   %indexes of the maximum elements of
the matrix "it_difference"

    %do iteration if this element is higher thana certain percentual of the
    %elements in the matrix (for example 1%)

    %selection of maximum value between two matrix at the index:
    it_index_pre=it_DHW_split_pre(it_raw_max,it_index2);
    it_index_last=it_DHW_split_last(it_raw_max,it_index2);
    it_vector_combined=[it_index_pre,it_index_last];
    it_maximum_value=max(it_vector_combined);

    it_percentage=(it_maximum2*100)/(it_maximum_value);

    %insert below the tolerance level (I put 1%)

    if it_percentage>=1   %tolerance level
        [it_DHW_split_last2]=iterative_merging_v3(it_DHW_split_last);
        it_number=it_number+1;
        it_DHW_split_pre=it_DHW_split_last;
        it_DHW_split_last=it_DHW_split_last2; %deve essere input della prossima
iterazione
    end


end
%per come è fatta l'iterazione devo tenerla così. alla fine, però, bisgona
%rimuovere tutte le colonne finali che hanno solo 0 come elementi

sizee_split_last2=size(it_DHW_split_last2);
col_last2=sizee_split_last2(1,2);
raw_last2=sizee_split_last2(1,1);
isu=1;
it_DHW_split_last_new=[];

for k=1:col_last2
    it_vettore=it_DHW_split_last2(:,k);
    it_control=zeros(raw_last2,1);

    if it_DHW_split_last2(1,k)~=0
       it_DHW_split_last_new(:,isu)=it_vettore;
       isu=isu+1;
    end

end
```

```
%iterative procedure in brutal force way. Now the cicle works, this is not
%useful
[DHW_split_it3]=iterative_merging_v3(DHW_split_it2);
[DHW_split_it4]=iterative_merging_v3(DHW_split_it3);
[DHW_split_it5]=iterative_merging_v3(DHW_split_it4);
[DHW_split_it6]=iterative_merging_v3(DHW_split_it5);
[DHW_split_it7]=iterative_merging_v3(DHW_split_it6);
[DHW_split_it8]=iterative_merging_v3(DHW_split_it7);
[DHW_split_it9]=iterative_merging_v3(DHW_split_it8);
[DHW_split_it10]=iterative_merging_v3(DHW_split_it9);
[DHW_split_it11]=iterative_merging_v3(DHW_split_it10);


end
```

"iterative_merging_v3.m"

```
%function for iterative merging v 3

%funziona ma diverge: per la sua stessa costruzione va a considerare alcuni
%gruppi più volte, quindi diverge. cambiare completamente impostazione e
%passare ad una dove il confronto tra gruppi è fatto nella modalità 1-2 poi
%3-4 e così via.. (non più 1-2, 2-3...)
%oppure mantenere la stessa impostazione, ma dopo il primo salvare solo il
%secondo vettore nel caso in cui nello step prima si sia fatto il merging:
%se 2-3 sono stati uniti nel gruppo 23, il passaggio dopo risulta che 3-4
%vanno tenuti separati, salvare solo il vettore 4. in questo modo non
%divergerà.

function [DHW_split_it2]=iterative_merging_v3(DHW_split)
%% itertive mergening con script per provare se va
number=2;
sizee_split=size(DHW_split);
ending=sizee_split(1,2);

%definition Tcr
Tcr_1_005=6.314;  %if n1+n2=3
Tcr_2_005=2.92;   %if n1+n2=4
Tcr_3_005=2.353;  %if n1+n2=5
Tcr_4_005=2.132;  %if n1+n2=6
Tcr_5_005=2.015;  %if n1+n2=7
Tcr_6_005=1.943;  %if n1+n2=8
Tcr_7_005=1.895;  %if n1+n2=9
Tcr_8_005=1.860;  %if n1+n2=10
Tcr_9_005=1.833;  %if n1+n2=11
Tcr_10_005=1.812; %if n1+n2=12
Tcr_11_005=1.796; %if n1+n2=13
Tcr_12_005=1.782; %if n1+n2=14
Tcr_13_005=1.771; %if n1+n2=15
Tcr_14_005=1.761; %if n1+n2=16
Tcr_15_005=1.753; %if n1+n2=17
Tcr_16_005=1.746; %if n1+n2=18
Tcr_17_005=1.740; %if n1+n2=19
Tcr_18_005=1.734; %if n1+n2=20
Tcr_19_005=1.729; %if n1+n2=21
Tcr_20_005=1.725; %if n1+n2=22
Tcr_21_005=1.721; %if n1+n2=23
```

```matlab
DHW_split_it2=[]; %matrix of DHW_split. It contains group merged from the first
iteration
%DHW_split_it2=zeros(1,1);

%cicle for controlling all colomns of DHW_split
for it=number:ending
%vector 1
sizee_split=size(DHW_split);
LUN1=sizee_split(1,1);
col_A=(it-1);    %number of the colomn i am considering
vector1=[];
for i=1:LUN1
    if DHW_split(i,col_A)~=0
        vector1(i,1)=DHW_split(i,col_A);
    else
    end
end

%vector 2
LUN2=sizee_split(1,1);
col_B=it;    %number of the colomn i am considering
vector2=[];
for i=1:LUN2
    if DHW_split(i,col_B)~=0
        vector2(i,1)=DHW_split(i,col_B);
    else
    end
end

% Student's t-test

%E: mean values
vector1_E=mean(vector1);
vector2_E=mean(vector2);

%n: number of elements
vector1_n=length(vector1);
vector2_n=length(vector2);

%S: standard deviation
vector1_S=std(vector1);
vector2_S=std(vector2);

%Tcal calculation
Tcal_iteration1=abs((vector1_E-
vector2_E)/(sqrt(((vector1_S^2)/vector1_n)+((vector2_S^2)/vector2_n))));

%choose the rigth value for Tcr
sum_n=vector1_n+vector2_n;
if sum_n==3
    Tcr=Tcr_1_005;
end
if sum_n==4
    Tcr=Tcr_2_005;
end
if sum_n==5
    Tcr=Tcr_3_005;
end
if sum_n==6
```

```matlab
        Tcr=Tcr_4_005;
end
if sum_n==7
    Tcr=Tcr_5_005;
end
if sum_n==8
    Tcr=Tcr_6_005;
end
if sum_n==9
    Tcr=Tcr_7_005;
end
if sum_n==10
    Tcr=Tcr_8_005;
end
if sum_n==11
    Tcr=Tcr_9_005;
end
if sum_n==12
    Tcr=Tcr_10_005;
end
if sum_n==13
    Tcr=Tcr_11_005;
end
if sum_n==14
    Tcr=Tcr_12_005;
end
if sum_n==15
    Tcr=Tcr_13_005;
end
if sum_n==16
    Tcr=Tcr_14_005;
end
if sum_n==17
    Tcr=Tcr_15_005;
end
if sum_n==18
    Tcr=Tcr_16_005;
end
if sum_n==19
    Tcr=Tcr_17_005;
end
if sum_n==20
    Tcr=Tcr_18_005;
end
if sum_n==21
    Tcr=Tcr_19_005;
end
if sum_n==22
    Tcr=Tcr_20_005;
end
if sum_n==23
    Tcr=Tcr_21_005;
end

%control for possible merging and creation of a new DHW_split matrix

%put vector 1 and vector 2 at the same dimension
if vector1_n~=vector2_n
    somma_n=[vector1_n,vector2_n];
    LUN_A=max(somma_n);
    LUN_B=min(somma_n);
    LUN_zeros=LUN_A-LUN_B;
```

```matlab
        corr_mat=zeros(LUN_zeros,1);
        if vector1_n<vector2_n
            vector1=[vector1;corr_mat];
        else
            vector2=[vector2;corr_mat];
        end
    end


%save vector1 and vector2 (or vector mean in ther matrix DHW_split_it2
i=1;
p=1;
LUN_vector=[vector1_n,vector2_n]; %combining length of the vectors to find the
maximum one
LUN=max(LUN_vector);     %find the length of the cicle for

sizee_split2=size(DHW_split_it2);    %for the first iteration this is 0
scorr=sizee_split2(1,2)+1;
control=0;

%operation for merging groups
 if Tcal_iteration1<=Tcr     %if they are not perfectly equal this cicle doesn't
start
     for fk=1:LUN
         if vector1(fk)~=0 && vector2(fk)~=0
         vector_union=[vector1(fk),vector2(fk)];
         vector_union_mean=mean(vector_union);
         DHW_split_it2(i,scorr)=vector_union_mean;
         i=i+1;
         control=3;
        else
            vector_union=[vector1(fk),vector2(fk)];
            vector_union_max=max(vector_union);
            DHW_split_it2(i,scorr)=vector_union_max;
            i=i+1;
            control=4;
        end
     end

 %operation if groups must be stay separate
 else
    lung_matrix=sizee_split2(1,1);
    larg_matrix=sizee_split2(1,2);
    lung_vector1=length(vector1);
    lung_vector2=length(vector2);

    if lung_vector1<lung_matrix
        diff=lung_matrix-lung_vector1;
        zero_corr=zeros(diff,1);
        vector1=[vector1;zero_corr];
    end

    if lung_vector2<lung_matrix
        diff=lung_matrix-lung_vector2;
        zero_corr=zeros(diff,1);
        vector2=[vector2;zero_corr];
    end

    if lung_vector1>lung_matrix
        diff=lung_vector1-lung_matrix;
        zero_corr=zeros(diff,larg_matrix);
        DHW_split_it2=[DHW_split_it2;zero_corr];
```

```
    end
    %it isn't necessary the same control for vector2: I put vector1 and
    %vector2 at the same dimensions.

    %code for not repeating the vector if in the previous case there was no
    %merging. problem with vector1

    colonna=sizee_split2(1,2);

    if colonna~=0
        DHW_split_it2=[DHW_split_it2,vector2];

    else DHW_split_it2=[DHW_split_it2,vector1]; %if we are at the beginning
(colonna=0) put two vector, without control of equality with last vector in
DHW_split_it2
        DHW_split_it2=[DHW_split_it2,vector2];
    end




end

end

end
```

"match.m"

```
%function per il calcolo della matrice match
function
[MATCH]=match(Tcr_46_005,fcr_46_005,W1_T_cal,W1_f_cal,W2_T_cal,W2_f_cal,W3_T_cal
,W3_f_cal,W4_T_cal,W4_f_cal,W5_T_cal,W5_f_cal,W6_T_cal,W6_f_cal,W7_T_cal,W7_f_ca
l,W8_T_cal,W8_f_cal,W9_T_cal,W9_f_cal,W10_T_cal,W10_f_cal,W11_T_cal,W11_f_cal,W1
2_T_cal,W12_f_cal,W13_T_cal,W13_f_cal)

MATCH=zeros(7,7);    %definition of match matrix
f=1;
for j=1:6
    for i=f:6
        value1=W1_T_cal(j,i);
        value2=W1_f_cal(j,i);
        if value1 <= Tcr_46_005 && value2 <= fcr_46_005
            MATCH((i+1),j)= MATCH((i+1),j)+1;
        end

    end
    f=f+1;
end

f=1;
for j=1:6
    for i=f:6
        value1=W2_T_cal(j,i);
        value2=W2_f_cal(j,i);
        if value1 <= Tcr_46_005 && value2 <= fcr_46_005
            MATCH((i+1),j)= MATCH((i+1),j)+1;
        end
```

```matlab
        end
    f=f+1;
end
f=1;
for j=1:6
    for i=f:6
        value1=W3_T_cal(j,i);
        value2=W3_f_cal(j,i);
        if value1 <= Tcr_46_005 && value2 <= fcr_46_005
            MATCH((i+1),j)= MATCH((i+1),j)+1;
        end

    end
    f=f+1;
end
f=1;
for j=1:6
    for i=f:6
        value1=W4_T_cal(j,i);
        value2=W4_f_cal(j,i);
        if value1 <= Tcr_46_005 && value2 <= fcr_46_005
            MATCH((i+1),j)= MATCH((i+1),j)+1;
        end

    end
    f=f+1;
end
f=1;
for j=1:6
    for i=f:6
        value1=W5_T_cal(j,i);
        value2=W5_f_cal(j,i);
        if value1 <= Tcr_46_005 && value2 <= fcr_46_005
            MATCH((i+1),j)= MATCH((i+1),j)+1;
        end

    end
    f=f+1;
end
f=1;
for j=1:6
    for i=f:6
        value1=W6_T_cal(j,i);
        value2=W6_f_cal(j,i);
        if value1 <= Tcr_46_005 && value2 <= fcr_46_005
            MATCH((i+1),j)= MATCH((i+1),j)+1;
        end

    end
    f=f+1;
end
f=1;
for j=1:6
    for i=f:6
        value1=W7_T_cal(j,i);
        value2=W7_f_cal(j,i);
        if value1 <= Tcr_46_005 && value2 <= fcr_46_005
            MATCH((i+1),j)= MATCH((i+1),j)+1;
        end

    end
```

```
        f=f+1;
    end
f=1;
for j=1:6
    for i=f:6
        value1=W8_T_cal(j,i);
        value2=W8_f_cal(j,i);
        if value1 <= Tcr_46_005 && value2 <= fcr_46_005
            MATCH((i+1),j)= MATCH((i+1),j)+1;
        end

    end
    f=f+1;
end
f=1;
for j=1:6
    for i=f:6
        value1=W9_T_cal(j,i);
        value2=W9_f_cal(j,i);
        if value1 <= Tcr_46_005 && value2 <= fcr_46_005
            MATCH((i+1),j)= MATCH((i+1),j)+1;
        end

    end
    f=f+1;
end
f=1;
for j=1:6
    for i=f:6
        value1=W10_T_cal(j,i);
        value2=W10_f_cal(j,i);
        if value1 <= Tcr_46_005 && value2 <= fcr_46_005
            MATCH((i+1),j)= MATCH((i+1),j)+1;
        end

    end
    f=f+1;
end
f=1;
for j=1:6
    for i=f:6
        value1=W11_T_cal(j,i);
        value2=W11_f_cal(j,i);
        if value1 <= Tcr_46_005 && value2 <= fcr_46_005
            MATCH((i+1),j)= MATCH((i+1),j)+1;
        end

    end
    f=f+1;
end
f=1;
for j=1:6
    for i=f:6
        value1=W12_T_cal(j,i);
        value2=W12_f_cal(j,i);
        if value1 <= Tcr_46_005 && value2 <= fcr_46_005
            MATCH((i+1),j)= MATCH((i+1),j)+1;
        end

    end
    f=f+1;
```

```
        end
f=1;
for j=1:6
    for i=f:6
        value1=W13_T_cal(j,i);
        value2=W13_f_cal(j,i);
        if value1 <= Tcr_46_005 && value2 <= fcr_46_005
            MATCH((i+1),j)= MATCH((i+1),j)+1;
        end

    end
    f=f+1;
end
end
```

"min_med_peak_v2.m"

```
%function for the identification of min med peak values

function
[DHW_E_min,DHW_E_max]=min_med_peak_v2(it_DHW_split_last_new,DHW_mean_group1,posi
tion1)
%% identification of peak, average and minimum heat use

%Assign the value of the last split matrix obtained
%(comment the unwanted matrix)
%DHW_split_merged=DHW_split_it11;  %if you want to use brutal force method
DHW_split_merged=it_DHW_split_last_new;  %if you want to use iterative procedure
method

%identification of group 1 and penultimate group
colonna1=DHW_split_merged(:,1); %group 1

sizee_split=size(DHW_split_merged);
k_1=sizee_split(1,2)-1;
colonnak1=DHW_split_merged(:,k_1); %k-1 group (penultimate)

%zero remotion
altezza=sizee_split(1,1);
i=1;
j=1;
group1=[];
groupk1=[];
for fk=1:altezza
    if colonna1(fk)~=0
        group1(i,1)=colonna1(fk);
        i=i+1;
    end

    if colonnak1(fk)~=0
        groupk1(j,1)=colonnak1(fk);
        j=j+1;
    end
end

%mean values
group1_E=mean(group1);
groupk1_E=mean(groupk1);

%number of elements
```

```matlab
group1_M=length(group1);
groupk1_M=length(groupk1);

%selection of Tcr
%definition Tcr
Tcr_1_001=31.821; %M=2 elements
Tcr_2_001=6.965 ; %M=3
Tcr_3_001=4.541;  %M=4
Tcr_4_001=3.747;  %M=5
Tcr_5_001=3.365;  %M=6
Tcr_6_001=3.143;  %M=7
Tcr_7_001=2.998;  %M=8
Tcr_8_001=2.896;  %M=9

%choose the right value of Tcr
%group1
if group1_M==2
    Tcr_group1=Tcr_1_001;
end
if group1_M==3
    Tcr_group1=Tcr_2_001;
end
if group1_M==4
    Tcr_group1=Tcr_3_001;
end
if group1_M==5
    Tcr_group1=Tcr_4_001;
end
if group1_M==6
    Tcr_group1=Tcr_5_001;
end
if group1_M==7
    Tcr_group1=Tcr_6_001;
end
if group1_M==8
    Tcr_group1=Tcr_7_001;
end
if group1_M==9
    Tcr_group1=Tcr_8_001;
end


%groupk1
if groupk1_M==2
    Tcr_groupk1=Tcr_1_001;
end
if groupk1_M==3
    Tcr_groupk1=Tcr_2_001;
end
if groupk1_M==4
    Tcr_groupk1=Tcr_3_001;
end
if groupk1_M==5
    Tcr_groupk1=Tcr_4_001;
end
if groupk1_M==6
    Tcr_groupk1=Tcr_5_001;
end
if groupk1_M==7
    Tcr_groupk1=Tcr_6_001;
end
if groupk1_M==8
```

```matlab
    Tcr_groupk1=Tcr_7_001;
end
if groupk1_M==9
    Tcr_groupk1=Tcr_8_001;
end

%standard deviation
group1_S=std(group1);
groupk1_S=std(groupk1);

%definition limit values
DHW_E_min=group1_E+Tcr_group1*sqrt(((group1_S)^2)/(group1_M));       %Emin
DHW_E_max=groupk1_E+Tcr_groupk1*sqrt(((groupk1_S)^2)/(groupk1_M)); %Emax

%with too low values we have high sensitivity to standard deviation. A
%small variation in this value may cause an error into the definition of
%Emin and Emax. For this reason it is inserted a correction. decrease Emax
%by 7% or increase Emin by 7% if no values are higher than Emax or lower
%than Emin. This is an error due to data and their standard deviation

counter_max=0;
counter_min=0;

for h=1:24
    if DHW_mean_group1(h,1)>DHW_E_max
        counter_max=counter_max+1;
    end
    if DHW_mean_group1(h,1)<DHW_E_min
        counter_min=counter_min+1;
    end
end

if counter_max==0
    DHW_E_max=DHW_E_max-0.07*DHW_E_max;
end

if counter_min==0
    DHW_E_min=DHW_E_min-0.07*DHW_E_min;
end



%plot typical day of group 1 with minimum value and peak value

plot_vector_Emin=zeros(1,24);
plot_vector_Emin=plot_vector_Emin+DHW_E_min;
plot_vector_Emax=zeros(1,24);
plot_vector_Emax=plot_vector_Emax+DHW_E_max;

%plot divided by color
variable=0;
j=1;
k=1;
l=1;
min_load=[];
medium_load=[];
peak_load=[];

index_min=[];
```

```matlab
index_medium=[];
index_peak=[];

for f=1:24
    variable=DHW_mean_group1(f,1);

    if variable <= DHW_E_min
        min_load(j,1)=variable;
        index_min(j,1)=f;
        j=j+1;
    end

    if variable> DHW_E_min && variable< DHW_E_max
        medium_load(k,1)=variable;
        index_medium(k,1)=f;
        k=k+1;
    end

    if variable>= DHW_E_max
        peak_load(l,1)=variable;
        index_peak(l,1)=f;
        l=l+1;
    end
end

%creation of the vector with

%complete vectors "index_min, index_medium and index_peak" with 0 where
%there are no different numbers

color1=[0,0.4470, 0.7410];
color2=[0.8500, 0.3250, 0.0980];
color3=[0.4940, 0.1840, 0.5560];

time=1:24;

fontSize=18;

figure (30)
subplot('Position',position1)
%bar(time,DHW_mean_group1_E) %plot if you want ordered day
bar(index_min,min_load,'FaceColor',color1)    %plot with real day profile
hold on
bar(index_medium,medium_load,'FaceColor',color2)
bar(index_peak,peak_load,'FaceColor',color3)
plot(time,plot_vector_Emin,'color',color2)
plot(time,plot_vector_Emax,'color',color3)
hold off
grid on
xlabel('TIME', 'FontSize', fontSize);
ylabel('AVG E USE', 'FontSize', fontSize);
title('DHW average consumption', 'FontSize', fontSize);
legend('Min load summer','Medium load summer','Peak load
summer','Emin','Emax','location','SouthEast')

end
```

"NEW_PROFILE.m"

```matlab
%function for creating SH and DHW vectors with new modified values day per
%day
function[NEW_MON_Y_DHW_WIN,NEW_MON_Y_SH_WIN]=NEW_PROFILE(DIFF_MON_WIN,E_DHW_Y_MO
N_win,E_SH_Y_MON_win,tolerance,decrease,o)


size_mon_win=size(DIFF_MON_WIN);
limit_4=size_mon_win(1,2);
NEW_MON_Y_DHW_WIN=E_DHW_Y_MON_win;
NEW_MON_Y_SH_WIN=E_SH_Y_MON_win;


for k=1:limit_4
    control_day=E_DHW_Y_MON_win(:,k);
    control_diff_vec=DIFF_MON_WIN(:,k);

    for j=1:24
        control_value=control_day(j);
        control_diff=control_diff_vec(j);

        if control_diff>=tolerance
            value_new=control_value-decrease*control_value;
            NEW_MON_Y_DHW_WIN(j,k)=value_new;
            NEW_MON_Y_SH_WIN(j,k)=NEW_MON_Y_SH_WIN(j,k)+decrease*control_value;
        end

    end
end

%GRAPHICAL VARIABLES
E_SH_Y_MON_win_plot=E_SH_Y_MON_win(:)';    %data from splitting model SH
E_DHW_Y_MON_win_plot=E_DHW_Y_MON_win(:)'; %data from splitting model DHW

fontSize=13;

limit_5=length(E_SH_Y_MON_win_plot);
guide=1:limit_5;

NEW_MON_Y_SH_WIN_plot=NEW_MON_Y_SH_WIN(:)';
NEW_MON_Y_DHW_WIN_plot=NEW_MON_Y_DHW_WIN(:)';

SEPARATION_LINE_X=[504 504];                        %line for separation between
firts 21 weeks and other 16 weeks (there is summer between days before and after
this line)
SEPARATION_LINE_Y=[0 max(E_SH_Y_MON_win_plot)];

axis_limit=[0 limit_5 0 150];

figure(32)
subplot(3,3,o)
plot(guide,E_SH_Y_MON_win_plot,'b')
hold on
plot(guide,E_DHW_Y_MON_win_plot,'color','#D95319')
plot(guide,NEW_MON_Y_SH_WIN_plot,'c')
plot(guide,NEW_MON_Y_DHW_WIN_plot,'color','#EDB120')
plot(SEPARATION_LINE_X,SEPARATION_LINE_Y,'r')
hold off
grid on
axis(axis_limit)
xlabel('TIME', 'FontSize', fontSize);
```

```matlab
ylabel('DEMAND', 'FontSize', fontSize);
title('WINTER DEMAND', 'FontSize', fontSize);
legend('SH from splitting','DHW from splitting','SH modified','DHW
modified','summer line','location','NorthEast')
% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);
end
```

"NEW_PROFILE_YEARLY.m"

```matlab
%FUNCTION for creation of yearly vector for DHW and SH dividied per day
function
[E_SH_Y_MON_NEW,E_DHW_Y_MON_NEW]=NEW_PROFILE_YEARLY(NEW_MON_Y_SH_WIN,E_SH_Y_MON,
NEW_MON_Y_DHW_WIN,E_DHW_Y_MON)

%SH
E_SH_Y_MON_NEW=[];
for k=1:21
    vector_corr=NEW_MON_Y_SH_WIN(:,k);
    E_SH_Y_MON_NEW=[E_SH_Y_MON_NEW,vector_corr];
end
for k=22:36
    vector_corr=E_SH_Y_MON(:,k);
    E_SH_Y_MON_NEW=[E_SH_Y_MON_NEW,vector_corr];
end
for k=22:37
    vector_corr=NEW_MON_Y_SH_WIN(:,k);
    E_SH_Y_MON_NEW=[E_SH_Y_MON_NEW,vector_corr];
end
%DHW
E_DHW_Y_MON_NEW=[];
for k=1:21
    vector_corr=NEW_MON_Y_DHW_WIN(:,k);
    E_DHW_Y_MON_NEW=[E_DHW_Y_MON_NEW,vector_corr];
end
for k=22:36
    vector_corr=E_DHW_Y_MON(:,k);
    E_DHW_Y_MON_NEW=[E_DHW_Y_MON_NEW,vector_corr];
end
for k=22:37
    vector_corr=NEW_MON_Y_DHW_WIN(:,k);
    E_DHW_Y_MON_NEW=[E_DHW_Y_MON_NEW,vector_corr];
end

end
```

"plot_winter_profile.m"

```matlab
%function for plotting winter profile of group1 and group2

function
[]=plot_winter_profile(index_min_group1,index_med_group1,index_max_group1,DHW_WI
N_STD_PROFILE_group1,DHW_E_min_WIN_group1,DHW_E_max_WIN_group1,g)

%selection min, med and max values from standard profile in winter
limit_1=length(index_min_group1);
limit_2=length(index_med_group1);
limit_3=length(index_max_group1);

min_group1_win=[];
med_group1_win=[];
```

```
max_group1_win=[];

for i=1:limit_1
    index=index_min_group1(i);
    min_group1_win(i)=DHW_WIN_STD_PROFILE_group1(index);
end

for i=1:limit_2
    index=index_med_group1(i);
    med_group1_win(i)=DHW_WIN_STD_PROFILE_group1(index);
end

for i=1:limit_3
    index=index_max_group1(i);
    max_group1_win(i)=DHW_WIN_STD_PROFILE_group1(index);
end

%GRAPHICAL VARIABLES
color1=[0,0.4470, 0.7410];
color2=[0.8500, 0.3250, 0.0980];
color3=[0.4940, 0.1840, 0.5560];

time=1:24;

fontSize=18;

plot_vector_Emin_win=zeros(1,24);
plot_vector_Emin_win=plot_vector_Emin_win+DHW_E_min_WIN_group1;
plot_vector_Emax_win=zeros(1,24);
plot_vector_Emax_win=plot_vector_Emax_win+DHW_E_max_WIN_group1;

%PLOT
figure(31)
subplot(3,2,g)
plot(time,plot_vector_Emin_win,'color',color2)
hold on
plot(time,plot_vector_Emax_win,'color',color3)
bar(index_min_group1,min_group1_win,'FaceColor',color1)   %plot with real day
profile
bar(index_med_group1,med_group1_win,'FaceColor',color2)
bar(index_max_group1,max_group1_win,'FaceColor',color3)
grid on
hold off
xlabel('TIME', 'FontSize', fontSize);
ylabel('STANDARD USE', 'FontSize', fontSize);
title('DHW STANDARD PROFILE WINTER', 'FontSize', fontSize);
legend('Emin','Emax','Min load winter','Medium load winter','Peak load
winter','location','SouthEast')
% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

end
```

"SHDHW_split_function.m"

```
%function for splitting DHW and SH, considering hourly vectors

function
[val_1_0AM,val_2_0AM,E_loss_0AM,E_SH_balanced_0AM_corr,E_DHW_0AM_corr]=SHDHW_spl
```

```matlab
it_function(Demand_NOR_SH_0AM,Demand_NOR_DHW_0AM,T_SH_0AM,T_DHW_0AM,DMT,ENERGY_G
LO,CPT_0AM,ENERGY_0AM,L)


Demand_shifted_SH=Demand_NOR_SH_0AM-min(Demand_NOR_DHW_0AM);
Demand_shifted_DHW=Demand_NOR_DHW_0AM-min(Demand_NOR_DHW_0AM);


ESC1_0AM=polyfit(T_SH_0AM,Demand_shifted_SH,1);
ESC2_0AM=polyfit(T_DHW_0AM,Demand_shifted_DHW,1);


val_1_0AM=polyval(ESC1_0AM,T_SH_0AM);
val_2_0AM=polyval(ESC2_0AM,T_DHW_0AM);


%Put values in chronological order with a cicle for
i=1;
value_ESC_0AM=[];

for k=1:365
    if DMT(k)<CPT_0AM
        value_ESC_0AM(i)=polyval(ESC1_0AM,DMT(k));
        i=i+1;
    else
        value_ESC_0AM(i)=polyval(ESC2_0AM,DMT(k));
        i=i+1;
    end
end



%DHW circuit losses evaluation. They are calculated as mean values during
%night consumption in summertime.

%selecte summer period (june-august) of the yearly demand
for k=1:24
    l=1;
    for i=152:1:243
        DHW_losses_data(l,k)=ENERGY_GLO(i,k);
        l=l+1;
    end
end

%Dailymean profile of summer consumption
DHW_losses_sum_pro_0AM=mean(DHW_losses_data);

%Losses determination

E_loss_0AM_mean=mean(DHW_losses_sum_pro_0AM(1:4));   %mean between 0:00 a 4:00

E_loss_year=min(DHW_losses_sum_pro_0AM);              %with minimum value

E_loss_0AM=min(ENERGY_0AM);                          %Choose here what method use
for losses evaluation

%E SH evaluation
E_SH_0AM=value_ESC_0AM-min(value_ESC_0AM);

%E DHW evaluation
E_DHW_0AM=[];
i=1;
for k=1:365
    if ENERGY_0AM(k)>E_SH_0AM(k)
```

```matlab
            E_DHW_0AM(i)=ENERGY_0AM(k)-E_SH_0AM(k)+E_loss_0AM;
            i=i+1;
        else
            E_DHW_0AM(i)=E_loss_0AM;
            i=i+1;
        end
end

%Balancing E SH
E_SH_balanced_0AM=ENERGY_0AM-E_DHW_0AM.';

%Correction for values <0
E_SH_0AM_corr=E_SH_0AM;
E_SH_balanced_0AM_corr=E_SH_balanced_0AM;
E_DHW_0AM_corr=E_DHW_0AM;

for k=1:365
    if E_SH_balanced_0AM(k)<0
        E_DHW_0AM_corr(k)=E_DHW_0AM_corr(k)+E_SH_balanced_0AM(k);
        E_SH_balanced_0AM_corr(k)=0;
    end
    if E_DHW_0AM(k)<0
        E_SH_balanced_0AM_corr(k)=E_SH_balanced_0AM_corr(k)+E_DHW_0AM(k);
        E_DHW_0AM_corr(k)=0;
    end
end

fontSize=18;

figure(L)
subplot(3,3,6)
plot(T_SH_0AM,val_1_0AM,'g')
hold on
grid on
plot(DMT,ENERGY_0AM,'.b')
plot(T_SH_0AM,Demand_NOR_SH_0AM,'r')
plot(T_DHW_0AM,val_2_0AM,'g')
plot(T_DHW_0AM,Demand_NOR_DHW_0AM,'r')
hold off
xlabel('DMT', 'FontSize', fontSize);
ylabel('Demand', 'FontSize', fontSize);
title('ESC complete', 'FontSize', fontSize);
legend('ESC shifted','Measured heat use 0AM','ESC of total heat use')
% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

guida=1:365;
E_loss_0AM_plot(1:365)=E_loss_0AM;
subplot(3,3,7)
plot(guida,ENERGY_0AM,'b')
hold on
grid on
plot(guida,E_loss_0AM_plot,'r')
xlabel('time (day)', 'FontSize', fontSize);
ylabel('Demand', 'FontSize', fontSize);
title('DHW losses evaluation', 'FontSize', fontSize);
legend('Demand','value for DHW losses')
hold off

subplot(3,3,8)
plot(guida,ENERGY_0AM,'b')
```

```
hold on
grid on
plot(guida,E_SH_balanced_0AM_corr,'g')
plot(guida,E_DHW_0AM_corr,'r')
xlabel('time (day)', 'FontSize', fontSize);
ylabel('Demand', 'FontSize', fontSize);
title('chronological demand balanced and correct', 'FontSize', fontSize);
legend('Measured Demand','SH demand balanced >0','DHW demand >0')
hold off

subplot(3,3,9)
plot(DMT,ENERGY_0AM,'.b')
hold on
grid on
plot(DMT,E_SH_balanced_0AM_corr,'.g')
plot(DMT,E_DHW_0AM_corr,'.r')
xlabel('DMT (°C)', 'FontSize', fontSize);
ylabel('Demand', 'FontSize', fontSize);
title('Demand splitted', 'FontSize', fontSize);
legend('Measured Demand','SH demand','DHW demand')
hold off
end
```

"SHDHW_split_function_year.m"

```
%function per la divisione tra DHW e SH per la domanda globale

function
[val_1_0AM,val_2_0AM,E_loss_0AM,E_SH_balanced_0AM_corr,E_DHW_0AM_corr,E_DHW_0AM]
=SHDHW_split_function_year(Demand_NOR_SH_0AM,Demand_NOR_DHW_0AM,T_SH_0AM,T_DHW_0
AM,T,ENERGY_GLO,CPT_0AM,ENERGY_0AM,L)

Demand_shifted_SH=Demand_NOR_SH_0AM-min(Demand_NOR_DHW_0AM);
Demand_shifted_DHW=Demand_NOR_DHW_0AM-min(Demand_NOR_DHW_0AM);

ESC1_0AM=polyfit(T_SH_0AM,Demand_shifted_SH,1);
ESC2_0AM=polyfit(T_DHW_0AM,Demand_shifted_DHW,1);

val_1_0AM=polyval(ESC1_0AM,T_SH_0AM);
val_2_0AM=polyval(ESC2_0AM,T_DHW_0AM);
%attenzione: i valori in val non sono in ordine cronologico ma di
%temperatura, per ottenerlo in ordine cronologico valutarlo sul DMT con
%ciclo for
i=1;
value_ESC_0AM=[];

for k=1:8759
    if T(k)<CPT_0AM
        value_ESC_0AM(i)=polyval(ESC1_0AM,T(k));
        i=i+1;
    else
        value_ESC_0AM(i)=polyval(ESC2_0AM,T(k));
        i=i+1;
    end
end

%ok, adesso ci siamo! ora si può usare il modello :)
%calcolo delle perdite nel circuito DHW. Avendo solo le ore 0 AM dei gionri
%ho deciso di considerare il valore minimo durante l'anno
```

```matlab
%determinazione losses nel circuito DHW. Possono essere calcolate come il
%valore medio del consumo di calore durante la notte.

%selezione periodo giugno-agosto della domanda totale:
for k=1:24
    l=1;
    for i=152:1:243
        DHW_losses_data(l,k)=ENERGY_GLO(i,k);
        l=l+1;
    end
end


%profilo medio giornaliero di domanda estiva:
DHW_losses_sum_pro_0AM=mean(DHW_losses_data);

%determinazione E losses 0 AM

E_loss_0AM_mean=mean(DHW_losses_sum_pro_0AM(1:4));  %ottenuto con media da 0:00
a 4:00


E_loss_year=min(DHW_losses_sum_pro_0AM);             %ottenuto con il valore
minimo

%E_loss_0AM=min(ENERGY_0AM);                         %avendo solo 1 ora,
considero il valore minimo, non vado a costruire il profilo perchè è basato su
tutto l'anno
E_loss_0AM=E_loss_0AM_mean;

%calcolo di E SH
E_SH_0AM=value_ESC_0AM-min(value_ESC_0AM);

%calcolo di E DHW
E_DHW_0AM=[];
i=1;
for k=1:8759
    if ENERGY_0AM(k)>E_SH_0AM(k)
        E_DHW_0AM(i)=ENERGY_0AM(k)-E_SH_0AM(k)+E_loss_0AM;
        i=i+1;
    else
        E_DHW_0AM(i)=E_loss_0AM;
        i=i+1;
    end
end


%bilanciamento di E SH
E_SH_balanced_0AM=ENERGY_0AM-E_DHW_0AM.';

%processo di correzione valori <0
E_SH_0AM_corr=E_SH_0AM;
E_SH_balanced_0AM_corr=E_SH_balanced_0AM;
E_DHW_0AM_corr=E_DHW_0AM;

for k=1:8759
    if E_SH_balanced_0AM(k)<0
        E_DHW_0AM_corr(k)=E_DHW_0AM_corr(k)+E_SH_balanced_0AM(k);
        E_SH_balanced_0AM_corr(k)=0;
    end
    if E_DHW_0AM(k)<0
        E_SH_balanced_0AM_corr(k)=E_SH_balanced_0AM_corr(k)+E_DHW_0AM(k);
        E_DHW_0AM_corr(k)=0;
```

```matlab
    end
end

fontSize=18;

figure(L)
subplot(3,3,6)
plot(T_SH_0AM,val_1_0AM,'g')
hold on
grid on
plot(T,ENERGY_0AM,'.b')
plot(T_SH_0AM,Demand_NOR_SH_0AM,'r')
plot(T_DHW_0AM,val_2_0AM,'g')
plot(T_DHW_0AM,Demand_NOR_DHW_0AM,'r')
hold off
xlabel('DMT', 'FontSize', fontSize);
ylabel('Demand', 'FontSize', fontSize);
title('ESC complete', 'FontSize', fontSize);
legend('ESC shifted','Measured heat use 0AM','ESC of total heat use')
% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

guida=1:8759;
E_loss_0AM_plot(1:8759)=E_loss_0AM;
subplot(3,3,7)
plot(guida,ENERGY_0AM,'b')
hold on
grid on
plot(guida,E_loss_0AM_plot,'r')
xlabel('time (day)', 'FontSize', fontSize);
ylabel('Demand', 'FontSize', fontSize);
title('Losses DHW circuit', 'FontSize', fontSize);
legend('Demand','value for DHW losses')
hold off

completo=[val_1_0AM,val_2_0AM];
% subplot(3,3,1)
% plot(guida,ENERGY_0AM,'b')
% hold on
% grid on
% plot(guida,E_SH_balanced_0AM,'g')
% plot(guida,E_DHW_0AM,'r')
% hold off
% xlabel('time (day)', 'FontSize', fontSize);
% ylabel('Demand 0AM', 'FontSize', fontSize);
% title('cronological Demand balanced but <0', 'FontSize', fontSize);
% legend('Measured Demand','SH demand balanced','DHW demand')

subplot(3,3,8)
plot(guida,ENERGY_0AM,'b')
hold on
grid on
plot(guida,E_SH_balanced_0AM_corr,'g')
plot(guida,E_DHW_0AM_corr,'r')
xlabel('time (day)', 'FontSize', fontSize);
ylabel('Demand', 'FontSize', fontSize);
title('cronological Demand balanced and correct', 'FontSize', fontSize);
legend('Measured Demand','SH demand balanced >0','DHW demand >0')
hold off

subplot(3,3,9)
```

```matlab
plot(T,ENERGY_0AM,'.b')
hold on
grid on
plot(T,E_SH_balanced_0AM_corr,'.g')
plot(T,E_DHW_0AM_corr,'.r')
xlabel('DMT (°C)', 'FontSize', fontSize);
ylabel('Demand', 'FontSize', fontSize);
title('Demand splitted', 'FontSize', fontSize);
legend('Measured Demand','SH demand','DHW demand')
hold off
end
```

"std_profile_v1"

```matlab
%function for the creation of the standard summer profile
function
[DHW_STD_PROFILE_group1,index_min,index_med,index_max]=std_profile_v1(DHW_mean_g
roup1,DHW_E_min_group1,DHW_E_max_group1,g)
%% creation of the standard profile for summer
limit_1=length(DHW_mean_group1);
i=1;
j=1;
f=1;
l=1;
DHW_STD_PROFILE_group1=[];
index_min=[];
index_med=[];
index_max=[];

min_load_STD_group1=[];
medium_load_STD_group1=[];
peak_load_STD_group1=[];

for k=1:limit_1
    control_E=DHW_mean_group1(k);

    if control_E<=DHW_E_min_group1
        DHW_STD_PROFILE_group1(i)=DHW_E_min_group1;
        i=i+1;
        min_load_STD_group1(j,1)=DHW_E_min_group1;
        index_min(j,1)=k;
        j=j+1;
    end

    if control_E>DHW_E_min_group1 && control_E<=DHW_E_max_group1
        DHW_STD_PROFILE_group1(i)=(DHW_E_min_group1+DHW_E_max_group1)/2;
        i=i+1;
        medium_load_STD_group1(f,1)=(DHW_E_min_group1+DHW_E_max_group1)/2;
        index_med(f,1)=k;
        f=f+1;
    end

    if control_E>DHW_E_max_group1
        DHW_STD_PROFILE_group1(i)=DHW_E_max_group1;
        i=i+1;
        peak_load_STD_group1(l,1)=DHW_E_max_group1;
        index_max(l,1)=k;
        l=l+1;
    end
end
```

```matlab
%GRAPHICAL VARIABLES
color1=[0,0.4470, 0.7410];
color2=[0.8500, 0.3250, 0.0980];
color3=[0.4940, 0.1840, 0.5560];

axis_limit=[0 25 0 20];

time=1:24;

fontSize=18;

plot_vector_Emin=zeros(1,24);
plot_vector_Emin=plot_vector_Emin+DHW_E_min_group1;
plot_vector_Emax=zeros(1,24);
plot_vector_Emax=plot_vector_Emax+DHW_E_max_group1;

%PLOT
figure(31)
subplot(3,2,g)
plot(time,plot_vector_Emin,'color',color2)
hold on
plot(time,plot_vector_Emax,'color',color3)
bar(index_min,min_load_STD_group1,'FaceColor',color1)   %plot with real day
profile
bar(index_med,medium_load_STD_group1,'FaceColor',color2)
bar(index_max,peak_load_STD_group1,'FaceColor',color3)
grid on
hold off
axis(axis_limit)
xlabel('TIME', 'FontSize', fontSize);
ylabel('STANDARD USE', 'FontSize', fontSize);
title('DHW STANDARD PROFILE SUMMER', 'FontSize', fontSize);
legend('Emin','Emax','Min load summer','Medium load summer','Peak load
summer','location','SouthEast')
% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

end
```

"stud_fish.m"

```matlab
%function for application of Student's t-test and Fisher criterion
%to apply for every week
function
[W1_T_cal,W1_f_cal]=stud_fish(DHW_monday,DHW_tuesday,DHW_wednesday,DHW_thursday,
DHW_friday,DHW_saturday,DHW_sunday,k)
%WEEK 1
W1_E_proof=[];
W1_n_proof=[];
W1_S_proof=[];

%E proof
W1_E_proof(1,1)=mean(DHW_monday(:,k));
W1_E_proof(1,2)=mean(DHW_tuesday(:,k));
W1_E_proof(1,3)=mean(DHW_wednesday(:,k));
W1_E_proof(1,4)=mean(DHW_thursday(:,k));
W1_E_proof(1,5)=mean(DHW_friday(:,k));
W1_E_proof(1,6)=mean(DHW_saturday(:,k));
W1_E_proof(1,7)=mean(DHW_sunday(:,k));
```

```matlab
%n proof
W1_n_proof(1,1:7)=24;

%S proof
W1_S_proof(1,1)=std(DHW_monday(:,k));
W1_S_proof(1,2)=std(DHW_tuesday(:,k));
W1_S_proof(1,3)=std(DHW_wednesday(:,k));
W1_S_proof(1,4)=std(DHW_thursday(:,k));
W1_S_proof(1,5)=std(DHW_friday(:,k));
W1_S_proof(1,6)=std(DHW_saturday(:,k));
W1_S_proof(1,7)=std(DHW_sunday(:,k));

%Tcal
f=2;
W1_T_cal=[];

for j=1:(7-1) %ok perchè sicuramente sono sempre 7 i giorni

    for i=f:7
        W1_T_cal(j,(i-1))=(W1_E_proof(1,j)-
W1_E_proof(1,i))/(sqrt(((W1_S_proof(1,j))^2/24)+((W1_S_proof(1,i))^2/24)));
    end

    f=f+1;
end

W1_T_cal=abs(W1_T_cal);

%definition critical values Student's t-test
%from https://www.itl.nist.gov/div898/handbook/eda/section3/eda3672.htm
Tcr_46_005=2.687;
Tcr_46_001=3.277;

%% FISHER'S CRITERION
%fcal
f=2;
W1_f_cal=[];

for j=1:(7-1) %ok perchè sicuramente sono sempre 7 i giorni

    for i=f:7
        vector(1,1)=(W1_S_proof(1,j))^2;
        vector(1,2)=(W1_S_proof(1,i))^2;
        W1_f_cal(j,(i-1))=(max(vector))/(min(vector));
    end

    f=f+1;
end

%definition critical value Fisher criterion
%from https://socr.umich.edu/Applets/F_Table.html

fcr_46_005=1.9838;

end
```

"T_cal_splitting.m"

```matlab
%function for splitting components of groups into vlues according to Tcal.

function [DHW_split_g2]=Tcal_splitting(Tcal_vector_1,DHW_mean_group1_E)
%% splitting in different group according to values in Tcal

%inserire nuovi vettori
% Tcal_vector_1=Tcal_vector_2;
% DHW_mean_group1_E=DHW_mean_group2_E;
% DHW_mean_group1=DHW_mean_group2;

%1 split <
i=1;
DHW_split_g2=[];
col=1;
Dimension_k=[];
p=1;
fk=1;
t=1;
for j=fk:23
    if i<23
    if Tcal_vector_1(i+1)<Tcal_vector_1(i)
        Dimension_k(p)=DHW_mean_group1_E(t);
        i=i+1;
        p=p+1;
        t=t+1;
        fk=fk+1;
    else end
    end
end
Dimension_k(p)=DHW_mean_group1_E(t); %insertive succesive value

Dimension_k=Dimension_k';

sizee_split=size(DHW_split_g2); %definition of length of matrix and vector
s_split=sizee_split(1,1);
s_Dimension=length(Dimension_k);

%ciclo if: se è più lungo il vettore Dimension_k uso quello come lunghezza
%e aggiungo zeri alla matrice DHW_split, altrimenti devo procede come
%descritto finora. inserire questa parte anche negli scrip precedenti.

if s_split>=s_Dimension
    vector_zeros_L=s_split-length(Dimension_k);
    vector_zero=zeros(vector_zeros_L,1);
    Dimension_k_mod=cat(1,Dimension_k,vector_zero);
    DHW_split_g2=[DHW_split_g2,Dimension_k_mod];
else
    raw=s_Dimension-s_split;
    colomn=sizee_split(1,2);
    matrix=zeros(raw,colomn);
    DHW_split_2=cat(1,DHW_split_g2,matrix);
    DHW_split_g2=[DHW_split_2,Dimension_k];
end

%2 split >
if i<23
Dimension_k=[];
p=1;
```

```matlab
%k=1;
for j=fk:23
    if i<23
    if Tcal_vector_1(i+1)>Tcal_vector_1(i)
        Dimension_k(p)=DHW_mean_group1_E(t);
        i=i+1;
        p=p+1;
        t=t+1;
        fk=fk+1;
    else end
    end
end
Dimension_k(p)=DHW_mean_group1_E(t); %insertive succesive value

Dimension_k=Dimension_k';

sizee_split=size(DHW_split_g2); %definition of length of matrix and vector
s_split=sizee_split(1,1);
s_Dimension=length(Dimension_k);

if s_split>=s_Dimension
    vector_zeros_L=s_split-length(Dimension_k);
    vector_zero=zeros(vector_zeros_L,1);
    Dimension_k_mod=cat(1,Dimension_k,vector_zero);
    DHW_split_g2=[DHW_split_g2,Dimension_k_mod];
else
    raw=s_Dimension-s_split;
    colomn=sizee_split(1,2);
    matrix=zeros(raw,colomn);
    DHW_split_2=cat(1,DHW_split_g2,matrix);
    DHW_split_g2=[DHW_split_2,Dimension_k];
end
end

%3 split <
if i<23
Dimension_k=[];
p=1;
%k=1;
for j=fk:23
    if i<23
    if Tcal_vector_1(i+1)<Tcal_vector_1(i)
        Dimension_k(p)=DHW_mean_group1_E(t);
        i=i+1;
        p=p+1;
        t=t+1;
        fk=fk+1;
    else end
    end
end
Dimension_k(p)=DHW_mean_group1_E(t); %insertive succesive value

Dimension_k=Dimension_k';

sizee_split=size(DHW_split_g2); %definition of length of matrix and vector
s_split=sizee_split(1,1);
s_Dimension=length(Dimension_k);

if s_split>=s_Dimension
    vector_zeros_L=s_split-length(Dimension_k);
    vector_zero=zeros(vector_zeros_L,1);
```

```matlab
    Dimension_k_mod=cat(1,Dimension_k,vector_zero);
    DHW_split_g2=[DHW_split_g2,Dimension_k_mod];
else
    raw=s_Dimension-s_split;
    colomn=sizee_split(1,2);
    matrix=zeros(raw,colomn);
    DHW_split_2=cat(1,DHW_split_g2,matrix);
    DHW_split_g2=[DHW_split_2,Dimension_k];
end
end


%4 split >
if i<23
Dimension_k=[];
p=1;
%k=1;
for j=fk:23
    if i<23
    if Tcal_vector_1(i+1)>Tcal_vector_1(i)
        Dimension_k(p)=DHW_mean_group1_E(t);
        i=i+1;
        p=p+1;
        t=t+1;
        fk=fk+1;
    else end
    end
end
Dimension_k(p)=DHW_mean_group1_E(t); %insertive succesive value

Dimension_k=Dimension_k';

sizee_split=size(DHW_split_g2); %definition of length of matrix and vector
s_split=sizee_split(1,1);
s_Dimension=length(Dimension_k);

if s_split>=s_Dimension
    vector_zeros_L=s_split-length(Dimension_k);
    vector_zero=zeros(vector_zeros_L,1);
    Dimension_k_mod=cat(1,Dimension_k,vector_zero);
    DHW_split_g2=[DHW_split_g2,Dimension_k_mod];
else
    raw=s_Dimension-s_split;
    colomn=sizee_split(1,2);
    matrix=zeros(raw,colomn);
    DHW_split_2=cat(1,DHW_split_g2,matrix);
    DHW_split_g2=[DHW_split_2,Dimension_k];
end
end



%5 split <
if i<23
Dimension_k=[];
p=1;
%k=1;
for j=fk:23
    if i<23
    if Tcal_vector_1(i+1)<Tcal_vector_1(i)
        Dimension_k(p)=DHW_mean_group1_E(t);
        i=i+1;
        p=p+1;
```

```matlab
        t=t+1;
        fk=fk+1;
    else end
    end
end
Dimension_k(p)=DHW_mean_group1_E(t); %insertive succesive value

Dimension_k=Dimension_k';

sizee_split=size(DHW_split_g2); %definition of length of matrix and vector
s_split=sizee_split(1,1);
s_Dimension=length(Dimension_k);

if s_split>=s_Dimension
    vector_zeros_L=s_split-length(Dimension_k);
    vector_zero=zeros(vector_zeros_L,1);
    Dimension_k_mod=cat(1,Dimension_k,vector_zero);
    DHW_split_g2=[DHW_split_g2,Dimension_k_mod];
else
    raw=s_Dimension-s_split;
    colomn=sizee_split(1,2);
    matrix=zeros(raw,colomn);
    DHW_split_2=cat(1,DHW_split_g2,matrix);
    DHW_split_g2=[DHW_split_2,Dimension_k];
end
end

%6 split >
if i<23
Dimension_k=[];
p=1;
%k=1;
for j=fk:23
    if i<23
    if Tcal_vector_1(i+1)>Tcal_vector_1(i)
        Dimension_k(p)=DHW_mean_group1_E(t);
        i=i+1;
        p=p+1;
        t=t+1;
        fk=fk+1;
    else end
    end
end
Dimension_k(p)=DHW_mean_group1_E(t); %insertive succesive value

Dimension_k=Dimension_k';

sizee_split=size(DHW_split_g2); %definition of length of matrix and vector
s_split=sizee_split(1,1);
s_Dimension=length(Dimension_k);

if s_split>=s_Dimension
    vector_zeros_L=s_split-length(Dimension_k);
    vector_zero=zeros(vector_zeros_L,1);
    Dimension_k_mod=cat(1,Dimension_k,vector_zero);
    DHW_split_g2=[DHW_split_g2,Dimension_k_mod];
else
    raw=s_Dimension-s_split;
    colomn=sizee_split(1,2);
    matrix=zeros(raw,colomn);
    DHW_split_2=cat(1,DHW_split_g2,matrix);
```

```matlab
    DHW_split_g2=[DHW_split_2,Dimension_k];
end
end


%7 split <
if i<23
Dimension_k=[];
p=1;
%k=1;
for j=fk:23
    if i<23
    if Tcal_vector_1(i+1)<Tcal_vector_1(i)
        Dimension_k(p)=DHW_mean_group1_E(t);
        i=i+1;
        p=p+1;
        t=t+1;
        fk=fk+1;
    else end
    end
end
Dimension_k(p)=DHW_mean_group1_E(t); %insertive succesive value

Dimension_k=Dimension_k';

sizee_split=size(DHW_split_g2); %definition of length of matrix and vector
s_split=sizee_split(1,1);
s_Dimension=length(Dimension_k);

if s_split>=s_Dimension
    vector_zeros_L=s_split-length(Dimension_k);
    vector_zero=zeros(vector_zeros_L,1);
    Dimension_k_mod=cat(1,Dimension_k,vector_zero);
    DHW_split_g2=[DHW_split_g2,Dimension_k_mod];
else
    raw=s_Dimension-s_split;
    colomn=sizee_split(1,2);
    matrix=zeros(raw,colomn);
    DHW_split_2=cat(1,DHW_split_g2,matrix);
    DHW_split_g2=[DHW_split_2,Dimension_k];
end
end

%8 split >
if i<23
Dimension_k=[];
p=1;
%k=1;
for j=fk:23
    if i<23
    if Tcal_vector_1(i+1)>Tcal_vector_1(i)
        Dimension_k(p)=DHW_mean_group1_E(t);
        i=i+1;
        p=p+1;
        t=t+1;
        fk=fk+1;
    else end
    end
end
Dimension_k(p)=DHW_mean_group1_E(t); %insertive succesive value
```

```matlab
Dimension_k=Dimension_k';

sizee_split=size(DHW_split_g2); %definition of length of matrix and vector
s_split=sizee_split(1,1);
s_Dimension=length(Dimension_k);

if s_split>=s_Dimension
    vector_zeros_L=s_split-length(Dimension_k);
    vector_zero=zeros(vector_zeros_L,1);
    Dimension_k_mod=cat(1,Dimension_k,vector_zero);
    DHW_split_g2=[DHW_split_g2,Dimension_k_mod];
else
    raw=s_Dimension-s_split;
    colomn=sizee_split(1,2);
    matrix=zeros(raw,colomn);
    DHW_split_2=cat(1,DHW_split_g2,matrix);
    DHW_split_g2=[DHW_split_2,Dimension_k];
end
end


%9 split <
if i<23
Dimension_k=[];
p=1;
%k=1;
for j=fk:23
    if i<23
    if Tcal_vector_1(i+1)<Tcal_vector_1(i)
        Dimension_k(p)=DHW_mean_group1_E(t);
        i=i+1;
        p=p+1;
        t=t+1;
        fk=fk+1;
    else
    end
    end
end
Dimension_k(p)=DHW_mean_group1_E(t); %insertive succesive value

Dimension_k=Dimension_k';

sizee_split=size(DHW_split_g2); %definition of length of matrix and vector
s_split=sizee_split(1,1);
s_Dimension=length(Dimension_k);

if s_split>=s_Dimension
    vector_zeros_L=s_split-length(Dimension_k);
    vector_zero=zeros(vector_zeros_L,1);
    Dimension_k_mod=cat(1,Dimension_k,vector_zero);
    DHW_split_g2=[DHW_split_g2,Dimension_k_mod];
else
    raw=s_Dimension-s_split;
    colomn=sizee_split(1,2);
    matrix=zeros(raw,colomn);
    DHW_split_2=cat(1,DHW_split_g2,matrix);
    DHW_split_g2=[DHW_split_2,Dimension_k];
end
end


%10 split >
if i<23
```

```
Dimension_k=[];
p=1;
fk=fk-1;
for j=fk:23
    if i<23

        if Tcal_vector_1(i+1)>Tcal_vector_1(i)
            Dimension_k(p)=DHW_mean_group1_E(t);
            i=i+1;
            p=p+1;
            t=t+1;
          fk=fk+1;
        else
        end
    else
        lun=length(Dimension_k);
        Dimension_k(lun+1)=DHW_mean_group1_E(24);

    end

end
Dimension_k(p)=DHW_mean_group1_E(t); %insertive succesive value

Dimension_k=Dimension_k';

sizee_split=size(DHW_split_g2); %definition of length of matrix and vector
s_split=sizee_split(1,1);
s_Dimension=length(Dimension_k);

if s_split>=s_Dimension
    vector_zeros_L=s_split-length(Dimension_k);
    vector_zero=zeros(vector_zeros_L,1);
    Dimension_k_mod=cat(1,Dimension_k,vector_zero);
    DHW_split_g2=[DHW_split_g2,Dimension_k_mod];
else
    raw=s_Dimension-s_split;
    colomn=sizee_split(1,2);
    matrix=zeros(raw,colomn);
    DHW_split_2=cat(1,DHW_split_g2,matrix);
    DHW_split_g2=[DHW_split_2,Dimension_k];
end
end

%insert also control i<23: in this way the stap is done only if we have not
%inserted all DHW_mean_group1_E values.

%do the t_test amond two different neighboring groups (adiacent coloumns of
%the matrix DHW_split) and compare Tcal with critical vallues to see if it
%is possible to merge neighboring groups:

end
```

"time_interval.m"

```
%function for the application of the step 2.2, the iterative process for
%calculating R1 and R2 from the initial E samples.

function
[Tcal_vector_1,Tcal_vector_2]=time_interval(DHW_mean_group1_E,DHW_mean_group2_E)
```

```matlab
%iterative process for the creation of R1 and R2
size1=size(DHW_mean_group1_E);
size2=size(DHW_mean_group2_E);
Tcal_vector_1=[];


for i=1:size1(1,2)
    dayy=DHW_mean_group1_E(:,i);

    for M=1:23
        %creation of vector R1 and R2 for every iteration
        R1=zeros(M,1);
        R2=zeros((M+1),1);

        %insert elements in R1 and R2
        for k=1:M
            R1(k,1)=dayy(k,1);
        end
        for o=1:(M+1)
            R2(o,1)=dayy(o,1);
        end

        %find E mean 1 and 2
        E_mean_R1=mean(R1);
        E_mean_R2=mean(R2);

        %find S1 and S2
        S_R1=std(R1);
        S_R2=std(R2);

        %find n of elements in 1 and 2
        n_R1=length(R1);
        n_R2=length(R2);

        %calculate and save Tcal
        Tcal_vector_1(M,i)=abs((E_mean_R1-
E_mean_R2)/(sqrt(((S_R1)^2/(n_R1))+((S_R2)^2/(n_R2)))));
        %Tcal_vector_1(M,i)=(E_mean_R1-
E_mean_R2)/(sqrt(((S_R1)^2/(n_R1))+((S_R2)^2/(n_R2))));
    end
end

Tcal_vector_2=[];
for i=1:size1(1,2)
    dayy=DHW_mean_group2_E(:,i);

    for M=1:23
        %creation of vector R1 and R2 for every iteration
        R1=zeros(M,1);
        R2=zeros((M+1),1);

        %insert elements in R1 and R2
        for k=1:M
            R1(k,1)=dayy(k,1);
        end
        for o=1:(M+1)
            R2(o,1)=dayy(o,1);
        end

        %find E mean 1 and 2
```

```matlab
        E_mean_R1=mean(R1);
        E_mean_R2=mean(R2);

        %find S1 and S2
        S_R1=std(R1);
        S_R2=std(R2);

        %find n of elements in 1 and 2
        n_R1=length(R1);
        n_R2=length(R2);

        %calculate and save Tcal
        Tcal_vector_2(M,i)=abs((E_mean_R1-
E_mean_R2)/(sqrt(((S_R1).^2/(n_R1))+((S_R2).^2/(n_R2)))));
        %Tcal_vector_2(M,i)=(E_mean_R1-
E_mean_R2)/(sqrt(((S_R1).^2/(n_R1))+((S_R2).^2/(n_R2))));
    end
end


end
```

"DHW_model_10922_v3.m"

```matlab
%script analysis and model of DHW
%the summer demand is supposed to be only DHW, infact it is the part
%T-independent of the ESC. From this profile the goal is to obtain a sort
%of model also for the rest of the year and to use it to smooth the demand
%and to improve the SH + DHW splitting.

%from:
%Development and analysis of hourly DHW heat use profiles in nursing homes in
Norway

%% step 1: Comparing similarity of DHW heat use profiles in different days of
the week

%selection and division of days of the week

%considered only summer period. From a graphical analysis it is evidente,
%for this building, that the summer period goes:
%from monday 20 may to sunday 15 september. It is possible that these are
%data for the specific year of 2019, so to obtain a more general value we
%assume the summer period is from june to august:
%from saturday 1 june to saturday 31 august.

time=1:24;

%SATURDAY
%change the value of colomn according to the calendar of the considered
%year
DHW_saturday=ones(24,14);
f=3624;     %the first hour of the day in summer period
for j=1:14
    for i=1:24
        DHW_saturday(i,j)=ENERGY(f);
        f=f+1;
    end
```

```matlab
    f=f+144;
end


%SUNDAY
%change the value of colomn according to the calendar of the considered
%year
DHW_sunday=ones(24,13);
f=3648;      %the first hour of the day in summer period
for j=1:13
    for i=1:24
        DHW_sunday(i,j)=ENERGY(f);
        f=f+1;
    end
    f=f+144;
end


%MONDAY
%change the value of colomn according to the calendar of the considered
%year
DHW_monday=ones(24,13);
f=3672;      %the first hour of the day in summer period
for j=1:13
    for i=1:24
        DHW_monday(i,j)=ENERGY(f);
        f=f+1;
    end
    f=f+144;
end


%TUESDAY
%change the value of colomn according to the calendar of the considered
%year
DHW_tuesday=ones(24,13);
f=3696;      %the first hour of the day in summer period
for j=1:13
    for i=1:24
        DHW_tuesday(i,j)=ENERGY(f);
        f=f+1;
    end
    f=f+144;
end


%WEDNESDAY
%change the value of colomn according to the calendar of the considered
%year
DHW_wednesday=ones(24,13);
f=3720;      %the first hour of the day in summer period
for j=1:13
    for i=1:24
        DHW_wednesday(i,j)=ENERGY(f);
        f=f+1;
    end
    f=f+144;
end


%THURSDAY
%change the value of colomn according to the calendar of the considered
%year
DHW_thursday=ones(24,13);
f=3744;      %the first hour of the day in summer period
for j=1:13
```

```
        for i=1:24
            DHW_thursday(i,j)=ENERGY(f);
            f=f+1;
        end
        f=f+144;
end


%FRIDAY
%change the value of colomn according to the calendar of the considered
%year
DHW_friday=ones(24,13);
f=3768;      %the first hour of the day in summer period
for j=1:13
        for i=1:24
            DHW_friday(i,j)=ENERGY(f);
            f=f+1;
        end
        f=f+144;
end



%% STUDENT'S T-TEST AND FISHER CRITERION (SAME WEEK, DIFFERENT DAYS)

%WEEK 1
[W1_T_cal,W1_f_cal]=stud_fish(DHW_monday,DHW_tuesday,DHW_wednesday,DHW_thursday,
DHW_friday,DHW_saturday,DHW_sunday,1);
%WEEK 2
[W2_T_cal,W2_f_cal]=stud_fish(DHW_monday,DHW_tuesday,DHW_wednesday,DHW_thursday,
DHW_friday,DHW_saturday,DHW_sunday,2);
%WEEK 3
[W3_T_cal,W3_f_cal]=stud_fish(DHW_monday,DHW_tuesday,DHW_wednesday,DHW_thursday,
DHW_friday,DHW_saturday,DHW_sunday,3);
%WEEK 4
[W4_T_cal,W4_f_cal]=stud_fish(DHW_monday,DHW_tuesday,DHW_wednesday,DHW_thursday,
DHW_friday,DHW_saturday,DHW_sunday,4);
%WEEK 5
[W5_T_cal,W5_f_cal]=stud_fish(DHW_monday,DHW_tuesday,DHW_wednesday,DHW_thursday,
DHW_friday,DHW_saturday,DHW_sunday,5);
%WEEK 6
[W6_T_cal,W6_f_cal]=stud_fish(DHW_monday,DHW_tuesday,DHW_wednesday,DHW_thursday,
DHW_friday,DHW_saturday,DHW_sunday,6);
%WEEK 7
[W7_T_cal,W7_f_cal]=stud_fish(DHW_monday,DHW_tuesday,DHW_wednesday,DHW_thursday,
DHW_friday,DHW_saturday,DHW_sunday,7);
%WEEK 8
[W8_T_cal,W8_f_cal]=stud_fish(DHW_monday,DHW_tuesday,DHW_wednesday,DHW_thursday,
DHW_friday,DHW_saturday,DHW_sunday,8);
%WEEK 9
[W9_T_cal,W9_f_cal]=stud_fish(DHW_monday,DHW_tuesday,DHW_wednesday,DHW_thursday,
DHW_friday,DHW_saturday,DHW_sunday,9);
%WEEK 10
[W10_T_cal,W10_f_cal]=stud_fish(DHW_monday,DHW_tuesday,DHW_wednesday,DHW_thursda
y,DHW_friday,DHW_saturday,DHW_sunday,10);
%WEEK 11
[W11_T_cal,W11_f_cal]=stud_fish(DHW_monday,DHW_tuesday,DHW_wednesday,DHW_thursda
y,DHW_friday,DHW_saturday,DHW_sunday,11);
%WEEK 12
[W12_T_cal,W12_f_cal]=stud_fish(DHW_monday,DHW_tuesday,DHW_wednesday,DHW_thursda
y,DHW_friday,DHW_saturday,DHW_sunday,12);
%WEEK 13
```

```matlab
[W13_T_cal,W13_f_cal]=stud_fish(DHW_monday,DHW_tuesday,DHW_wednesday,DHW_thursda
y,DHW_friday,DHW_saturday,DHW_sunday,13);

%definition critical values Student's t-test
%from https://www.itl.nist.gov/div898/handbook/eda/section3/eda3672.htm
Tcr_46_005=2.687;
Tcr_46_001=3.277;

%definition critical value Fisher criterion
%from https://socr.umich.edu/Applets/F_Table.html

fcr_46_005=1.9838;

%% conclusion of the tests: comparison with critical values
[MATCH]=match(Tcr_46_001,fcr_46_005,W1_T_cal,W1_f_cal,W2_T_cal,W2_f_cal,W3_T_cal
,W3_f_cal,W4_T_cal,W4_f_cal,W5_T_cal,W5_f_cal,W6_T_cal,W6_f_cal,W7_T_cal,W7_f_ca
l,W8_T_cal,W8_f_cal,W9_T_cal,W9_f_cal,W10_T_cal,W10_f_cal,W11_T_cal,W11_f_cal,W1
2_T_cal,W12_f_cal,W13_T_cal,W13_f_cal);
%choose the value for Tcr with k=0.01 for a major similiraty spectrum

%insert diagonal values
fk=13; %insert here number of the weeks
for j=1:7
    MATCH(j,j)=fk;
end

%found percentage
MATCH=MATCH.*(100/fk);

%% Analysis of MATCH matrix
%according to the article and by considering the fact that we are using a
%smaller sample of values, the acceptable error is 25% (in article is 14%)

%from this we can assume that there are 2 profiles:
%friday
%rest of the day

%mathematical approach to the division of the profiles of DHW.
%select the profiles you want to use for the analysis

%combining together days included in the same profile
DHW_group1=DHW_friday;                       %insert here the days of the week
belonging to group 1
DHW_group2=[DHW_monday,DHW_tuesday,DHW_wednesday,DHW_thursday,DHW_saturday,DHW_s
unday]; %insert here the days of the week belonging to group 2
%include eventually extra groups

%calculating mean profiles
DHW_mean_group1=[];
DHW_mean_group2=[];
size1=size(DHW_group1);
size2=size(DHW_group2);
s_split=1;
for i=1:24
    DHW_mean_group1(s_split,1)=mean(DHW_group1(i,:));
    s_split=s_split+1;
end

s_split=1;
for i=1:24
```

```matlab
        DHW_mean_group2(s_split,1)=mean(DHW_group2(i,:));
        s_split=s_split+1;
end


%plot of the results from the division
%add plots if there are more profiles or if there is the possibility to
%have also winter consumption.

figure (30)
subplot(4,2,1)
bar(time,DHW_mean_group1)
grid on
xlabel('TIME', 'FontSize', fontSize);
ylabel('AVG E USE', 'FontSize', fontSize);
title('Group1 DHW average consumption', 'FontSize', fontSize);
legend('Summer')

subplot(4,2,2)
bar(time,DHW_mean_group2)
grid on
xlabel('TIME', 'FontSize', fontSize);
ylabel('AVG E USE', 'FontSize', fontSize);
title('Group2 DHW average consumption', 'FontSize', fontSize);
legend('Summer')

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

%% step 2: Determining the time zones with peak, minimum, and average heat load
for daily profiles of DHW heat use.
%division the hours of DHW heat use into several groups with statistically
%different mean values within each group.

%STEP 2.1 Sorting the elements of the sample in the order of their increase

%group 1
DHW_group1_E=zeros(24,13);
DHW_group2_E=zeros(24,79);

for j=1:13
    coloumn=DHW_group1(:,j);
    coloumn_sorted=sort(coloumn);
    DHW_group1_E(:,j)=coloumn_sorted;
end

%group2
for j=1:79
    coloumn=DHW_group2(:,j);
    coloumn_sorted=sort(coloumn);
    DHW_group2_E(:,j)=coloumn_sorted;
end

%STEP 2.2 Identifying the initial groups for the elements that could be
considered statistically similar
%with average profile in group 1 and group 2

%Find the average of the sorted group1 and group2
DHW_mean_group1_E=zeros(24,1);
DHW_mean_group2_E=zeros(24,1);
s_split=1;
```

```matlab
for i=1:24
    DHW_mean_group1_E(s_split,1)=mean(DHW_group1_E(i,:));
    s_split=s_split+1;
end

s_split=1;
for i=1:24
    DHW_mean_group2_E(s_split,1)=mean(DHW_group2_E(i,:));
    s_split=s_split+1;
end

DHW_mean_group1_E=sort(DHW_mean_group1);
DHW_mean_group2_E=sort(DHW_mean_group2);
%function for finding Tcal group1 and Tcal group 2
[Tcal_vector_1,Tcal_vector_2]=time_interval(DHW_mean_group1_E,DHW_mean_group2_E)
;

%At this point we have the vector Tcal_vector containing Tcal between
%different R1 and R2 groups. the next step is to use these values to divide
%the sample into multiple groups and check the possibility to merge them.
%The goal is to obtain the lowest number of group as possibile within each
%sample.

%plot Tcal group1
figure (30)
subplot (4,2,3)
plot(time(1:23),Tcal_vector_1)
grid on
xlabel('TIME', 'FontSize', fontSize);
ylabel('Tcal', 'FontSize', fontSize);
title('Tcal group 1', 'FontSize', fontSize);

%plot Tcal group 2
figure (30)
subplot (4,2,4)
plot(time(1:23),Tcal_vector_2)
grid on
xlabel('TIME', 'FontSize', fontSize);
ylabel('Tcal', 'FontSize', fontSize);
title('Tcal group 2', 'FontSize', fontSize);

%% splitting in different group according to values in Tcal
%fucntion for the division of the DHW_mean_E vectors into groups according
%to the values of Tcal. in this case we have 2 groups (data obtained from
%the previous part of the analysis), so we apply the function to group1 and
%group2.function: "Tcal_splitting"

[DHW_split_group1]=Tcal_splitting(Tcal_vector_1,DHW_mean_group1_E);
[DHW_split_group2]=Tcal_splitting(Tcal_vector_2,DHW_mean_group2_E);

%% iterative merging process
[DHW_split_it11_group1,it_DHW_split_lastnew_group1]=iterative_merging_process(DH
W_split_group1);
[DHW_split_it11_group2,it_DHW_split_lastnew_group2]=iterative_merging_process(DH
W_split_group2);

%% identification of peak, average and minimum heat use
position1=[0.13 0.1 0.34 0.35]; %only for graphical purpose
position2=[0.55 0.1 0.34 0.35];
```

```matlab
[DHW_E_min_group1,DHW_E_max_group1]=min_med_peak_v2(it_DHW_split_lastnew_group1,
DHW_mean_group1,position1);
[DHW_E_min_group2,DHW_E_max_group2]=min_med_peak_v2(it_DHW_split_lastnew_group2,
DHW_mean_group2,position2);


%completed the analysis for the summer profile of DHW consumption.

%% creation of the standard profile for summer

[DHW_STD_PROFILE_group1,index_min_group1,index_med_group1,index_max_group1]=std_
profile_v1(DHW_mean_group1,DHW_E_min_group1,DHW_E_max_group1,1);
[DHW_STD_PROFILE_group2,index_min_group2,index_med_group2,index_max_group2]=std_
profile_v1(DHW_mean_group2,DHW_E_min_group2,DHW_E_max_group2,2);

%% creation of winter DHW standard profile

%mean values of DHW consumption during winter. Keep the division of the days of
the week done in
%this analysis also for winter season.

%creation of group1 and group2 for the whole year using DHW consumption
%from the previous analysis

%SPLITTING DHW AND SH DEMAND INTO DAYS OF THE WEEK
E_SH_Y_MON=[];
E_SH_Y_TUE=[];
E_SH_Y_WED=[];
E_SH_Y_THU=[];
E_SH_Y_FRI=[];
E_SH_Y_SAT=[];
E_SH_Y_SUN=[];
E_DHW_Y_MON=[];
E_DHW_Y_TUE=[];
E_DHW_Y_WED=[];
E_DHW_Y_THU=[];
E_DHW_Y_FRI=[];
E_DHW_Y_SAT=[];
E_DHW_Y_SUN=[];

limit_2=length(E_SH_balanced_year_corr);
E_DHW_year_corr_tra=E_DHW_year_corr';

%TUE
start=1;
finish=start+23;
for k=1:limit_2
    if finish<=limit_2
        vector_E_SH=E_SH_balanced_year_corr(start:finish);
        vector_E_DHW=E_DHW_year_corr_tra(start:finish);

        E_SH_Y_TUE=[E_SH_Y_TUE,vector_E_SH];
        E_DHW_Y_TUE=[E_DHW_Y_TUE,vector_E_DHW];

        start=start+168;
        finish=finish+168;
    end
end

%WED
start=25;
```

```matlab
finish=start+23;
for k=1:limit_2
    if finish<=limit_2
        vector_E_SH=E_SH_balanced_year_corr(start:finish);
        vector_E_DHW=E_DHW_year_corr_tra(start:finish);

        E_SH_Y_WED=[E_SH_Y_WED,vector_E_SH];
        E_DHW_Y_WED=[E_DHW_Y_WED,vector_E_DHW];

        start=start+168;
        finish=finish+168;
    end
end

%THU
start=49;
finish=start+23;
for k=1:limit_2
    if finish<=limit_2
        vector_E_SH=E_SH_balanced_year_corr(start:finish);
        vector_E_DHW=E_DHW_year_corr_tra(start:finish);

        E_SH_Y_THU=[E_SH_Y_THU,vector_E_SH];
        E_DHW_Y_THU=[E_DHW_Y_THU,vector_E_DHW];

        start=start+168;
        finish=finish+168;
    end
end

%FRI
start=73;
finish=start+23;
for k=1:limit_2
    if finish<=limit_2
        vector_E_SH=E_SH_balanced_year_corr(start:finish);
        vector_E_DHW=E_DHW_year_corr_tra(start:finish);

        E_SH_Y_FRI=[E_SH_Y_FRI,vector_E_SH];
        E_DHW_Y_FRI=[E_DHW_Y_FRI,vector_E_DHW];

        start=start+168;
        finish=finish+168;
    end
end

%SAT
start=97;
finish=start+23;
for k=1:limit_2
    if finish<=limit_2
        vector_E_SH=E_SH_balanced_year_corr(start:finish);
        vector_E_DHW=E_DHW_year_corr_tra(start:finish);

        E_SH_Y_SAT=[E_SH_Y_SAT,vector_E_SH];
        E_DHW_Y_SAT=[E_DHW_Y_SAT,vector_E_DHW];

        start=start+168;
        finish=finish+168;
    end
```

```matlab
    end

%SUN
start=121;
finish=start+23;
for k=1:limit_2
    if finish<=limit_2
        vector_E_SH=E_SH_balanced_year_corr(start:finish);
        vector_E_DHW=E_DHW_year_corr_tra(start:finish);

        E_SH_Y_SUN=[E_SH_Y_SUN,vector_E_SH];
        E_DHW_Y_SUN=[E_DHW_Y_SUN,vector_E_DHW];

        start=start+168;
        finish=finish+168;
    end
end

%MON
start=145;
finish=start+23;
for k=1:limit_2
    if finish<=limit_2
        vector_E_SH=E_SH_balanced_year_corr(start:finish);
        vector_E_DHW=E_DHW_year_corr_tra(start:finish);

        E_SH_Y_MON=[E_SH_Y_MON,vector_E_SH];
        E_DHW_Y_MON=[E_DHW_Y_MON,vector_E_DHW];

        start=start+168;
        finish=finish+168;
    end
end

%% MEAN VALUE FOR STAD PROFILE OF GROUP 1 and GROUP 2
mean_std_profile_group1=mean(DHW_STD_PROFILE_group1);
mean_std_profile_group2=mean(DHW_STD_PROFILE_group2);
%% SELECT WINTER SEASON
%for the winter season the period considered is from week 1 to week 21 and
%from week 37 to week 52.

E_SH_Y_MON_win=[];
E_SH_Y_TUE_win=[];
E_SH_Y_WED_win=[];
E_SH_Y_THU_win=[];
E_SH_Y_FRI_win=[];
E_SH_Y_SAT_win=[];
E_SH_Y_SUN_win=[];
E_DHW_Y_MON_win=[];
E_DHW_Y_TUE_win=[];
E_DHW_Y_WED_win=[];
E_DHW_Y_THU_win=[];
E_DHW_Y_FRI_win=[];
E_DHW_Y_SAT_win=[];
E_DHW_Y_SUN_win=[];

%TUE
for k=1:21
    control_day_SH=E_SH_Y_TUE(:,k);
    control_day_DHW=E_DHW_Y_TUE(:,k);
```

```matlab
    E_SH_Y_TUE_win=[E_SH_Y_TUE_win,control_day_SH];
    E_DHW_Y_TUE_win=[E_DHW_Y_TUE_win,control_day_DHW];
end
for k=37:52
    control_day_SH=E_SH_Y_TUE(:,k);
    control_day_DHW=E_DHW_Y_TUE(:,k);
    E_SH_Y_TUE_win=[E_SH_Y_TUE_win,control_day_SH];
    E_DHW_Y_TUE_win=[E_DHW_Y_TUE_win,control_day_DHW];
end


%WED
for k=1:21
    control_day_SH=E_SH_Y_WED(:,k);
    control_day_DHW=E_DHW_Y_WED(:,k);
    E_SH_Y_WED_win=[E_SH_Y_WED_win,control_day_SH];
    E_DHW_Y_WED_win=[E_DHW_Y_WED_win,control_day_DHW];
end
for k=37:52
    control_day_SH=E_SH_Y_WED(:,k);
    control_day_DHW=E_DHW_Y_WED(:,k);
    E_SH_Y_WED_win=[E_SH_Y_WED_win,control_day_SH];
    E_DHW_Y_WED_win=[E_DHW_Y_WED_win,control_day_DHW];
end


%THU
for k=1:21
    control_day_SH=E_SH_Y_THU(:,k);
    control_day_DHW=E_DHW_Y_THU(:,k);
    E_SH_Y_THU_win=[E_SH_Y_THU_win,control_day_SH];
    E_DHW_Y_THU_win=[E_DHW_Y_THU_win,control_day_DHW];
end
for k=37:52
    control_day_SH=E_SH_Y_THU(:,k);
    control_day_DHW=E_DHW_Y_THU(:,k);
    E_SH_Y_THU_win=[E_SH_Y_THU_win,control_day_SH];
    E_DHW_Y_THU_win=[E_DHW_Y_THU_win,control_day_DHW];
end


%FRI
for k=1:21
    control_day_SH=E_SH_Y_FRI(:,k);
    control_day_DHW=E_DHW_Y_FRI(:,k);
    E_SH_Y_FRI_win=[E_SH_Y_FRI_win,control_day_SH];
    E_DHW_Y_FRI_win=[E_DHW_Y_FRI_win,control_day_DHW];
end
for k=37:52
    control_day_SH=E_SH_Y_FRI(:,k);
    control_day_DHW=E_DHW_Y_FRI(:,k);
    E_SH_Y_FRI_win=[E_SH_Y_FRI_win,control_day_SH];
    E_DHW_Y_FRI_win=[E_DHW_Y_FRI_win,control_day_DHW];
end


%SAT
for k=1:21
    control_day_SH=E_SH_Y_SAT(:,k);
    control_day_DHW=E_DHW_Y_SAT(:,k);
    E_SH_Y_SAT_win=[E_SH_Y_SAT_win,control_day_SH];
    E_DHW_Y_SAT_win=[E_DHW_Y_SAT_win,control_day_DHW];
end
for k=37:52
    control_day_SH=E_SH_Y_SAT(:,k);
```

```matlab
        control_day_DHW=E_DHW_Y_SAT(:,k);
        E_SH_Y_SAT_win=[E_SH_Y_SAT_win,control_day_SH];
        E_DHW_Y_SAT_win=[E_DHW_Y_SAT_win,control_day_DHW];
end


%SUN
for k=1:21
        control_day_SH=E_SH_Y_SUN(:,k);
        control_day_DHW=E_DHW_Y_SUN(:,k);
        E_SH_Y_SUN_win=[E_SH_Y_SUN_win,control_day_SH];
        E_DHW_Y_SUN_win=[E_DHW_Y_SUN_win,control_day_DHW];
end
for k=37:52
        control_day_SH=E_SH_Y_SUN(:,k);
        control_day_DHW=E_DHW_Y_SUN(:,k);
        E_SH_Y_SUN_win=[E_SH_Y_SUN_win,control_day_SH];
        E_DHW_Y_SUN_win=[E_DHW_Y_SUN_win,control_day_DHW];
end


%MON
for k=1:21
        control_day_SH=E_SH_Y_MON(:,k);
        control_day_DHW=E_DHW_Y_MON(:,k);
        E_SH_Y_MON_win=[E_SH_Y_MON_win,control_day_SH];
        E_DHW_Y_MON_win=[E_DHW_Y_MON_win,control_day_DHW];
end
for k=37:52
        control_day_SH=E_SH_Y_MON(:,k);
        control_day_DHW=E_DHW_Y_MON(:,k);
        E_SH_Y_MON_win=[E_SH_Y_MON_win,control_day_SH];
        E_DHW_Y_MON_win=[E_DHW_Y_MON_win,control_day_DHW];
end


%% MEAN VALUES IN WINTER SEASON
mean_MON_DHW_win=mean(E_DHW_Y_MON_win(:));
mean_MON_SH_win=mean(E_SH_Y_MON_win(:));
mean_TUE_DHW_win=mean(E_DHW_Y_TUE_win(:));
mean_TUE_SH_win=mean(E_SH_Y_TUE_win(:));
mean_WED_DHW_win=mean(E_DHW_Y_WED_win(:));
mean_WED_SH_win=mean(E_SH_Y_WED_win(:));
mean_THU_DHW_win=mean(E_DHW_Y_THU_win(:));
mean_THU_SH_win=mean(E_SH_Y_THU_win(:));
mean_FRI_DHW_win=mean(E_DHW_Y_FRI_win(:));
mean_FRI_SH_win=mean(E_SH_Y_FRI_win(:));
mean_SAT_DHW_win=mean(E_DHW_Y_SAT_win(:));
mean_SAT_SH_win=mean(E_SH_Y_SAT_win(:));
mean_SUN_DHW_win=mean(E_DHW_Y_SUN_win(:));
mean_SUN_SH_win=mean(E_SH_Y_SUN_win(:));


%% DIFFERENCE IN % BETWEEN MEAN VALUES IN SUMMER AND IN WINTER
diff_MON=(mean_MON_DHW_win-mean_std_profile_group2)/mean_MON_DHW_win;
diff_TUE=(mean_TUE_DHW_win-mean_std_profile_group2)/mean_TUE_DHW_win;
diff_WED=(mean_WED_DHW_win-mean_std_profile_group2)/mean_WED_DHW_win;
diff_THU=(mean_THU_DHW_win-mean_std_profile_group2)/mean_THU_DHW_win;
diff_FRI=(mean_FRI_DHW_win-mean_std_profile_group1)/mean_FRI_DHW_win;
diff_SAT=(mean_SAT_DHW_win-mean_std_profile_group2)/mean_SAT_DHW_win;
diff_SUN=(mean_SUN_DHW_win-mean_std_profile_group2)/mean_SUN_DHW_win;


%% CREATION OF STANDARD PROFILE FOR WINTER

%determination of mean per cent difference between values in winter and in
```

```matlab
%summer in group1 and group2
mean_diff_perc_group2=(diff_MON+diff_TUE+diff_WED+diff_THU+diff_SAT+diff_SUN)/6;
mean_diff_perc_group1=diff_FRI;


%creation of new standard profile for winter: hourly values of the standard
%profile are increased by mean per cent difference between summer and
%winter of the corresponding group
vector_diff_group1=DHW_STD_PROFILE_group1.*mean_diff_perc_group1;
vector_diff_group2=DHW_STD_PROFILE_group2.*mean_diff_perc_group2;


DHW_WIN_STD_PROFILE_group1=DHW_STD_PROFILE_group1+vector_diff_group1;
DHW_WIN_STD_PROFILE_group2=DHW_STD_PROFILE_group2+vector_diff_group2;


DHW_E_min_WIN_group1=DHW_E_min_group1+mean_diff_perc_group1*DHW_E_min_group1;
DHW_E_max_WIN_group1=DHW_E_max_group1+mean_diff_perc_group1*DHW_E_max_group1;


DHW_E_min_WIN_group2=DHW_E_min_group2+mean_diff_perc_group2*DHW_E_min_group2;
DHW_E_max_WIN_group2=DHW_E_max_group2+mean_diff_perc_group2*DHW_E_max_group2;

%% PLOT NEW WINTER STANDARD PROFILE

plot_winter_profile(index_min_group1,index_med_group1,index_max_group1,DHW_WIN_S
TD_PROFILE_group1,DHW_E_min_WIN_group1,DHW_E_max_WIN_group1,3)
plot_winter_profile(index_min_group2,index_med_group2,index_max_group2,DHW_WIN_S
TD_PROFILE_group2,DHW_E_min_WIN_group2,DHW_E_max_WIN_group2,4)

%% FLATTENING WINTER DHW PROFILE USING STANDARD WINTER PROFILE AS A GUIDE

%create a matrix full of standard profiles in every colomn
MATRIX_STANDARD_WIN_group1=[];
for k=1:37

MATRIX_STANDARD_WIN_group1=[MATRIX_STANDARD_WIN_group1,DHW_WIN_STD_PROFILE_group
1'];
end

MATRIX_STANDARD_WIN_group2=[];
for k=1:37

MATRIX_STANDARD_WIN_group2=[MATRIX_STANDARD_WIN_group2,DHW_WIN_STD_PROFILE_group
2'];
end

%matrix with difference in % between value of DHW consumption in winter
%obtained from splitting model and standard porfile created for winter
%season.
DIFF_MON_WIN=(E_DHW_Y_MON_win-MATRIX_STANDARD_WIN_group2)./E_DHW_Y_MON_win;
DIFF_TUE_WIN=(E_DHW_Y_TUE_win-MATRIX_STANDARD_WIN_group2)./E_DHW_Y_TUE_win;
DIFF_WED_WIN=(E_DHW_Y_WED_win-MATRIX_STANDARD_WIN_group2)./E_DHW_Y_WED_win;
DIFF_THU_WIN=(E_DHW_Y_THU_win-MATRIX_STANDARD_WIN_group2)./E_DHW_Y_THU_win;
DIFF_FRI_WIN=(E_DHW_Y_FRI_win-MATRIX_STANDARD_WIN_group1)./E_DHW_Y_FRI_win;
DIFF_SAT_WIN=(E_DHW_Y_SAT_win-MATRIX_STANDARD_WIN_group2)./E_DHW_Y_SAT_win;
DIFF_SUN_WIN=(E_DHW_Y_SUN_win-MATRIX_STANDARD_WIN_group2)./E_DHW_Y_SUN_win;


%check what are the highest % differences
MAX_DIFF_MON_ABS=max(max(DIFF_MON_WIN));
MAX_DIFF_TUE_ABS=max(max(DIFF_TUE_WIN));
MAX_DIFF_WED_ABS=max(max(DIFF_WED_WIN));
MAX_DIFF_THU_ABS=max(max(DIFF_THU_WIN));
MAX_DIFF_FRI_ABS=max(max(DIFF_FRI_WIN));
```

```matlab
MAX_DIFF_SAT_ABS=max(max(DIFF_SAT_WIN));
MAX_DIFF_SUN_ABS=max(max(DIFF_SUN_WIN));


%check the mean % differences between winter profile from splitting and
%standard winter profile

MEAN_DIFF_MON=mean(mean(DIFF_MON_WIN));
MEAN_DIFF_TUE=mean(mean(DIFF_TUE_WIN));
MEAN_DIFF_WED=mean(mean(DIFF_WED_WIN));
MEAN_DIFF_THU=mean(mean(DIFF_THU_WIN));
MEAN_DIFF_FRI=mean(mean(DIFF_FRI_WIN));
MEAN_DIFF_SAT=mean(mean(DIFF_SAT_WIN));
MEAN_DIFF_SUN=mean(mean(DIFF_SUN_WIN));
%% propose a way for flattening DHW profile in winter
tolerance=0.3;
decrease=0.25;
[NEW_MON_Y_DHW_WIN,NEW_MON_Y_SH_WIN]=NEW_PROFILE(DIFF_MON_WIN,E_DHW_Y_MON_win,E_
SH_Y_MON_win,tolerance,decrease,1);
[NEW_TUE_Y_DHW_WIN,NEW_TUE_Y_SH_WIN]=NEW_PROFILE(DIFF_TUE_WIN,E_DHW_Y_TUE_win,E_
SH_Y_TUE_win,tolerance,decrease,2);
[NEW_WED_Y_DHW_WIN,NEW_WED_Y_SH_WIN]=NEW_PROFILE(DIFF_WED_WIN,E_DHW_Y_WED_win,E_
SH_Y_WED_win,tolerance,decrease,3);
[NEW_THU_Y_DHW_WIN,NEW_THU_Y_SH_WIN]=NEW_PROFILE(DIFF_THU_WIN,E_DHW_Y_THU_win,E_
SH_Y_THU_win,tolerance,decrease,4);
[NEW_FRI_Y_DHW_WIN,NEW_FRI_Y_SH_WIN]=NEW_PROFILE(DIFF_FRI_WIN,E_DHW_Y_FRI_win,E_
SH_Y_FRI_win,tolerance,decrease,5);
[NEW_SAT_Y_DHW_WIN,NEW_SAT_Y_SH_WIN]=NEW_PROFILE(DIFF_SAT_WIN,E_DHW_Y_SAT_win,E_
SH_Y_SAT_win,tolerance,decrease,6);
[NEW_SUN_Y_DHW_WIN,NEW_SUN_Y_SH_WIN]=NEW_PROFILE(DIFF_SUN_WIN,E_DHW_Y_SUN_win,E_
SH_Y_SUN_win,tolerance,decrease,7);


%% reconstruction of yearly demand with modified values
[E_SH_Y_MON_NEW,E_DHW_Y_MON_NEW]=NEW_PROFILE_YEARLY(NEW_MON_Y_SH_WIN,E_SH_Y_MON,
NEW_MON_Y_DHW_WIN,E_DHW_Y_MON);
[E_SH_Y_TUE_NEW,E_DHW_Y_TUE_NEW]=NEW_PROFILE_YEARLY(NEW_TUE_Y_SH_WIN,E_SH_Y_TUE,
NEW_TUE_Y_DHW_WIN,E_DHW_Y_TUE);
[E_SH_Y_WED_NEW,E_DHW_Y_WED_NEW]=NEW_PROFILE_YEARLY(NEW_WED_Y_SH_WIN,E_SH_Y_WED,
NEW_WED_Y_DHW_WIN,E_DHW_Y_WED);
[E_SH_Y_THU_NEW,E_DHW_Y_THU_NEW]=NEW_PROFILE_YEARLY(NEW_THU_Y_SH_WIN,E_SH_Y_THU,
NEW_THU_Y_DHW_WIN,E_DHW_Y_THU);
[E_SH_Y_FRI_NEW,E_DHW_Y_FRI_NEW]=NEW_PROFILE_YEARLY(NEW_FRI_Y_SH_WIN,E_SH_Y_FRI,
NEW_FRI_Y_DHW_WIN,E_DHW_Y_FRI);
[E_SH_Y_SAT_NEW,E_DHW_Y_SAT_NEW]=NEW_PROFILE_YEARLY(NEW_SAT_Y_SH_WIN,E_SH_Y_SAT,
NEW_SAT_Y_DHW_WIN,E_DHW_Y_SAT);
[E_SH_Y_SUN_NEW,E_DHW_Y_SUN_NEW]=NEW_PROFILE_YEARLY(NEW_SUN_Y_SH_WIN,E_SH_Y_SUN,
NEW_SUN_Y_DHW_WIN,E_DHW_Y_SUN);


%% Creation of whole year demand (put together all days of the week)
%this year starts from tuesday)

E_DHW_Y_NEW=[];
E_SH_Y_NEW=[];

for k=1:52
    control_vector=E_SH_Y_TUE_NEW(:,k);
    E_SH_Y_NEW=[E_SH_Y_NEW,control_vector];
    control_vector=E_SH_Y_WED_NEW(:,k);
    E_SH_Y_NEW=[E_SH_Y_NEW,control_vector];
    control_vector=E_SH_Y_THU_NEW(:,k);
    E_SH_Y_NEW=[E_SH_Y_NEW,control_vector];
    control_vector=E_SH_Y_FRI_NEW(:,k);
```

```matlab
        E_SH_Y_NEW=[E_SH_Y_NEW,control_vector];
        control_vector=E_SH_Y_SAT_NEW(:,k);
        E_SH_Y_NEW=[E_SH_Y_NEW,control_vector];
        control_vector=E_SH_Y_SUN_NEW(:,k);
        E_SH_Y_NEW=[E_SH_Y_NEW,control_vector];
        control_vector=E_SH_Y_MON_NEW(:,k);
        E_SH_Y_NEW=[E_SH_Y_NEW,control_vector];
end

for k=1:52
        control_vector=E_DHW_Y_TUE_NEW(:,k);
        E_DHW_Y_NEW=[E_DHW_Y_NEW,control_vector];
        control_vector=E_DHW_Y_WED_NEW(:,k);
        E_DHW_Y_NEW=[E_DHW_Y_NEW,control_vector];
        control_vector=E_DHW_Y_THU_NEW(:,k);
        E_DHW_Y_NEW=[E_DHW_Y_NEW,control_vector];
        control_vector=E_DHW_Y_FRI_NEW(:,k);
        E_DHW_Y_NEW=[E_DHW_Y_NEW,control_vector];
        control_vector=E_DHW_Y_SAT_NEW(:,k);
        E_DHW_Y_NEW=[E_DHW_Y_NEW,control_vector];
        control_vector=E_DHW_Y_SUN_NEW(:,k);
        E_DHW_Y_NEW=[E_DHW_Y_NEW,control_vector];
        control_vector=E_DHW_Y_MON_NEW(:,k);
        E_DHW_Y_NEW=[E_DHW_Y_NEW,control_vector];
end

E_DHW_Y_NEW_line=E_DHW_Y_NEW(:)';
E_SH_Y_NEW_line=E_SH_Y_NEW(:)';

%% plot representaiton

%graphical variables
limit_4=length(E_DHW_Y_NEW_line);
limit_5=length(E_SH_balanced_year_corr);
time_1=1:limit_4;
time_2=1:limit_5;

fontSize=13;

figure (32)
subplot(3,3,8)
plot(time_2,E_DHW_year_corr,'color','#D95319')
hold on
plot(time_2,E_SH_balanced_year_corr,'b')
plot(time_1,E_DHW_Y_NEW_line,'color','#EDB120')
plot(time_1,E_SH_Y_NEW_line,'c')
hold off
grid on
xlabel('TIME', 'FontSize', fontSize);
ylabel('DEMAND', 'FontSize', fontSize);
title('SPLITTED DEMAND VS TIME', 'FontSize', fontSize);
legend('DHW from splitting','SH from splitting','DHW modified','SH
MODIFIED','location','NorthEast')

%% last plot Demand vs T out
%manca solo l'ultimo grafico: le 4 domande ottenute ma vs la temperatura

%graphical parameters
T_NEW=T(1:length(E_SH_Y_NEW_line));
```

```matlab
figure (32)
subplot(3,3,9)
plot(T,E_DHW_year_corr,'.','color','#D95319')
hold on
plot(T,E_SH_balanced_year_corr,'.b')
plot(T_NEW,E_DHW_Y_NEW_line,'.','color','#EDB120')
plot(T_NEW,E_SH_Y_NEW_line,'.c')
hold off
grid on
xlabel('T out', 'FontSize', fontSize);
ylabel('DEMAND', 'FontSize', fontSize);
title('SPLITTED DEMAND VS T out', 'FontSize', fontSize);
legend('DHW from splitting','SH from splitting','DHW modified','SH
MODIFIED','location','NorthEast')
```

A2: Partial database (Verona case study)

"CI_v2.m"

```matlab
%function per il CI
function [y_CI_up,y_CI_down]=CI_v2(coeff_DEF_SH_0AM,T_SH_0AM,Demand_SH_0AM,i,k)

%valutazione punti con il modello lineare
SH_value_0AM=polyval(coeff_DEF_SH_0AM,T_SH_0AM);
n_lib_deg=length(SH_value_0AM)-2;

%caclolo parametri necessari
Se_SH_0AM=sqrt((n_lib_deg+1)*((std(Demand_SH_0AM)^2)-
(coeff_DEF_SH_0AM(1)^2)*std(T_SH_0AM))/(n_lib_deg));
%usare il valore di porbabilità cumulata (cumulative density function)
%nella t-student per includere tutti i valori all'interno della probabilità
%scelt, il 95% solitamente basta. Avendo un alto numero di gradi di libertà
%coincide con la distribuzione gaussiana.
t_stud_value=tcdf(0.95,n_lib_deg);
prod_factor=sqrt(1+(1/length(SH_value_0AM))+(((T_SH_0AM-
mean(T_SH_0AM)).^2)/(length(SH_value_0AM)*(std(T_SH_0AM)^2))));

%calcolo y degli intervalli di confidenza
y_CI_up=SH_value_0AM+Se_SH_0AM.*t_stud_value.*prod_factor;
y_CI_down=SH_value_0AM-Se_SH_0AM.*t_stud_value.*prod_factor;

fontSize=18;
axis_limit=[-5 20 0 60];

%grafico
figure(i)
subplot(3,3,k)
% pos=[0.65 0.2 0.3 0.3];
% subplot('Position',pos)
plot(T_SH_0AM,y_CI_up,'g')
grid on
hold on
plot(T_SH_0AM,y_CI_down,'g')
plot(T_SH_0AM,Demand_SH_0AM,'.b')
plot(T_SH_0AM,SH_value_0AM,'r')
hold off
axis(axis_limit)
xlabel('T', 'FontSize', fontSize);
ylabel('ENERGY USE', 'FontSize', fontSize);
title('CONFIDENCE INTERVAL AND FITTING', 'FontSize', fontSize);
```

```matlab
legend('UPPER CI 95%','LOWER CI 95%','experimental points','model')

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

end
```

"CI_v3.m"

```matlab
%function per il CI
function [y_CI_up,y_CI_down]=CI_v3(coeff_DEF_SH_0AM,T_SH_0AM,Demand_SH_0AM,i,k)

%valutazione punti con il modello lineare
SH_value_0AM=polyval(coeff_DEF_SH_0AM,T_SH_0AM);
n_lib_deg=length(SH_value_0AM)-2;

%caclolo parametri necessari
Se_SH_0AM=sqrt((n_lib_deg+1)*((std(Demand_SH_0AM)^2)-
(coeff_DEF_SH_0AM(1)^2)*std(T_SH_0AM))/(n_lib_deg));
%usare il valore di porbabilità cumulata (cumulative density function)
%nella t-student per includere tutti i valori all'interno della probabilità
%scelt, il 95% solitamente basta. Avendo un alto numero di gradi di libertà
%coincide con la distribuzione gaussiana.
t_stud_value=tcdf(0.95,n_lib_deg);
prod_factor=sqrt(1+(1/length(SH_value_0AM))+(((T_SH_0AM-
mean(T_SH_0AM)).^2)/(length(SH_value_0AM)*(std(T_SH_0AM)^2))));

%calcolo y degli intervalli di confidenza
y_CI_up=SH_value_0AM+Se_SH_0AM.*t_stud_value.*prod_factor;
y_CI_down=SH_value_0AM-Se_SH_0AM.*t_stud_value.*prod_factor;

fontSize=18;

axis_limit=[-5 20 0 60];

%grafico
figure(i)
subplot(3,3,k)
% pos=[0.65 0.2 0.3 0.3];
% subplot('Position',pos)
plot(T_SH_0AM,y_CI_up,'g')
grid on
hold on
plot(T_SH_0AM,y_CI_down,'g')
plot(T_SH_0AM,Demand_SH_0AM,'.b')
plot(T_SH_0AM,SH_value_0AM,'r')
hold off
axis(axis_limit)
xlabel('T', 'FontSize', fontSize);
ylabel('ENERGY USE', 'FontSize', fontSize);
title('CONFIDENCE INTERVAL AND FITTING', 'FontSize', fontSize);
legend('UPPER CI 95%','LOWER CI 95%','experimental points','model')

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

end
```

"DMT_calc.m"

```matlab
%function for DMT calculation

function [DMT,T]=DMT_calc(HOURS_D_NaN)
DMT=[];
scr=1;

T=HOURS_D_NaN;       %vector T contains T out with removal of entire days with at
least one NaN value
limit=length(T);

for j=1:364
    if (scr+23)<= limit

dailyMT=(T(scr,1)+T(scr+1,1)+T(scr+2,1)+T(scr+3,1)+T(scr+4,1)+T(scr+5,1)+T(scr+6
,1)+T(scr+7,1)+T(scr+8,1)+T(scr+9,1)+
T(scr+10,1)+T(scr+11,1)+T(scr+12,1)+T(scr+13,1)+T(scr+14,1)+T(scr+15,1)+T(scr+16
,1)+T(scr+17,1)+T(scr+18,1)+T(scr+19,1)+T(scr+20,1)+T(scr+21,1)+T(scr+22,1)+T(sc
r+23,1))/24;
    DMT(j,1)=dailyMT;

    scr=scr+24;
    end
end

end
```

"ESC_School2.m"

```matlab
%script for the analysis of incomplete dataset (less than 8760 values). inm
%particular this is made for buildings where only heating season
%consumption is known (from november to march).

%based on Verona DHN dataset

%close:
close all
clear all
clc

%data reading
ENERGY=xlsread('dataset_School2.xlsx','R2:R8760');
%importare dati non doppi, processare in excel prima
HOURS=xlsread('dataset_School2.xlsx','Q2:Q8760');
%vettore che contiene dati della temperatura
CONTROL_POWER=xlsread('dataset_School2.xlsx','G2:G8760');
%for controlling where is NaN in the original dataset

%% processing dataset: delete only NaN values
%delete every NaN element from the ENERGY vestor. in the original dataset
%there are some empty raws: probably meter was off or there were some
%errors in the data reading.

%do the same operation in the vector HOURS to be suyre that the Tout
%corresponds to the value of demand in the considered hour.
```

```matlab
%
[ENERGY_N_NaN,HOURS_N_NaN,ENERGY_D_NaN,HOURS_D_NaN]=NaN_removal(ENERGY,HOURS,CON
TROL_POWER);

%starting: insert number of the first day in the database:
%(monday=1; tuesday=2; wednesday=3... sunday=7);
starting=1;


[ENERGY_N_NaN,HOURS_N_NaN,ENERGY_D_NaN_test,HOURS_D_NaN_test,NUMBER_DAY,ENERGY_D
_NaN_ORDER,WEEKEND_N_NaN_E_COL,WEEKDAYS_N_NaN_E_COL,WEEKEND_N_NaN_T_COL,WEEKDAYS
_N_NaN_T_COL]=NaN_removal_v2(ENERGY,HOURS,CONTROL_POWER,starting);

%% TEST CONSIDERING ONLY WEEKDAYS (NO WEEKEND)
ENERGY_D_NaN=WEEKDAYS_N_NaN_E_COL;
HOURS_D_NaN=WEEKDAYS_N_NaN_T_COL;


% ENERGY_D_NaN=ENERGY_D_NaN_test;
% HOURS_D_NaN=HOURS_D_NaN_test;

%it is much better to consider only weekdays, most conistent results

%% HOUR SPLITTING (energy)
%(possible only with ENERGY_D_NaN
[ENERGY_0AM]=HOUR_split(ENERGY_D_NaN,1);
[ENERGY_1AM]=HOUR_split(ENERGY_D_NaN,2);
[ENERGY_2AM]=HOUR_split(ENERGY_D_NaN,3);
[ENERGY_3AM]=HOUR_split(ENERGY_D_NaN,4);
[ENERGY_4AM]=HOUR_split(ENERGY_D_NaN,5);
[ENERGY_5AM]=HOUR_split(ENERGY_D_NaN,6);
[ENERGY_6AM]=HOUR_split(ENERGY_D_NaN,7);
[ENERGY_7AM]=HOUR_split(ENERGY_D_NaN,8);
[ENERGY_8AM]=HOUR_split(ENERGY_D_NaN,9);
[ENERGY_9AM]=HOUR_split(ENERGY_D_NaN,10);
[ENERGY_10AM]=HOUR_split(ENERGY_D_NaN,11);
[ENERGY_11AM]=HOUR_split(ENERGY_D_NaN,12);
[ENERGY_12AM]=HOUR_split(ENERGY_D_NaN,13);
[ENERGY_13PM]=HOUR_split(ENERGY_D_NaN,14);
[ENERGY_14PM]=HOUR_split(ENERGY_D_NaN,15);
[ENERGY_15PM]=HOUR_split(ENERGY_D_NaN,16);
[ENERGY_16PM]=HOUR_split(ENERGY_D_NaN,17);
[ENERGY_17PM]=HOUR_split(ENERGY_D_NaN,18);
[ENERGY_18PM]=HOUR_split(ENERGY_D_NaN,19);
[ENERGY_19PM]=HOUR_split(ENERGY_D_NaN,20);
[ENERGY_20PM]=HOUR_split(ENERGY_D_NaN,21);
[ENERGY_21PM]=HOUR_split(ENERGY_D_NaN,22);
[ENERGY_22PM]=HOUR_split(ENERGY_D_NaN,23);
[ENERGY_23PM]=HOUR_split(ENERGY_D_NaN,24);


%% HOUR SPLITTING (temperature)
%(possible only with HOURS_D_NaN
[T_0AM]=HOUR_split(HOURS_D_NaN,1);
[T_1AM]=HOUR_split(HOURS_D_NaN,2);
[T_2AM]=HOUR_split(HOURS_D_NaN,3);
[T_3AM]=HOUR_split(HOURS_D_NaN,4);
[T_4AM]=HOUR_split(HOURS_D_NaN,5);
[T_5AM]=HOUR_split(HOURS_D_NaN,6);
[T_6AM]=HOUR_split(HOURS_D_NaN,7);
[T_7AM]=HOUR_split(HOURS_D_NaN,8);
[T_8AM]=HOUR_split(HOURS_D_NaN,9);
[T_9AM]=HOUR_split(HOURS_D_NaN,10);
[T_10AM]=HOUR_split(HOURS_D_NaN,11);
```

```matlab
[T_11AM]=HOUR_split(HOURS_D_NaN,12);
[T_12AM]=HOUR_split(HOURS_D_NaN,13);
[T_13PM]=HOUR_split(HOURS_D_NaN,14);
[T_14PM]=HOUR_split(HOURS_D_NaN,15);
[T_15PM]=HOUR_split(HOURS_D_NaN,16);
[T_16PM]=HOUR_split(HOURS_D_NaN,17);
[T_17PM]=HOUR_split(HOURS_D_NaN,18);
[T_18PM]=HOUR_split(HOURS_D_NaN,19);
[T_19PM]=HOUR_split(HOURS_D_NaN,20);
[T_20PM]=HOUR_split(HOURS_D_NaN,21);
[T_21PM]=HOUR_split(HOURS_D_NaN,22);
[T_22PM]=HOUR_split(HOURS_D_NaN,23);
[T_23PM]=HOUR_split(HOURS_D_NaN,24);

%% daily mean temperature
[DMT,T]=DMT_calc(HOURS_D_NaN); %vector T contains T out with removal of entire
days with at least one NaN value, same as HOURS_D_NaN

%% Check which points are below the most common pattern:
fit=polyfit(T,ENERGY_D_NaN,1);
eval=polyval(fit,T);
%ok if the line stay in the middle between 2 profiles, otherwise fine a
%line from the plot and insert coefficient by hand in the coefficient
%polynomial "fit"

%% plot DMT vs ENERGY_hour (0AM, 1AM...)
%creation of 1 figure per vector, usa as base for further considerations

fontSize = 13;
markerSize = 20;
axis_limit_1=[-5 20 0 max(ENERGY_D_NaN)];

figure (1)
subplot(6,4,1)
plot(DMT,ENERGY_0AM,'.')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 0AM', 'FontSize', fontSize);
title('ENERGY 0AM vs DMT', 'FontSize', fontSize);
hold on
plot(T,eval,'r');
hold off
axis(axis_limit_1)

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (1)
subplot(6,4,2)
plot(DMT,ENERGY_1AM,'.')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 1AM', 'FontSize', fontSize);
title('ENERGY 1AM vs DMT', 'FontSize', fontSize);
hold on
plot(T,eval,'r');
hold off
axis(axis_limit_1)

% METTERE A SCHERMO INTERO
```

221

```matlab
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (1)
subplot(6,4,3)
plot(DMT,ENERGY_2AM,'.')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 2AM', 'FontSize', fontSize);
title('ENERGY 2AM vs DMT', 'FontSize', fontSize);
hold on
plot(T,eval,'r');
hold off
axis(axis_limit_1)

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (1)
subplot(6,4,4)
plot(DMT,ENERGY_3AM,'.')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 3AM', 'FontSize', fontSize);
title('ENERGY 3AM vs DMT', 'FontSize', fontSize);
hold on
plot(T,eval,'r');
hold off
axis(axis_limit_1)

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (1)
subplot(6,4,5)
plot(DMT,ENERGY_4AM,'.')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 4 AM', 'FontSize', fontSize);
title('ENERGY 4 AM vs DMT', 'FontSize', fontSize);
hold on
plot(T,eval,'r');
hold off
axis(axis_limit_1)

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (1)
subplot(6,4,6)
plot(DMT,ENERGY_5AM,'.')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 5AM', 'FontSize', fontSize);
title('ENERGY 5AM vs DMT', 'FontSize', fontSize);
hold on
plot(T,eval,'r');
hold off
axis(axis_limit_1)

% METTERE A SCHERMO INTERO
```

```matlab
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (1)
subplot(6,4,7)
plot(DMT,ENERGY_6AM,'.')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 6AM', 'FontSize', fontSize);
title('ENERGY 6AM vs DMT', 'FontSize', fontSize);
hold on
plot(T,eval,'r');
hold off
axis(axis_limit_1)

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (1)
subplot(6,4,8)
plot(DMT,ENERGY_7AM,'.')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 7AM', 'FontSize', fontSize);
title('ENERGY 7AM vs DMT', 'FontSize', fontSize);
hold on
plot(T,eval,'r');
hold off
axis(axis_limit_1)

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (1)
subplot(6,4,9)
plot(DMT,ENERGY_8AM,'.')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 8AM', 'FontSize', fontSize);
title('ENERGY 8AM vs DMT', 'FontSize', fontSize);
hold on
plot(T,eval,'r');
hold off
axis(axis_limit_1)

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (1)
subplot(6,4,10)
plot(DMT,ENERGY_9AM,'.')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 9AM', 'FontSize', fontSize);
title('ENERGY 9AM vs DMT', 'FontSize', fontSize);
hold on
plot(T,eval,'r');
hold off
axis(axis_limit_1)

% METTERE A SCHERMO INTERO
```

```matlab
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (1)
subplot(6,4,11)
plot(DMT,ENERGY_10AM,'.')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 10AM', 'FontSize', fontSize);
title('ENERGY 10AM vs DMT', 'FontSize', fontSize);
hold on
plot(T,eval,'r');
hold off
axis(axis_limit_1)

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (1)
subplot(6,4,12)
plot(DMT,ENERGY_11AM,'.')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 11AM', 'FontSize', fontSize);
title('ENERGY 11AM vs DMT', 'FontSize', fontSize);
hold on
plot(T,eval,'r');
hold off
axis(axis_limit_1)

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (1)
subplot(6,4,13)
plot(DMT,ENERGY_12AM,'.')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 12AM', 'FontSize', fontSize);
title('ENERGY 12AM vs DMT', 'FontSize', fontSize);
hold on
plot(T,eval,'r');
hold off
axis(axis_limit_1)

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (1)
subplot(6,4,14)
plot(DMT,ENERGY_13PM,'.')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 13PM', 'FontSize', fontSize);
title('ENERGY 13PM vs DMT', 'FontSize', fontSize);
hold on
plot(T,eval,'r');
hold off
axis(axis_limit_1)

% METTERE A SCHERMO INTERO
```

```matlab
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (1)
subplot(6,4,15)
plot(DMT,ENERGY_14PM,'.')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 14PM', 'FontSize', fontSize);
title('ENERGY 14PM vs DMT', 'FontSize', fontSize);
hold on
plot(T,eval,'r');
hold off
axis(axis_limit_1)

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (1)
subplot(6,4,16)
plot(DMT,ENERGY_15PM,'.')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 15PM', 'FontSize', fontSize);
title('ENERGY 15PM vs DMT', 'FontSize', fontSize);
hold on
plot(T,eval,'r');
hold off
axis(axis_limit_1)

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (1)
subplot(6,4,17)
plot(DMT,ENERGY_16PM,'.')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 16PM', 'FontSize', fontSize);
title('ENERGY 16PM vs DMT', 'FontSize', fontSize);
hold on
plot(T,eval,'r');
hold off
axis(axis_limit_1)

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (1)
subplot(6,4,18)
plot(DMT,ENERGY_17PM,'.')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 17PM', 'FontSize', fontSize);
title('ENERGY 17PM vs DMT', 'FontSize', fontSize);
hold on
plot(T,eval,'r');
hold off
axis(axis_limit_1)

% METTERE A SCHERMO INTERO
```

```matlab
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (1)
subplot(6,4,19)
plot(DMT,ENERGY_18PM,'.')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 18PM', 'FontSize', fontSize);
title('ENERGY 18PM vs DMT', 'FontSize', fontSize);
hold on
plot(T,eval,'r');
hold off
axis(axis_limit_1)

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (1)
subplot(6,4,20)
plot(DMT,ENERGY_19PM,'.')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 19PM', 'FontSize', fontSize);
title('ENERGY 19PM vs DMT', 'FontSize', fontSize);
hold on
plot(T,eval,'r');
hold off
axis(axis_limit_1)

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (1)
subplot(6,4,21)
plot(DMT,ENERGY_20PM,'.')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 20PM', 'FontSize', fontSize);
title('ENERGY 20PM vs DMT', 'FontSize', fontSize);
hold on
plot(T,eval,'r');
hold off
axis(axis_limit_1)

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (1)
subplot(6,4,22)
plot(DMT,ENERGY_21PM,'.')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 21PM', 'FontSize', fontSize);
title('ENERGY 21PM vs DMT', 'FontSize', fontSize);
hold on
plot(T,eval,'r');
hold off
axis(axis_limit_1)

% METTERE A SCHERMO INTERO
```

```matlab
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (1)
subplot(6,4,23)
plot(DMT,ENERGY_22PM,'.')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 22PM', 'FontSize', fontSize);
title('ENERGY 22PM vs DMT', 'FontSize', fontSize);
hold on
plot(T,eval,'r');
hold off
axis(axis_limit_1)

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

figure (1)
subplot(6,4,24)
plot(DMT,ENERGY_23PM,'.')
grid on
xlabel('DMT', 'FontSize', fontSize);
ylabel('ENERGY 23PM', 'FontSize', fontSize);
title('ENERGY 23PM vs DMT', 'FontSize', fontSize);
hold on
plot(T,eval,'r');
hold off
axis(axis_limit_1)

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

%% calcolo di ESC totale, non solo quella divisa in ore

figure (27)
%grafico punti sperimentali
subplot(3,3,1)
plot(T,ENERGY_D_NaN,'.')
grid on
xlabel('T', 'FontSize', fontSize);
ylabel('ENERGY USE', 'FontSize', fontSize);
title('ENERGY USE vs T', 'FontSize', fontSize);
hold on
plot(T,eval,'r');
hold off

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

%plot of complete year using n NaN (removal only NaN, not entire days)
figure (28)
%grafico punti sperimentali
subplot(3,3,1)
plot(HOURS_N_NaN,ENERGY_N_NaN,'.')
grid on
xlabel('T', 'FontSize', fontSize);
ylabel('ENERGY USE', 'FontSize', fontSize);
title('ENERGY USE vs T', 'FontSize', fontSize);
hold on
plot(T,eval,'r');
```

```matlab
hold off

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

%% Profile separation
% Verona is different from Tartu: there is a milder climate, so it isn't
% necessary to keep SH on for the entire day. There are rules from comune
% di Verona with hours for the SH plant operation. thanks to hours
% splitting we can simply find 2 profiles:
%-SH on: creation of ESC and considerations done for Tartu
%-SH off: useful for DHW losses considerations (typical is at night time)

%find the corresponding group for each hour with an analysis from the plot.

%separation line is the interpolation line of the complete dataset. points
%aboce are SH on, points below when SH is off.

%splitting by hand, observing the plot
%SH on (%14 h of SH on, ok with Municipality of Verona's rules for
%2019-2020)
ENERGY_SH_ON=[ENERGY_4AM,ENERGY_5AM,ENERGY_6AM,ENERGY_9AM,ENERGY_10AM,ENERGY_11A
M,ENERGY_12AM,ENERGY_13PM,ENERGY_14PM,ENERGY_15PM,ENERGY_16PM,ENERGY_17PM,ENERGY
_18PM,ENERGY_19PM];

%SH off
ENERGY_SH_OFF=[ENERGY_0AM,ENERGY_1AM,ENERGY_2AM,ENERGY_3AM,ENERGY_7AM,ENERGY_8AM
,ENERGY_20PM,ENERGY_21PM,ENERGY_22PM,ENERGY_23PM];

%Automatic splitting
percentage=0.55;  %try different tolerance (between 50-85) to find the best
rapresentation

COMPLETE_ENERGY=[ENERGY_0AM,ENERGY_1AM,ENERGY_2AM,ENERGY_3AM,ENERGY_4AM,ENERGY_5
AM,ENERGY_6AM,ENERGY_7AM,ENERGY_8AM,ENERGY_9AM,ENERGY_10AM,ENERGY_11AM,ENERGY_12
AM,ENERGY_13PM,ENERGY_14PM,ENERGY_15PM,ENERGY_16PM,ENERGY_17PM,ENERGY_18PM,ENERG
Y_19PM,ENERGY_20PM,ENERGY_21PM,ENERGY_22PM,ENERGY_23PM];
COMPLETE_T=[T_0AM,T_1AM,T_2AM,T_3AM,T_4AM,T_5AM,T_6AM,T_7AM,T_8AM,T_9AM,T_10AM,T
_11AM,T_12AM,T_13PM,T_14PM,T_15PM,T_16PM,T_17PM,T_18PM,T_19PM,T_20PM,T_21PM,T_22
PM,T_23PM];

[ENERGY_SH_ON_A,ENERGY_SH_OFF_A,T_SH_ON_A,T_SH_OFF_A]=SH_ONOFF(COMPLETE_ENERGY,C
OMPLETE_T,percentage,fit);

%% FITTING SH ON PROFILE AND SH OFF PROFILE
COMPLETE_T_COL=COMPLETE_T(:);   %trasforming matrix into vector
COMPLETE_ENERGY_COL=COMPLETE_ENERGY(:);
ENERGY_SH_ON_A_COL=ENERGY_SH_ON_A(:);
ENERGY_SH_OFF_A_COL=ENERGY_SH_OFF_A(:);
T_SH_ON_A_COL=T_SH_ON_A(:);
T_SH_OFF_A_COL=T_SH_OFF_A(:);

fit_SH_ON=polyfit(T_SH_ON_A_COL,ENERGY_SH_ON_A_COL,1);
fit_SH_OFF=polyfit(T_SH_OFF_A_COL,ENERGY_SH_OFF_A_COL,1);

eval_SH_ON=polyval(fit_SH_ON,T_SH_ON_A_COL);
eval_SH_OFF=polyval(fit_SH_OFF,T_SH_OFF_A_COL);

axis_limit=[-5 20 0 60];
```

```matlab
figure(29)
subplot(3,3,1)
plot(T_SH_ON_A_COL,ENERGY_SH_ON_A_COL,'g.')
hold on
plot(T_SH_OFF_A_COL,ENERGY_SH_OFF_A_COL,'b.')
plot(T_SH_ON_A_COL,eval_SH_ON,'g')
plot(T_SH_OFF_A_COL,eval_SH_OFF,'b')
grid on
axis(axis_limit)
xlabel('T out', 'FontSize', fontSize);
ylabel('ENERGY 0AM', 'FontSize', fontSize);
title('ENERGY vs T out', 'FontSize', fontSize);
legend('Demand SH on','Demand SH off');


% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);


%% CI APPLICATION
[y_CI_up_on,y_CI_down_on]=CI_v2(fit_SH_ON,T_SH_ON_A_COL,ENERGY_SH_ON_A_COL,29,2)
;
[y_CI_up_off,y_CI_down_off]=CI_v3(fit_SH_OFF,T_SH_OFF_A_COL,ENERGY_SH_OFF_A_COL,
29,3);

%% ESC reconstruction
%consider last values, find a mean values between them and interpolate. it
%is necessary the vector that contains temperature of Verona in the whole
%year. we reconstruct the summer demand with constant values.

%create vector with Tout from the complete year

%for choosing CPT build frequence distribution of T out and cancel tails of
%the distribution (in particular higher tail). Use typical 95% of
%tolerance

LUNG_T_ON=length(T_SH_ON_A_COL);
LUNG_T_OFF=length(T_SH_OFF_A_COL);

[DIST_T_SH_ON_ASS,X_T_SH_ON]=hist(T_SH_ON_A_COL,24);
[DIST_T_SH_OFF_ASS,X_T_SH_OFF]=hist(T_SH_OFF_A_COL,24);
DIST_T_SH_ON=DIST_T_SH_ON_ASS/LUNG_T_ON;
DIST_T_SH_OFF=DIST_T_SH_OFF_ASS/LUNG_T_OFF;

%norm pdf mean and std from T dataset
MEAN_T_SH_ON=mean(T_SH_ON_A_COL);
STD_T_SH_ON=std(T_SH_ON_A_COL);
MIN_T_SH_ON=min(T_SH_ON_A_COL);
MAX_T_SH_ON=max(T_SH_ON_A_COL);
vector_norm=MIN_T_SH_ON:0.1:MAX_T_SH_ON;
NORM_T_SH_ON=normpdf(vector_norm,MEAN_T_SH_ON,STD_T_SH_ON);
NORM_CDF_T_SH_ON=normcdf(vector_norm,MEAN_T_SH_ON,STD_T_SH_ON);

%build vector of CDF for DIST_T_SH_ON

figure(29)
subplot(3,3,4)
plot(vector_norm,NORM_T_SH_ON,'b')
```

```matlab
grid on
hold on
% plot(X_T_SH_ON,DIST_T_SH_ON,'g')
bar(X_T_SH_ON,DIST_T_SH_ON,'g')
hold off
xlabel('T out', 'FontSize', fontSize);
ylabel('probability', 'FontSize', fontSize);
title('PDF T OUT', 'FontSize', fontSize);
legend('normal PDF','T PDF');

figure(29)
subplot(3,3,5)
plot(vector_norm,NORM_CDF_T_SH_ON)
hold on
%cdfplot(T_SH_ON_A_COL)
ecdf(T_SH_ON_A_COL,'Bounds','on')
grid on
xlabel('T out', 'FontSize', fontSize);
ylabel('probability', 'FontSize', fontSize);
title('CDF T OUT', 'FontSize', fontSize);
legend('normal CDF','T data CDF','Location','southeast');

figure(29)
subplot(3,3,6)
normplot(T_SH_ON_A_COL)

%3 different ways to prove that T out distribution follows normal
%distribution (with PDF; CDF and with normplot test (iof data are normal,
%plot is linear, as it happens here)

%find T where CDF=0.95
[CDF_DATA_SH_ON,T_CDF_DATA_SH_ON]=ecdf(T_SH_ON_A_COL);  %find CDF value of
empirical T profile

limit=length(CDF_DATA_SH_ON);
j=limit;
save_T=[];
i=1;
for k=limit:(-1):1
    control_P=CDF_DATA_SH_ON(j);
    control_T=T_CDF_DATA_SH_ON(j);
    if control_P>=0.95
        save_T(i)=control_T;
        i=i+1;
        j=j-1;
    else
        j=j-1;
    end
end

limit_2=length(save_T);
CPT=save_T(1,limit_2);

%insert CPT into plot of Demand vs Tout
vector_Y_CPT=[0,max(ENERGY_SH_ON_A_COL)];
vector_X_CPT=[CPT,CPT];

figure(29)
subplot(3,3,1)
plot(vector_X_CPT,vector_Y_CPT,'r')
```

```matlab
hold off
legend('Demand SH on','Demand SH off','fitting SH on','fitting SH off','CPT');

%find interception between CPT and fitting of SH on
CPD=polyval(fit_SH_ON,CPT); %change point demand

%find the maximum temperature of the year:
%(considered year 2020)

TEMPERATURE_2020=xlsread('weather data verona 2020.xlsx','B3:B8763');
MAX_T=max(TEMPERATURE_2020);
SUMMER_Y=[CPD,CPD];
SUMMER_X=[CPT, MAX_T];

%% use CPD to delete part of 0 demand to create a better model
limit_1= length(ENERGY_SH_ON_A_COL);
ENERGY_SH_ON_A_COL_mod=[];
T_SH_ON_A_COL_mod=[];
i=1;
for k=1:limit_1
    control_E=ENERGY_SH_ON_A_COL(k);
    control_T=T_SH_ON_A_COL(k);
    if control_E>=CPD
        ENERGY_SH_ON_A_COL_mod(i)=control_E;
        T_SH_ON_A_COL_mod(i)=control_T;
        i=i+1;
    end
end

%% interpolation line from Tmin to CPT
T_min=min(T_SH_ON_A_COL);
T_SH_ON_vector_X=[T_min, CPT];
T_SH_ON_vector_Y=[polyval(fit_SH_ON,T_min), CPD];



%PLOT complete demand, only SH on
figure(29)
subplot(3,3,7)
plot(T_SH_ON_A_COL,ENERGY_SH_ON_A_COL,'g.')
hold on
plot(T_SH_ON_vector_X,T_SH_ON_vector_Y,'b')
plot(vector_X_CPT,vector_Y_CPT,'r')
plot(SUMMER_X,SUMMER_Y,'b')
grid on
xlabel('T out', 'FontSize', fontSize);
ylabel('Demand', 'FontSize', fontSize);
title('Complete rebuilt consumption profile', 'FontSize', fontSize);
legend('experimental data SH on','interpolation','CPT','rebuilt summer
profile');

%% interpolation of modified vector of demand and plot
fit_SH_on_mod=polyfit(T_SH_ON_A_COL_mod,ENERGY_SH_ON_A_COL_mod,1);
CPD_mod=polyval(fit_SH_on_mod,CPT);
val_SH_on_mod=polyval(fit_SH_on_mod,T_SH_ON_A_COL_mod);

T_min_mod=min(T_SH_ON_A_COL_mod);
T_SH_ON_vector_X_mod=[T_min_mod, CPT];
T_SH_ON_vector_Y_mod=[polyval(fit_SH_on_mod,T_min_mod), CPD_mod];

SUMMER_Y_mod=[CPD_mod,CPD_mod];
```

```matlab
SUMMER_X_mod=[CPT, MAX_T];

%R2 calcualtion
SSE_SH_mod=sum((ENERGY_SH_ON_A_COL_mod-val_SH_on_mod).^2);
SSYY_SH_mod=sum((ENERGY_SH_ON_A_COL_mod-mean(ENERGY_SH_ON_A_COL_mod)).^2);
R2_nor_SH_mod=1-(SSE_SH_mod/SSYY_SH_mod);
fit_SH_on_mod_prova=fitlm(T_SH_ON_A_COL_mod,ENERGY_SH_ON_A_COL_mod,'linear');
fit_SH_on_prova=fitlm(T_SH_ON_A_COL,ENERGY_SH_ON_A_COL,'linear');


%PLOT complete demand, only SH on modified values
figure(29)
subplot(3,3,8)
plot(T_SH_ON_A_COL_mod,ENERGY_SH_ON_A_COL_mod,'g.')
hold on
plot(T_SH_ON_vector_X_mod,T_SH_ON_vector_Y_mod,'b')
plot(vector_X_CPT,vector_Y_CPT,'r')
plot(SUMMER_X_mod,SUMMER_Y_mod,'b')
grid on
xlabel('T out', 'FontSize', fontSize);
ylabel('Demand', 'FontSize', fontSize);
title('Complete rebuilt consumption profile without low demand', 'FontSize',
fontSize);
legend('experimental data SH on, no low values','interpolation','CPT','rebuilt
summer profile');
%% Calcolo SH e DHW
%T dependent part of the ESC is increased by a costant coefficient. The
%goal is to calculate it in order to move down the ESC to the x axis.

%Build T vector for summer period (T_SHW_year)
%from 11 march to 26 november 2020 (dates depend on dataset in analysis)

T_DHW_year=xlsread('weather data verona 2020.xlsx','B1682:B7945'); %but for DHW
we need only summer part (higher than CPT)
limit_5=length(T_DHW_year);
i=1;
T_DHW_year_summer=[];
for k=1:limit_5
    control_T=T_DHW_year(k);
    if control_T>=CPT
        T_DHW_year_summer(i)=control_T;
        i=i+1;
    end
end

T_DHW_year=T_DHW_year_summer';   %correction on the vector "T_DHW_year": insert
only values >= CPT

T_SH_year=T_SH_ON_A_COL; %but we need value up to CPT, no higher
limit_6=length(T_SH_year);
i=1;
T_SH_year_winter=[];
ENERGY_SH_ON_winter=[]; %even for experimental points it is necessary a vector
with no points higher than CPT (sam elength as T vector)
for k=1:limit_6
    control_T=T_SH_year(k);
    control_E=ENERGY_SH_ON_A_COL(k);
    if control_T<=CPT
        T_SH_year_winter(i)=control_T;
        ENERGY_SH_ON_winter(i)=control_E;
        i=i+1;
```

```matlab
    end
end
T_SH_year=T_SH_year_winter;


lung=length(T_DHW_year);
Demand_NOR_SH_year=polyval(fit_SH_on_mod,T_SH_year); %values obtained from model
for SH
Demand_NOR_DHW_year=ones(1,lung);
Demand_NOR_DHW_year(:,:)=CPD_mod;                        %values from model for DHW

T_year=[T_SH_year';T_DHW_year]; %complete vector with T data
%%
figure(40)
subplot(3,3,1)
plot(T_SH_year,Demand_NOR_SH_year)
hold on
plot(T_DHW_year,Demand_NOR_DHW_year)
grid on
xlabel('T out', 'FontSize', fontSize);
ylabel('ENERGY', 'FontSize', fontSize);
title('ENERGY vs T out', 'FontSize', fontSize);

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

%vector with energy consumption for the year:
%winter: ENERGY_SH_ON_winter (real data from meter but with T<= CPT)
%summer: Demand_NOR_DHW_year (simulated data for the rest of the year)

ENERGY_year=[ENERGY_SH_ON_winter,Demand_NOR_DHW_year];  %complete vector with E
data (experimental)

figure(40)
subplot(3,3,1)
plot(T_year,ENERGY_year,'.b')
hold on

%%
% Demand_shifted_SH=Demand_NOR_SH_year-min(Demand_NOR_DHW_year);  %shifting
demand down
% Demand_shifted_DHW=Demand_NOR_DHW_year-min(Demand_NOR_DHW_year);

Demand_shifted_SH=Demand_NOR_SH_year;  %test senza shiftare giù
Demand_shifted_DHW=Demand_NOR_DHW_year;

figure (40)
subplot(3,3,1)
plot(T_SH_year,Demand_shifted_SH)
plot(T_DHW_year,Demand_shifted_DHW)
hold off
legend('Demand winter','Demand summer','experimental points','Demand winter
down','Demand summer down');

ESC1_year=polyfit(T_SH_year,Demand_shifted_SH,1);
ESC2_year=polyfit(T_DHW_year,Demand_shifted_DHW,1);

val_1_year=polyval(ESC1_year,T_SH_year);
val_2_year=polyval(ESC2_year,T_DHW_year);

%%
```

```matlab
%Put values in chronological order with cicle for
%in this construction they are in chronological order. from
%november-november
limit_8=length(ENERGY_year);
time_vector=1:limit_8;


%to have a more clear plot and analysis it is possibile to remove some
%points from summer consumption (itis constant, so no problem)
%I choose 900 values to consider
ENERGY_year_lim=ENERGY_year(1:900);
limit_9=length(ENERGY_year_lim);
time_vector_lim=1:limit_9;


figure(40)
subplot(3,3,2)
plot(time_vector_lim,ENERGY_year_lim,'b')
grid on
xlabel('Time', 'FontSize', fontSize);
ylabel('ENERGY', 'FontSize', fontSize);
title('ENERGY Demand', 'FontSize', fontSize);
legend('energy demand for 900 hours')

%% Losses DHW evaluation
%Losses in the DHW circuit: not possible to apply method used for Tartu
%(summer consumption during night is unknown) but it is possible to
%estimate losses using moment when SH is off during winter. Consider night
%time: between 0.00 am and 4.00 am

%create vector with demand values from 0.00 am to 4.00 am
ENERGY_night=[ENERGY_0AM;ENERGY_1AM;ENERGY_2AM;ENERGY_3AM];
TEMPERATURE_night=[T_0AM;T_1AM;T_2AM;T_3AM];

%Removal of 5% tail of lowest temperature, then for losses considered mean
%values of night time when SH is off (only for points considered, after
%remotion)

[CDF_E_SH_OFF,VALUES_CDF_E_SH_OFF]=ecdf(TEMPERATURE_night);  %find CDF value of
empirical Demand night profile

limit=length(CDF_E_SH_OFF);
j=limit;
save_T_night=[];
i=1;
for k=limit:(-1):1
    control_P=CDF_E_SH_OFF(j);
    control_T=VALUES_CDF_E_SH_OFF(j);
    if control_P>=0.90
        save_T_night(i)=control_T;
        i=i+1;
        j=j-1;
    else
        j=j-1;
    end
end

limit_2=length(save_T_night);
T_DHW_losses_limit=save_T_night(1,limit_2); %T out limit for considering losses
(demand higher than that is ok to consider)

%fitting of night data
```

```matlab
fit_SH_OFF_night=polyfit(TEMPERATURE_night,ENERGY_night,1);
values_SH_OFF_night=polyval(fit_SH_OFF_night,TEMPERATURE_night);

%vector for plotting values of T_DHW_losses_limit
vector_T_DHW_losses_limit_x=[T_DHW_losses_limit,T_DHW_losses_limit];
vector_T_DHW_losses_limit_y=[0,max(ENERGY_night)];

%DHW circuit losses during summer estimation (mean value in the hottest
%part of the SH off period at night)
DHW_ENERGY_for_losses=[];
limit=length(TEMPERATURE_night);
k=limit;
i=1;
for j=limit:(-1):1
    control_E=ENERGY_night(j);
    control_T=TEMPERATURE_night(j);
    if control_T>=T_DHW_losses_limit
        DHW_ENERGY_for_losses(i)=control_E;
        i=i+1;
    else
    end
end

DHW_losses=mean(DHW_ENERGY_for_losses); %losses as mean value of highest demand
point during winter night time
DHW_losses_x=[T_DHW_losses_limit, max(TEMPERATURE_night)]; %vector for plotting
lines
DHW_losses_y=[DHW_losses, DHW_losses];

%norm pdf mean and std from T out at night dataset
MEAN_T_SH_OFF=mean(TEMPERATURE_night);
STD_T_SH_OFF=std(TEMPERATURE_night);
MIN_T_SH_OFF=min(TEMPERATURE_night);
MAX_T_SH_OFF=max(TEMPERATURE_night);
vector_norm_night=MIN_T_SH_OFF:0.1:MAX_T_SH_OFF;
NORM_CDF_T_SH_OFF=normcdf(vector_norm_night,MEAN_T_SH_OFF,STD_T_SH_OFF);

axis_limit=[-6 13 0 65];  %set axes limit to have same dimensions as plot before

figure (40)
subplot(3,3,5)
plot(TEMPERATURE_night,ENERGY_night,'.k')
hold on
plot(TEMPERATURE_night,values_SH_OFF_night,'b')
plot(vector_T_DHW_losses_limit_x,vector_T_DHW_losses_limit_y,'r')
plot(DHW_losses_x,DHW_losses_y,'g')
grid on
hold off
axis(axis_limit_1);
xlabel('T out', 'FontSize', fontSize);
ylabel('ENERGY AT NIGHT', 'FontSize', fontSize);
title('ENERGY AT NIGHT vs T out', 'FontSize', fontSize);
legend('Demand between 0.00 am - 4.00 am','fitting','T limit for DHW losses
estimation','DHW losses');

%%
figure (40)
subplot(3,3,3)
plot(VALUES_CDF_E_SH_OFF,CDF_E_SH_OFF)
hold on
plot(vector_norm_night,NORM_CDF_T_SH_OFF)
```

```matlab
grid on
xlabel('T out at night', 'FontSize', fontSize);
ylabel('Probability', 'FontSize', fontSize);
title('T out AT NIGHT', 'FontSize', fontSize);
legend('T out night','Normal distribution');

figure(40)
subplot(3,3,4)
normplot(TEMPERATURE_night)

%%
E_loss_year=DHW_losses;

%put in chronological order points obtained from ESC model
i=1;
value_ESC_year=[];
limit_10=length(T_year);

for k=1:limit_10
    if T_year(k)<=CPT
        value_ESC_year(i)=polyval(ESC1_year,T_year(k));
        i=i+1;
    else
        value_ESC_year(i)=polyval(ESC2_year,T_year(k));
        i=i+1;
    end
end

%calcolo di E SH
% E_SH_year=value_ESC_year-min(value_ESC_year);
E_SH_year=value_ESC_year; %non la sposto in giù

%calcolo di E DHW
E_DHW_year=[];
i=1;
limit_3=length(ENERGY_year);
for k=1:limit_3
    if ENERGY_year(k)>E_SH_year(k)
        E_DHW_year(i)=ENERGY_year(k)-E_SH_year(k)+E_loss_year;
        i=i+1;
    else
        E_DHW_year(i)=E_loss_year;
        i=i+1;
    end
end

%%
%Balnacing E_SH and E_SHW
E_SH_balanced_year=ENERGY_year-E_DHW_year;

%Correction process for values <0
%"corr" vector are balanced + all positive vectors
E_SH_year_corr=E_SH_year;
E_SH_balanced_year_corr=E_SH_balanced_year;
E_DHW_year_corr=E_DHW_year;

limit_4=length(E_SH_balanced_year);

for k=1:limit_4
    if E_SH_balanced_year(k)<0
```

```matlab
        E_DHW_year_corr(k)=E_DHW_year_corr(k)+E_SH_balanced_year(k);
        E_SH_balanced_year_corr(k)=0;
    end
    if E_DHW_year(k)<0
        E_SH_balanced_year_corr(k)=E_SH_balanced_year_corr(k)+E_DHW_year(k);
        E_DHW_year_corr(k)=0;
    end
end


%%
fontSize=18;

figure(40)
subplot(3,3,6)
plot(T_SH_year,val_1_year,'g')
hold on
grid on
plot(T_year,ENERGY_year,'.b')
plot(T_SH_year,Demand_NOR_SH_year,'r')
plot(T_DHW_year,val_2_year,'g')
plot(T_DHW_year,Demand_NOR_DHW_year,'r')
hold off
xlabel('DMT', 'FontSize', fontSize);
ylabel('Demand', 'FontSize', fontSize);
title('ESC complete', 'FontSize', fontSize);
legend('ESC shifted','Measured heat use 0AM','ESC of total heat use')
% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

guida=1:limit_3;
E_loss_year_plot(1:limit_3)=E_loss_year;
subplot(3,3,7)
plot(guida,ENERGY_year,'b')
hold on
grid on
plot(guida,E_loss_year_plot,'r')
xlabel('time (day)', 'FontSize', fontSize);
ylabel('Demand 0AM', 'FontSize', fontSize);
title('minimum value of demand at 0AM', 'FontSize', fontSize);
legend('Demand at 0AM','value for DHW losses')
hold off
%creation of smaller values for more clear plot (900 values)
guida_lim=1:900;
E_SH_balanced_year_corr_lim=E_SH_balanced_year_corr(1:900);
E_DHW_year_corr_lim=E_DHW_year_corr(1:900);

subplot(3,3,8)
plot(guida_lim,ENERGY_year_lim,'b')
hold on
grid on
plot(guida_lim,E_SH_balanced_year_corr_lim,'g')
plot(guida_lim,E_DHW_year_corr_lim,'r')
xlabel('time (day)', 'FontSize', fontSize);
ylabel('Demand 0AM', 'FontSize', fontSize);
title('cronological Demand balanced and correct', 'FontSize', fontSize);
legend('Measured Demand','SH demand balanced >0','DHW demand >0')
hold off

subplot(3,3,9)
plot(T_year,ENERGY_year,'.b')
hold on
```

```matlab
grid on
plot(T_year,E_SH_balanced_year_corr,'.g')
plot(T_year,E_DHW_year_corr,'.r')
xlabel('DMT (°C)', 'FontSize', fontSize);
ylabel('Demand 0AM', 'FontSize', fontSize);
title('Demand splitted', 'FontSize', fontSize);
legend('Measured Demand','SH demand','DHW demand')
hold off

%model is ok, but DHW consumption is higher than SH. this is ok because it
%is evident that slope of SH model is low and by not knowing data i estimate
%DHW consumption in summer but when heating season was not finished

%maybe it is possibile to estimate DHW consumption during summer from
%points with SH off. Try to use these values to built a new ESC with
%significantly lower DHW consumption in summer

%% Improvement with outlier removal?
%identify when points of extremely high demand happens
%% TMA dependency control, corrected
%TMA determination from 1 to 48 hours before. only with SH part of demand
% T_SH_ON_A_COL,ENERGY_SH_ON_A_COL
BETA=[];
i=1;
for k=1:48
    TMA_hour=movmean(T_SH_ON_A_COL,[k 0]);
    pol_hour=polyfit(TMA_hour,ENERGY_SH_ON_A_COL,1);
    control_beta=pol_hour(1,1);
    BETA(i)=control_beta;
    i=i+1;
end

guida_1=1:length(BETA);

figure (26)
subplot(2,1,1)
plot(T_SH_ON_A_COL,ENERGY_SH_ON_A_COL,'.')
grid on
xlabel('T OUTDOOR', 'FontSize', fontSize);
ylabel('ENERGY USE', 'FontSize', fontSize);
title('T OUTDOOR vs ENERGY USE', 'FontSize', fontSize);

subplot(2,1,2)
plot(guida_1,BETA)
grid on
xlabel('Lag of Temperature MA (hours)', 'FontSize', fontSize);
ylabel('Correlation between ENERGY and TMA', 'FontSize', fontSize);
title('TMA vs ENERGY USE', 'FontSize', fontSize);

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);
%% new script but this time consider initial vector without data outside CI (too
high data)
%select and remove data outside confidence interval
%starting from vector "ENERGY_SH_ON_A_COL"

ENERGY_OK_VALUES=[];
TEMPERATURE_OK_VALUES=[];
i=1;
limit_11=length(ENERGY_SH_ON_A_COL);
```

```
for k=1:limit_11
    control_E=ENERGY_SH_ON_A_COL(k);
    control_T=T_SH_ON_A_COL(k);
    control_CI_up=y_CI_up_on(k);
    control_CI_down=y_CI_down_on(k);
    if control_E>=control_CI_down && control_E<=control_CI_up
        ENERGY_OK_VALUES(i)=control_E;
        TEMPERATURE_OK_VALUES(i)=control_T;
        i=i+1;
    end
end

figure (41)
subplot(3,3,1)
plot(T_SH_ON_A_COL,ENERGY_SH_ON_A_COL,'.b')
hold on
plot(TEMPERATURE_OK_VALUES,ENERGY_OK_VALUES,'.g')
plot(T_SH_ON_A_COL,y_CI_up_on,'g')
plot(T_SH_ON_A_COL,y_CI_down_on,'g')
grid on
xlabel('T out', 'FontSize', fontSize);
ylabel('Energy', 'FontSize', fontSize);
title('CI on ENERGY', 'FontSize', fontSize);
legend('experimental points','values inside CI','UPPER 95% CI','LOWER 95% CI');
% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

%% ESC reconstruction
%consider last values, find a mean values between them and interpolate. it
%is necessary the vector that contains temperature of Verona in the whole
%year. we reconstruct the summer demand with constant values.

%create vector with Tout from the complete year

%for choosing CPT build frequence distribution of T out and cancel tails of
%the distribution (in particular higher tail). Use typical 95% of
%tolerance

LUNG_T_ON_CI=length(TEMPERATURE_OK_VALUES);
LUNG_T_OFF_CI=length(T_SH_OFF_A_COL);

[DIST_T_SH_ON_ASS_CI,X_T_SH_ON_CI]=hist(TEMPERATURE_OK_VALUES,24);
[DIST_T_SH_OFF_ASS,X_T_SH_OFF]=hist(T_SH_OFF_A_COL,24);
DIST_T_SH_ON_CI=DIST_T_SH_ON_ASS_CI/LUNG_T_ON_CI;
DIST_T_SH_OFF=DIST_T_SH_OFF_ASS/LUNG_T_OFF_CI;

%norm pdf mean and std from T dataset
MEAN_T_SH_ON_CI=mean(TEMPERATURE_OK_VALUES);
STD_T_SH_ON_CI=std(TEMPERATURE_OK_VALUES);
MIN_T_SH_ON_CI=min(TEMPERATURE_OK_VALUES);
MAX_T_SH_ON_CI=max(TEMPERATURE_OK_VALUES);
vector_norm_CI=MIN_T_SH_ON_CI:0.1:MAX_T_SH_ON_CI;
NORM_T_SH_ON_CI=normpdf(vector_norm_CI,MEAN_T_SH_ON_CI,STD_T_SH_ON_CI);
NORM_CDF_T_SH_ON_CI=normcdf(vector_norm_CI,MEAN_T_SH_ON_CI,STD_T_SH_ON_CI);

%build vector of CDF for DIST_T_SH_ON

figure(41)
subplot(3,3,2)
plot(vector_norm_CI,NORM_T_SH_ON_CI,'b')
```

```matlab
grid on
hold on
% plot(X_T_SH_ON,DIST_T_SH_ON,'g')
bar(X_T_SH_ON_CI,DIST_T_SH_ON_CI,'g')
hold off
xlabel('T out', 'FontSize', fontSize);
ylabel('probability', 'FontSize', fontSize);
title('PDF T OUT (CI on)', 'FontSize', fontSize);
legend('normal PDF','T PDF');

figure(41)
subplot(3,3,3)
plot(vector_norm_CI,NORM_CDF_T_SH_ON_CI)
hold on
%cdfplot(T_SH_ON_A_COL)
ecdf(TEMPERATURE_OK_VALUES,'Bounds','on')
grid on
xlabel('T out', 'FontSize', fontSize);
ylabel('probability', 'FontSize', fontSize);
title('CDF T OUT (CI on)', 'FontSize', fontSize);
legend('normal CDF','T data CDF','Location','southeast');

figure(41)
subplot(3,3,4)
normplot(TEMPERATURE_OK_VALUES)

%3 different ways to prove that T out distribution follows normal
%distribution (with PDF; CDF and with normplot test (iof data are normal,
%plot is linear, as it happens here)

%find T where CDF=0.95
[CDF_DATA_SH_ON_CI,T_CDF_DATA_SH_ON_CI]=ecdf(TEMPERATURE_OK_VALUES);  %find CDF
value of empirical T profile

limit=length(CDF_DATA_SH_ON_CI);
j=limit;
save_T_CI=[];
i=1;
for k=limit:(-1):1
    control_P=CDF_DATA_SH_ON_CI(j);
    control_T=T_CDF_DATA_SH_ON_CI(j);
    if control_P>=0.95
        save_T_CI(i)=control_T;
        i=i+1;
        j=j-1;
    else
        j=j-1;
    end
end

limit_2=length(save_T_CI);
CPT_CI=save_T_CI(1,limit_2);

%insert CPT into plot of Demand vs Tout
vector_Y_CPT_CI=[0,max(ENERGY_OK_VALUES)];
vector_X_CPT_CI=[CPT_CI,CPT_CI];

figure(41)
subplot(3,3,1)
plot(vector_X_CPT_CI,vector_Y_CPT_CI,'r')
```

```matlab
hold off
legend('Demand SH on','Demand SH off','fitting SH on','fitting SH off','CPT');

%find new fitting coefficient
fit_SH_on_CI=polyfit(TEMPERATURE_OK_VALUES,ENERGY_OK_VALUES,1);

%find interception between CPT and fitting of SH on
CPD_CI=polyval(fit_SH_on_CI,CPT_CI); %change point demand

%find the maximum temperature of the year:
%(considered year 2020)

TEMPERATURE_2020=xlsread('weather data verona 2020.xlsx','B3:B8763');
MAX_T=max(TEMPERATURE_2020);
SUMMER_Y=[CPD_CI,CPD_CI];
SUMMER_X=[CPT_CI, MAX_T];

%interpolation line from Tmin to CPT
T_min=min(T_SH_ON_A_COL);
T_SH_ON_vector_X=[T_min, CPT_CI];
T_SH_ON_vector_Y=[polyval(fit_SH_on_CI,T_min), CPD_CI];


%PLOT complete demand, only SH on
figure(41)
subplot(3,3,5)
plot(TEMPERATURE_OK_VALUES,ENERGY_OK_VALUES,'g.')
hold on
plot(T_SH_ON_vector_X,T_SH_ON_vector_Y,'b')
plot(vector_X_CPT_CI,vector_Y_CPT_CI,'r')
plot(SUMMER_X,SUMMER_Y,'b')
grid on
xlabel('T out', 'FontSize', fontSize);
ylabel('Demand', 'FontSize', fontSize);
title('Complete rebuilt consumption profile (CI on)', 'FontSize', fontSize);
legend('Data with CI processing','interpolation','CPT','rebuilt summer
profile');

%% Calcolo SH e DHW
%T dependent part of the ESC is increased by a costant coefficient. The
%goal is to calculate it in order to move down the ESC to the x axis.

%Build T vector for summer period (T_SHW_year)
%from 11 march to 26 november 2020 (dates depend on dataset in analysis)

T_DHW_year_CI=xlsread('weather data verona 2020.xlsx','B1682:B7945'); %but for
DHW we need only summer part (higher than CPT)
limit_5=length(T_DHW_year_CI);
i=1;
T_DHW_year_summer_CI=[];
for k=1:limit_5
    control_T=T_DHW_year_CI(k);
    if control_T>=CPT_CI
        T_DHW_year_summer_CI(i)=control_T;
        i=i+1;
    end
end

T_DHW_year_CI=T_DHW_year_summer_CI';  %correction on the vector "T_DHW_year":
insert only values >= CPT
```

```matlab
T_SH_year_CI=TEMPERATURE_OK_VALUES; %but we need value up to CPT, no higher
limit_6=length(T_SH_year_CI);
i=1;
T_SH_year_winter_CI=[];
ENERGY_SH_ON_winter_CI=[]; %even for experimental points it is necessary a
vector with no points higher than CPT (sam elength as T vector)
for k=1:limit_6
    control_T=T_SH_year_CI(k);
    control_E=ENERGY_OK_VALUES(k);
    if control_T<=CPT_CI
        T_SH_year_winter_CI(i)=control_T;
        ENERGY_SH_ON_winter_CI(i)=control_E;
        i=i+1;
    end
end
T_SH_year_CI=T_SH_year_winter_CI;

lung=length(T_DHW_year_CI);
Demand_NOR_SH_year_CI=polyval(fit_SH_on_CI,T_SH_year_CI); %values obtained from
model for SH
Demand_NOR_DHW_year_CI=ones(1,lung);
Demand_NOR_DHW_year_CI(:,:)=CPD_CI;                       %values from model for
DHW

T_year_CI=[T_SH_year_CI';T_DHW_year_CI]; %complete vector with T data

%%
figure(42)
subplot(3,3,1)
plot(T_SH_year_CI,Demand_NOR_SH_year_CI)
hold on
plot(T_DHW_year_CI,Demand_NOR_DHW_year_CI)
grid on
xlabel('T out', 'FontSize', fontSize);
ylabel('ENERGY', 'FontSize', fontSize);
title('ENERGY vs T out (CI on)', 'FontSize', fontSize);

% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

%vector with energy consumption for the year:
%winter: ENERGY_SH_ON_winter (real data from meter but with T<= CPT)
%summer: Demand_NOR_DHW_year (simulated data for the rest of the year)

ENERGY_year_CI=[ENERGY_SH_ON_winter_CI,Demand_NOR_DHW_year_CI];  %complete
vector with E data (experimental)

figure(42)
subplot(3,3,1)
plot(T_year_CI,ENERGY_year_CI,'.b')
hold on

%%
% Demand_shifted_SH_CI=Demand_NOR_SH_year_CI-min(Demand_NOR_DHW_year_CI);
%shifting demand down
% Demand_shifted_DHW_CI=Demand_NOR_DHW_year_CI-min(Demand_NOR_DHW_year_CI);

Demand_shifted_SH_CI=Demand_NOR_SH_year_CI;  %no shifting down
Demand_shifted_DHW_CI=Demand_NOR_DHW_year_CI;
```

```matlab
figure (42)
subplot(3,3,1)
plot(T_SH_year_CI,Demand_shifted_SH_CI)
plot(T_DHW_year_CI,Demand_shifted_DHW_CI)
hold off
legend('Demand winter','Demand summer','experimental points','Demand winter
down','Demand summer down');

ESC1_year_CI=polyfit(T_SH_year_CI,Demand_shifted_SH_CI,1);
ESC2_year_CI=polyfit(T_DHW_year_CI,Demand_shifted_DHW_CI,1);

val_1_year_CI=polyval(ESC1_year_CI,T_SH_year_CI);
val_2_year_CI=polyval(ESC2_year_CI,T_DHW_year_CI);

%%
%Put values in chronological order with cicle for
%in this construction they are in chronological order. from
%november-november
limit_8=length(ENERGY_year_CI);
time_vector=1:limit_8;

%to have a more clear plot and analysis it is possibile to remove some
%points from summer consumption (itis constant, so no problem)
%I choose 900 values to consider
ENERGY_year_lim_CI=ENERGY_year_CI(1:900);
limit_9=length(ENERGY_year_lim_CI);
time_vector_lim_CI=1:limit_9;

figure(42)
subplot(3,3,2)
plot(time_vector_lim_CI,ENERGY_year_lim_CI,'b')
grid on
xlabel('Time', 'FontSize', fontSize);
ylabel('ENERGY', 'FontSize', fontSize);
title('ENERGY Demand (CI on)', 'FontSize', fontSize);
legend('energy demand for 900 hours')

%% Losses DHW evaluation
%Losses in the DHW circuit: not possible to apply method used for Tartu
%(summer consumption during night is unknown) but it is possible to
%estimate losses using moment when SH is off during winter. Consider night
%time: between 0.00 am and 4.00 am

%create vector with demand values from 0.00 am to 4.00 am
ENERGY_night=[ENERGY_0AM;ENERGY_1AM;ENERGY_2AM;ENERGY_3AM];
TEMPERATURE_night=[T_0AM;T_1AM;T_2AM;T_3AM];

%Removal of 5% tail of lowest temperature, then for losses considered mean
%values of night time when SH is off (only for points considered, after
%remotion)

[CDF_E_SH_OFF,VALUES_CDF_E_SH_OFF]=ecdf(TEMPERATURE_night);  %find CDF value of
empirical Demand night profile

limit=length(CDF_E_SH_OFF);
j=limit;
save_T_night=[];
i=1;
for k=limit:(-1):1
```

```matlab
        control_P=CDF_E_SH_OFF(j);
        control_T=VALUES_CDF_E_SH_OFF(j);
        if control_P>=0.90
            save_T_night(i)=control_T;
            i=i+1;
            j=j-1;
        else
            j=j-1;
        end
end


limit_2=length(save_T_night);
T_DHW_losses_limit=save_T_night(1,limit_2); %T out limit for considering losses
(demand higher than that is ok to consider)


%fitting of night data
fit_SH_OFF_night=polyfit(TEMPERATURE_night,ENERGY_night,1);
values_SH_OFF_night=polyval(fit_SH_OFF_night,TEMPERATURE_night);


%vector for plotting values of T_DHW_losses_limit
vector_T_DHW_losses_limit_x=[T_DHW_losses_limit,T_DHW_losses_limit];
vector_T_DHW_losses_limit_y=[0,max(ENERGY_night)];


%DHW circuit losses during summer estimation (mean value in the hottest
%part of the SH off period at night)
DHW_ENERGY_for_losses=[];
limit=length(TEMPERATURE_night);
k=limit;
i=1;
for j=limit:(-1):1
    control_E=ENERGY_night(j);
    control_T=TEMPERATURE_night(j);
    if control_T>=T_DHW_losses_limit
        DHW_ENERGY_for_losses(i)=control_E;
        i=i+1;
    else
    end
end


DHW_losses=mean(DHW_ENERGY_for_losses); %losses as mean value of highest demand
point during winter night time
DHW_losses_x=[T_DHW_losses_limit, max(TEMPERATURE_night)]; %vector for plotting
lines
DHW_losses_y=[DHW_losses, DHW_losses];


%norm pdf mean and std from T out at night dataset
MEAN_T_SH_OFF=mean(TEMPERATURE_night);
STD_T_SH_OFF=std(TEMPERATURE_night);
MIN_T_SH_OFF=min(TEMPERATURE_night);
MAX_T_SH_OFF=max(TEMPERATURE_night);
vector_norm_night=MIN_T_SH_OFF:0.1:MAX_T_SH_OFF;
NORM_CDF_T_SH_OFF=normcdf(vector_norm_night,MEAN_T_SH_OFF,STD_T_SH_OFF);


axis_limit=[-6 13 0 65];   %set axes limit to have same dimensions as plot before


figure (42)
subplot(3,3,5)
plot(TEMPERATURE_night,ENERGY_night,'.k')
hold on
plot(TEMPERATURE_night,values_SH_OFF_night,'b')
```

```matlab
plot(vector_T_DHW_losses_limit_x,vector_T_DHW_losses_limit_y,'r')
plot(DHW_losses_x,DHW_losses_y,'g')
grid on
hold off
axis(axis_limit_1);
xlabel('T out', 'FontSize', fontSize);
ylabel('ENERGY AT NIGHT', 'FontSize', fontSize);
title('ENERGY AT NIGHT vs T out (CI on)', 'FontSize', fontSize);
legend('Demand between 0.00 am - 4.00 am','fitting','T limit for DHW losses
estimation','DHW losses');

%%
figure (42)
subplot(3,3,3)
plot(VALUES_CDF_E_SH_OFF,CDF_E_SH_OFF)
hold on
plot(vector_norm_night,NORM_CDF_T_SH_OFF)
grid on
xlabel('T out at night', 'FontSize', fontSize);
ylabel('Probability', 'FontSize', fontSize);
title('T out AT NIGHT (CI on)', 'FontSize', fontSize);
legend('T out night','Normal distribution');

figure(42)
subplot(3,3,4)
normplot(TEMPERATURE_night)

%%
E_loss_year=DHW_losses;

%put in chronological order points obtained from ESC model
i=1;
value_ESC_year_CI=[];
limit_10=length(T_year_CI);

for k=1:limit_10
    if T_year_CI(k)<=CPT_CI
        value_ESC_year_CI(i)=polyval(ESC1_year_CI,T_year_CI(k));
        i=i+1;
    else
        value_ESC_year_CI(i)=polyval(ESC2_year_CI,T_year_CI(k));
        i=i+1;
    end
end

%calcolo di E SH
% E_SH_year_CI=value_ESC_year_CI-min(value_ESC_year_CI);
E_SH_year_CI=value_ESC_year_CI; %without subtraction in school

%calcolo di E DHW
E_DHW_year_CI=[];
i=1;
limit_3=length(ENERGY_year_CI);
for k=1:limit_3
    if ENERGY_year_CI(k)>E_SH_year_CI(k)
        E_DHW_year_CI(i)=ENERGY_year_CI(k)-E_SH_year_CI(k)+E_loss_year;
        i=i+1;
    else
        E_DHW_year_CI(i)=E_loss_year;
        i=i+1;
    end
```

```matlab
end

%%
%Balnacing E_SH and E_SHW
E_SH_balanced_year_CI=ENERGY_year_CI-E_DHW_year_CI;

%Correction process for values <0
%"corr" vector are balanced + all positive vectors
E_SH_year_corr_CI=E_SH_year_CI;
E_SH_balanced_year_corr_CI=E_SH_balanced_year_CI;
E_DHW_year_corr_CI=E_DHW_year_CI;

limit_4=length(E_SH_balanced_year_CI);

for k=1:limit_4
    if E_SH_balanced_year_CI(k)<0
        E_DHW_year_corr_CI(k)=E_DHW_year_corr_CI(k)+E_SH_balanced_year_CI(k);
        E_SH_balanced_year_corr_CI(k)=0;
    end
    if E_DHW_year_CI(k)<0

E_SH_balanced_year_corr_CI(k)=E_SH_balanced_year_corr_CI(k)+E_DHW_year_CI(k);
        E_DHW_year_corr_CI(k)=0;
    end
end

%%
fontSize=18;

figure(42)
subplot(3,3,6)
plot(T_SH_year_CI,val_1_year_CI,'g')
hold on
grid on
plot(T_year_CI,ENERGY_year_CI,'.b')
plot(T_SH_year_CI,Demand_NOR_SH_year_CI,'r')
plot(T_DHW_year_CI,val_2_year_CI,'g')
plot(T_DHW_year_CI,Demand_NOR_DHW_year_CI,'r')
hold off
xlabel('DMT', 'FontSize', fontSize);
ylabel('Demand', 'FontSize', fontSize);
title('ESC complete (CI on)', 'FontSize', fontSize);
legend('ESC shifted','Measured heat use 0AM','ESC of total heat use')
% METTERE A SCHERMO INTERO
set(gcf, 'Units', 'Normalized', 'OuterPosition', [0, 0.04, 1, 0.96]);

guida_CI=1:limit_3;
E_loss_year_plot_CI(1:limit_3)=E_loss_year;
subplot(3,3,7)
plot(guida_CI,ENERGY_year_CI,'b')
hold on
grid on
plot(guida_CI,E_loss_year_plot_CI,'r')
xlabel('time (day)', 'FontSize', fontSize);
ylabel('Demand 0AM', 'FontSize', fontSize);
title('minimum value of demand at 0AM (CI on)', 'FontSize', fontSize);
legend('Demand at 0AM','value for DHW losses')
hold off
%creation of smaller values for more clear plot (900 values)
guida_lim=1:900;
E_SH_balanced_year_corr_lim_CI=E_SH_balanced_year_corr_CI(1:900);
```

```
E_DHW_year_corr_lim_CI=E_DHW_year_corr_CI(1:900);

subplot(3,3,8)
plot(guida_lim,ENERGY_year_lim_CI,'b')
hold on
grid on
plot(guida_lim,E_SH_balanced_year_corr_lim_CI,'g')
plot(guida_lim,E_DHW_year_corr_lim_CI,'r')
xlabel('time (day)', 'FontSize', fontSize);
ylabel('Demand 0AM', 'FontSize', fontSize);
title('cronological Demand balanced and correct (CI on)', 'FontSize', fontSize);
legend('Measured Demand','SH demand balanced >0','DHW demand >0')
hold off

subplot(3,3,9)
plot(T_year_CI,ENERGY_year_CI,'.b')
hold on
grid on
plot(T_year_CI,E_SH_balanced_year_corr_CI,'.g')
plot(T_year_CI,E_DHW_year_corr_CI,'.r')
xlabel('DMT (°C)', 'FontSize', fontSize);
ylabel('Demand 0AM', 'FontSize', fontSize);
title('Demand splitted (CI on)', 'FontSize', fontSize);
legend('Measured Demand','SH demand','DHW demand')
hold off

%model is ok, but DHW consumption is higher than SH. this is ok because it
%is evident that slope of SH model is low and by not knowing data i estimate
%DHW consumption in summer but when heating season was not finished

%maybe it is possibile to estimate DHW consumption during summer from
%points with SH off. Try to use these values to built a new ESC with
%significantly lower DHW consumption in summer
```

"NaN_removal_v2.m"

```
%function for NaN cleaning from the original dataset
%2 different outputs:
%ENERGY_N_NaN: vector with energy consumption without empty data
%ENERGY_D_NaN: vector with energyu consumption without days that contains
%at least one empty box in the original dataset

%delete every NaN element from the ENERGY vestor. in the original dataset
%there are some empty raws: probably meter was off or there were some
%errors in the data reading.

%do the same operation in the vector HOURS to be suyre that the Tout
%corresponds to the value of demand in the considered hour.

function
[ENERGY_N_NaN,HOURS_N_NaN,ENERGY_D_NaN,HOURS_D_NaN,NUMBER_DAY,ENERGY_D_NaN_ORDER
,WEEKEND_N_NaN_E_COL,WEEKDAYS_N_NaN_E_COL,WEEKEND_N_NaN_T_COL,WEEKDAYS_N_NaN_T_C
OL]=NaN_removal_v2(ENERGY,HOURS,CONTROL_POWER,starting)

limit=length(ENERGY);
ENERGY_N_NaN=[];
f=0;
i=1;
HOURS_N_NaN=[];
```

```matlab
for j=1:limit
    controlP=CONTROL_POWER(j);
    controlE=ENERGY(j);
    controlH=HOURS(j);
    if isnan(controlP)
        f=f+1;
    else
        ENERGY_N_NaN(i,1)=controlE;
        HOURS_N_NaN(i,1)=controlH;
        i=i+1;
    end
end

%% processing dataset: delete days with at least one NaN values

ENERGY_D_NaN=[];
HOURS_D_NaN=[];

%monday=1,%tuesday=2,%wednesday=3...Sunday=7
%insert here from which day the database start

NUMBER_DAY=[];
vector_number=starting-1;

day=1;

for a=1:limit
    day_end=day+23;

    if day_end<=limit
        f=1;
        vector_day=[];
        vector_hours=[];
        vector_control=[];
        vector_number=vector_number+1;
        if vector_number>7
            vector_number=vector_number-7;
        end
        for i=day:day_end
            vector_day(f)=ENERGY(i);
            vector_hours(f)=HOURS(i);
            vector_control(f)=CONTROL_POWER(i);
            f=f+1;
        end
    s=0;
    g=0;
    for k=1:24
        variable=vector_control(k);
        if isnan(variable)
         s=s+1;
        else
         g=g+1;
        end
    end

    if g==24
        ENERGY_D_NaN=[ENERGY_D_NaN,vector_day];
        HOURS_D_NaN=[HOURS_D_NaN,vector_hours];
        NUMBER_DAY=[NUMBER_DAY,vector_number];
```

248

```matlab
        end


    day=day+24;
    end
end


ENERGY_D_NaN=ENERGY_D_NaN';
HOURS_D_NaN=HOURS_D_NaN';
NUMBER_DAY=NUMBER_DAY';

%ok vector with number of days is good. Now split in colomn of 24 values
%ENERGY_D_NaN
start=1;
finish=24;
limit_2=length(ENERGY_D_NaN);
limit_3=(limit_2/24)+1;
ENERGY_D_NaN_ORDER=[];
HOURS_D_NaN_ORDER=[];

for k=1:limit_3
if finish<=limit_2
    vector_E=ENERGY_D_NaN(start:finish);
    vector_T=HOURS_D_NaN(start:finish);
    ENERGY_D_NaN_ORDER=[ENERGY_D_NaN_ORDER,vector_E];
    HOURS_D_NaN_ORDER=[HOURS_D_NaN_ORDER,vector_T];
    start=start+24;
    finish=finish+24;
end
end


%split days of weekend (saturday+sunday) and of rest of the week
limit_4=length(NUMBER_DAY);
i=1;
WEEKEND_N_NaN_E=[];
WEEKDAYS_N_NaN_E=[];
WEEKEND_N_NaN_T=[];
WEEKDAYS_N_NaN_T=[];

for k=1:limit_4
    control_day=NUMBER_DAY(i);
    control_vector_E=ENERGY_D_NaN_ORDER(:,i);
    control_vector_T=HOURS_D_NaN_ORDER(:,i);

    if control_day==6 || control_day==7
        WEEKEND_N_NaN_E=[WEEKEND_N_NaN_E,control_vector_E];
        WEEKEND_N_NaN_T=[WEEKEND_N_NaN_T,control_vector_T];
    else
        WEEKDAYS_N_NaN_E=[WEEKDAYS_N_NaN_E,control_vector_E];
        WEEKDAYS_N_NaN_T=[WEEKDAYS_N_NaN_T,control_vector_T];
    end
    i=i+1;

end


WEEKEND_N_NaN_E_COL=WEEKEND_N_NaN_E(:);
WEEKDAYS_N_NaN_E_COL=WEEKDAYS_N_NaN_E(:);

WEEKEND_N_NaN_T_COL=WEEKEND_N_NaN_T(:);
WEEKDAYS_N_NaN_T_COL=WEEKDAYS_N_NaN_T(:);
end
```

"SH_ONOFF.m"

```matlab
%function for splitting complete demand profile into SH on and SH off hours

function[ENERGY_SH_ON_A,ENERGY_SH_OFF_A,T_SH_ON_A,T_SH_OFF_A]=SH_ONOFF(COMPLETE_
ENERGY,COMPLETE_T,percentage,fit)

ENERGY_SH_ON_A=[];
ENERGY_SH_OFF_A=[];
T_SH_ON_A=[];
T_SH_OFF_A=[];
sizee=size(COMPLETE_ENERGY);
limit=sizee(1,1);

for k=1:24
    up=0;
    low=0;
    T_vector=COMPLETE_T(:,k);
    E_vector=COMPLETE_ENERGY(:,k);

    for j=1:limit
        T=T_vector(j);
        E=E_vector(j);
        compare=polyval(fit,T);

        if E>=compare
            up=up+1;
        else
            low=low+1;
        end
    end

    if up>=(percentage*limit)
            ENERGY_SH_ON_A=[ENERGY_SH_ON_A,E_vector];
            T_SH_ON_A=[T_SH_ON_A,T_vector];
    else
            ENERGY_SH_OFF_A=[ENERGY_SH_OFF_A,E_vector];
            T_SH_OFF_A=[T_SH_OFF_A,T_vector];
    end

end

end
```

References.

1. T. Catalina, V. Iordache, B. Caracaleanu, Multiple regression model for fast prediction of the heating energy demand, Energy Build. 57 (2013) 302–312. [3] S.S. Kwok, R.K. Yuen, E.W. Lee, An intelligent approach to assessing the effect of building occupancy on building cooling load prediction, Build. Environ. 46 (8) (2011) 1681–1690.

2. W. Zeyu, S.S. Ravi, A review of artificial intelligence-based building energy use prediction: Contrasting the capabilities of single and ensemble prediction models, Renew. Sustain. Energy Rev. 75 (2017) 796–808 75

3. A. Foucquier, S. Robert, F. Suard, L. Stéphan, A. Jay, State of the art in building modelling and energy performances prediction: a review, Renew. Sustain. Energy Rev. 23 (2013) 272–288.

4. Wall M. Distribution of solar radiation in glazed spaces and adjacent buildings. A comparison of simulation programs. Energy and Buildings 1997; 26:129–35.

5. McDowell Timothy P, Thornton Jeff W, Emmerich Steven, Walton George. Integration of airflow and energy simulation using CONTAM and TRNSYS. ASHRAE Transactions 2003;109 KC-03-10-2

6. Woloszyn Monika, Rode Carsten. Tools for performance simulation of heat, air, and moisture conditions of whole buildings. Building Simulation 2008; 1:5–24.

7. Parti Michael, Parti Cynthia. The total and appliance-specific conditional demand for electricity in the household sector. Bell Journal of Economics 1980; 11:309–21

8. Aydinalp-Koksal Merih, Ugursal V Ismet. Comparison of neural network, conditional demand analysis, and engineering approaches for modelling end-use energy consumption in the residential sector. Applied Energy 2008; 85:271–96.

9. am Joseph C, Hui Sam CM, Chan Apple LS. Regression analysis of high-rise fully air-conditioned office buildings. Energy and Buildings 1997; 26:189–97

10. Freire Roberto Z, Oliveira Gustavo HC, Mendes Nathan. Development of regression equations for predicting energy and hygrothermal performance of buildings. Energy and Buildings 2008; 40:810–20.

11. Lafrance G, Perron D. Evolution of residential electricity demand by end-use in Quebec 1979–1989: a conditional demand analysis. Energy Studies Review 1994;6(Art 4).

12. Tiedermann KH. Using conditional demand analysis to estimate residential energy use and energy savings. European Council for an Energy Efficient Economy Summer Study 2007; 6:1279–83

13. Znouda Essia, Ghrab-Morcos Nadia, Hadj-Alouane Atidel. Optimization of Mediterranean building design using genetic algorithms. Energy and Buildings 2007; 39:148–53.

14. Winkelmann FC, Birsdall BE, Bull WF, Ellington KL, Erdem AE, Hirsh JJ. DOE- 2 supplements, version 2.1e, Technical Report lbl-34947. Lawrence Berkeley National Laboratory, Berkeley, CA; 1993.

15. Nicol JF, Humphreys MA. Adaptive thermal comfort and sustainable thermal standards for buildings. Energy Build 2002;34(6):563e72.

16. Schneider S, Hollmuller P, Chambers J, Patel M. A heat demand load curve model of the Swiss national territory. In: IOP conference series: earth and environmental science; 2019. p. 1e8.

17. Ivanko D, Nord N, Sorensen AL, Sartori Plesser TSW, Walnum HT. Prediction of DHW energy use in a hotel in Norway. IOP Conf Ser Mater Sci Eng 2019; 609:1e7.

18. Gerin O, Bleys B, De Cuyper K. Seasonal variation of hot and cold-water consumption in apartment buildings. In: Proceedings of international symposium of CIB W062 water supply and drainage for buildings, Brazil; 2014. p. 1e9.

19. Pedersen L. Doctoral thesis. Load modelling of buildings in Mixed energy distribution systems. Trondheim: NTNU: Department of Energy and Process Engineering; 2007.

20. Csoknyai T, Legardeur J, Akle AA, Horvath M. Analysis of energy consumption profiles in residential buildings and impact assessment of a serious game on occupants' behaviour. Energy Build 2019; 196:1e20.

21. J. D. Balcomb, Energy Signatures: A Proposed New Design Tool, ASES '86-Annual Meeting and Annual Passive Conference Boulder, Colorado, 1986.

22. Aronsson, S. (1996); "Heat loads of buildings supplied by district heating, an analysis based on measurements of 50 buildings", doctoral dissertation, Chalmers University of Technology, Department of Building Services Engineering, Document D35:1996, ISBN 91-7197-383-4

23. Bauer M, Scartezzini JL. A simplified correlation method accounting for heating and cooling loads in energy-efficient buildings. Energy and Buildings. 1998;27(2):147-54.

24. European Commission, Directive (EU) 2018/844 of the European Parliament and of the Council, Available from: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32018L0844 (Accessed March 2023).

25. European Commission, Smart grids and meters, Available from: https://energy.ec.europa.eu/topics/markets-and-consumers/smart-grids-and-meters_en

26. Volkova A, Latosov E, Masatin V, Siirde A. Development of a user-friendly mobile app for the national level promotion of the 4th generation district heating. International Journal of Sustainable Energy Planning and Management 2019;20.

27. Bacher P, de Saint-Aubain PA, Christiansen LE, Madsen H. Non-parametric method for separating domestic hot water heating spikes and space heating. Energy Build 2016; 130:107e12.

28. Burzynski R, Crane M, Yao R, Becerra V. Space heating and hot water demand analysis of dwellings connected to district heating scheme in UK. J Cent S Univ 2012;19(6):1629e38.

29. Tereshchenko T, Ivanko D, Nord N, Sartori I. Analysis of energy signatures and planning of heating and domestic hot water energy use in buildings in Norway. E3S Web of Conferences 2019;111(6009):1e8.

30. The government's standard assessment procedure for energy rating of dwellings. In: Building research establishment, Watford, UK, BRE garston, Watford; 2019. p. 180.

31. Marszal AJ, Zhang C, Pomianowski MZ, Heiselberg PK, Gram-Hanssen K, Hansen AR. Simple methodology to estimate the mean hourly and the daily profiles of domestic hot water demand from hourly total heating readings. Energy Build 2019; 184:53e64

32. Neu O, Oxizidis S, Flynn D. Finn D utilising time of use surveys to predict domestic hot water consumption and heat demand profiles of residential building stocks. Br J Environ Clime Change 2016;6(2):77e89. 2016.

33. Verhaert I, Bleys B, Binnemans S, Janssen E. A methodology to design domestic hot water production systems based on tap patterns. In: Proceedings of CLIMA 2016: 12th REHVA world congress, Aalborg, Denmark; 2016. p. 1e10.

34. Tronchin L, Fabbri K. Energy performance building evaluation in Mediterranean countries: comparison between software simulations and operating rating simulations. Energy Build 2008;40(7):1176e87.

35. Aydinalp M, Ugursal VI, Fung AS. Modelling of the space and domestic hot water heating energy-consumption in the residential sector using neural networks. Apply Energy 2004;79(2):159e78.

36. Sorensen AL, Lindberg KB, Walnum HT, Sartori I, Aakenes UR, Andresen I. Heat analysis for energy management in neighbourhoods: case study of a large housing cooperative in Norway. IOP Conf Ser Mater Sci Eng. 2019; 609:1e7.

37. Riachi Y, Rangod N, Guillemot B. A numerical model for determining hourly heating and DHW loads in district heating systems. In: The 14th international symposium on district heating and cooling; 2014. p. 1e5.

38. Dmytro Ivanko, Åse Lekang Sørensen, Natasa Nord. Splitting measurements of the total heat demand in a hotel into domestic hot water and space heating heat use. Energy 219 (2021), 119685

39. Mikel Lumbreras, Roberto Garay-Martinez, Benat Arregi, Koldobika Martin-Escudero, Gonzalo Diarce, Margus Rau, Indrek Hagu. Data driven model for heat load prediction in buildings connected to District Heating by using smart heat meters. Energy 239 (2022) 122318

40. J. Wiehagen, J. Sikora, Domestic hot water system modelling for the design of energy efficient systems, NAHB Research Centre, Upper Marlboro, Technical report, (2002) 67

41. E. Abel, P.-E. Nilsson, L. Ekberg, P. Fahlen, L. Jagemar, R. Clark, O. Fanger, K. Fitzner, L. Gunnarsen, P.V. Nielsen, Achieving the desired indoor climate-energy efficiency aspects of system design, Lund: Studentlitteratur: Commtech Group (2003) .

42. T. Peters, J. Weyer, Architectural design for low energy housing: Experiences from two recent affordable housing projects in Denmark, in: Proceedings of Sustainable cities and buildings: 7th Passive House Norden Conference, DTU Byg and KADK Copenhagen, Denmark, 2015, pp. 1-10.

43. J. Hirvonen, J. Jokisalo, J. Heljo, R. Kosonen, Towards the EU emissions targets of 2050: optimal energy renovation measures of Finnish apartment buildings, International Journal of Sustainable Energy 38 (7) (2019) 649–672.

44. S. Attia, P. Eleftheriou, F. Xeni, R. Morlot, C. Ménézo, V. Kostopoulos, M. Betsi, I. Kalaitzoglou, L. Pagliano, M. Almeida, Overview of challenges of residential nearly Zero Energy Buildings (nZEB) in Southern Europe, in, Sustainable Buildings Design Lab, Technical Report, Liege, Belgium, 2016, pp. 87.

45. Ministero della Salute, Governo italiano. Linee guida per la prevenzione ed il controllo della legionellosi, available from: https://www.salute.gov.it/imgs/C_17_pubblicazioni_2362_allegato.pdf

46. T.-A. Koiv, A. Mikola, A. Toode, DHW design flow rates and consumption pro- files in educational, office buildings and shopping centres, Smart Grid and Renewable Energy 4 (03) (2013) 287–296

47. A. Reynaud, M. Pons, C. Pesado, Household Water Demand in Andorra: Impact of Individual Metering and Seasonality, Water 10 (3) (2018) 321.

48. H. Ahmad, Energy Efficiency Solutions for Domestic Hot Water, Online publisher: bigee.net, (2017) p. 1-5.

49. E. Wilson, C. Engebrecht-Metzger, S. Horowitz, R. Hendron, 2014 Building America house simulation protocols, in, National Renewable Energy Laboratory, Technical Report, Golden, United States, 2014, pp. 91

50. N. Good, L. Zhang, A. Navarro-Espinosa, P. Mancarella, High resolution modelling of multi-energy domestic demand profiles, Applied Energy 137 (2015) 193–210.

51. D. Fischer, T. Wolf, J. Scherer, B. Wille-Haussmann, A stochastic bottom-up model for space heating and domestic hot water load profiles for German households, Energy and Buildings 124 (2016) 120–128.

52. L.G. Swan, V.I. Ugursal, I. Beausoleil-Morrison, Occupant related household energy consumption in Canada: Estimation using a bottom-up neural-network technique, Energy and Buildings 43 (2011) 326–337.

53. Aurore Lomet, Frédéric Suard, David Chèze. Statistical modelling for real domestic hot water consumption forecasting. Energy Procedia 70 (2015) 379 – 387.

54. John C. Evarts, Lukas G. Swan. Domestic hot water consumption estimates for solar thermal system sizing. Energy and Buildings 58 (2013) 58–65.

55. R. Hendron and J. Burch. Development of Standardized Domestic Hot Water Event Schedules for Residential, Conference Paper NREL/CP-550-40874 August 2008.

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

57. Aranda A, Ferreira G, Mainar-Toledo MD, Scarpellini S, Llera Sastresa E. Multiple regression models to predict the annual energy consumption in the Spanish banking sector. Energy and Buildings. 2012; 49:380-7.

58. Lindberg KB, Doorman G. Hourly load modelling of non-residential building stock. Conference Hourly load modelling of non-residential building stock. p. 1-6.

59. Stokke, G. Doorman, T. Ericson. "An analysis of demand charge electricity grid tariff in the residential sector." Energy Efficiency, vol. 3, pp. 267-282, Jan. 2010.

60. NIST Standard Reference Database 69: NIST Chemistry WebBook, Thermophysical Properties of Fluid Systems, Isobaric Properties for Water, 2008.

61. Kaiguang Zhaoa, Michael A. Wulder, Tongxi Hu, Ryan Bright, Qiusheng Wu, Haiming Qin, Yang Li, Elizabeth Toman, Bani Mallick, Xuesong Zhang, Molly Brown. Detecting change-point, trend, and seasonality in satellite time series data to track abrupt changes and nonlinear dynamics: A Bayesian ensemble algorithm. Remote Sensing of Environment 232 (2019) 111181.

62. Altman DG, Gardner MJJ Bmj. Statistics in Medicine: Calculating confidence intervals for regression and correlation. 1988;296(6631):1238

63. K. Winqvist (Ed.), How Europeans Spend Their Time Everyday Life of Women and Men, Office for Official Publications of the European Communities, Luxembourg, 2004.

64. National Institute of Standards and Technology (NIST), Critical Values of the Student's t Distribution. Available at: https://www.itl.nist.gov/div898/handbook/eda/section3/eda3672.htm

65. Statistics Online Computational Resources (SOCR). F distribution tables. Available at: https://socr.umich.edu/Applets/F_Table.html

66. V.F. Nakhodov, A.I. Zamulko, M. Al Sharari, D.O. Medintseva, Analysis of duration and border of existing tariff zones, Power engineering: Economics, Technology, Ecology 2 (2016) 88–97.

67. Vlatko Milić & Patrik Rohdin (2023): Screening of thermal characteristics and assessment of comparative energy efficiency potential in a residential district, Advances in Building Energy Research, DOI: 10.1080/17512549.2023.2183522

68. Site of municipality of Verona: https://www.comune.verona.it/nqcontent.cfm?a_id=1

69. Kamstrup, https://www.kamstrup.com/en-us/heat-solutions/heatmeters/meters/multical-603 (Accessed in March 2023).

70. EN 1434-1:2015, Heat meters. Part 1: general requirements.

71. University of Tartu, Institute of Physics, Laboratory of Environmental Physics, http://meteo.physic.ut.ee/?lang¼en (Accessed in 2019)

72. Li Aihua, Feng Mengyan, Li Yanruyu, Liu Zhidong. Application of outlier mining in insider identification based on boxplot method. Procedia Computer Science 2016; 91:245e51. https://doi.org/10.1016/j.procs.2016.07.069. ISSN1877-0509.

73. Schwertman Neil C, Ann Owens Margaret, Robiah Adnan. A simple more general boxplot method for identifying outliers. Comput Stat Data Anal2004;47(Issue 1):165e74. https://doi.org/10.1016/j.csda.2003.10.012. ISSN0167-9473.

74. Ente italiano di normazione UNI. UNI 10349-1:2016 Riscaldamento e raffrescamento degli edifici - Dati climatici - Parte 1: Medie mensili per la valutazione della prestazione termo-energetica dell'edificio e metodi per ripartire l'irradianza solare nella frazione diretta e diffusa. 2016.

75. Martin Eriksson, Jan Akander, Bahram Moshfegh. Development and validation of energy signature method – Case study on a multi-family building in Sweden before and after deep renovation. Energy & Buildings 210 (2020) 109756.

76. Wahlqvist M., 2017: Modellering av energianvändning i flerbostadshus. Enstudie över möjligheter och begränsningar med olika sätt att beräkna energisignatur, examensarbete UU (in Swedish).

77. Department of Education of UK government, School Output Specification, Technical Annex 2H: Energy. November 2022. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1119606/GDB_Annex_2H-Energy-A-C13.pdf

78. Thonipara A, Runst P, Ochsner C, Bizer K. Energy efficiency of residential buildings in the European Union-An exploratory analysis of cross-country consumption patterns. Energy Pol 2019; 129:1156e67.

79. European Commission, 2018, Energy Performance of Buildings Directive, Available from: https://energy.ec.europa.eu/topics/energy-efficiency/energy-efficient-buildings/energy-performance-buildings-directive_en (visited in March 2023)

80. Eurostat, 2018. Energy Consumption in Households

81. European Commission. (2019). Clean Energy for all Europeans. Available from: https://ec.europa.eu/energy/topics/energy-strategy/clean-energy-all-europeans_en (visited in March 2023)

82. European Parliament. (2021). Energy efficiency in buildings. Available from: https://www.europarl.europa.eu/factsheets/en/sheet/146/energy-efficiency-in-buildings (visited in March 2023)