POLITECNICO DI TORINO

PhD STUDENT:
Marco Savino Piscitelli

TUTORS:
Prof. Alfonso Capozzoli, Prof. Marco Perino

Energy management in buildings through data analytics technologies

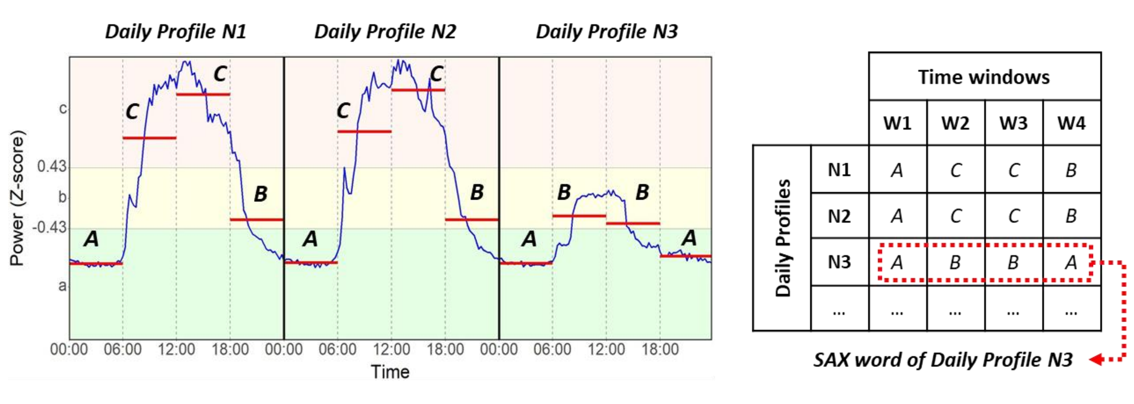
1. Capozzoli A, Piscitelli M S, Brandi S., Grassi D., Chicco G. *Automated load pattern learning and anomaly detection for enhancing energy management in smart buildings*. Energy 2018; 157: 336-352

Scientific background

In the context of smart buildings in smart cities, the growing spread of ICT (Information and Communication Technologies) and IOT (Internet Of Things) technologies is enabling the collection of a huge amount of building related-data. Consequently, the energy management in buildings is becoming more and more a data-centric task also considering the effectiveness of novel data analytics technologies in driving the development of ready-to-implement energy conservation measures in such a complex systems. In this framework, my research activity is aimed at exploiting the way in which a data analytics based approach is changing the paradigm of energy management leading to a significant reduction of energy consumption and energy wastes during the building operation. Even though some applications are well developed in the literature (e.g., Fault detection and diagnosis), the diversity of data analytics techniques and their combination still remain challenging to be handled in the building physics sector. The effective coupling of building physics and data science needs significant contributions aimed at exploring robust and generalizable analytical frameworks.

Temporal data mining in buildings

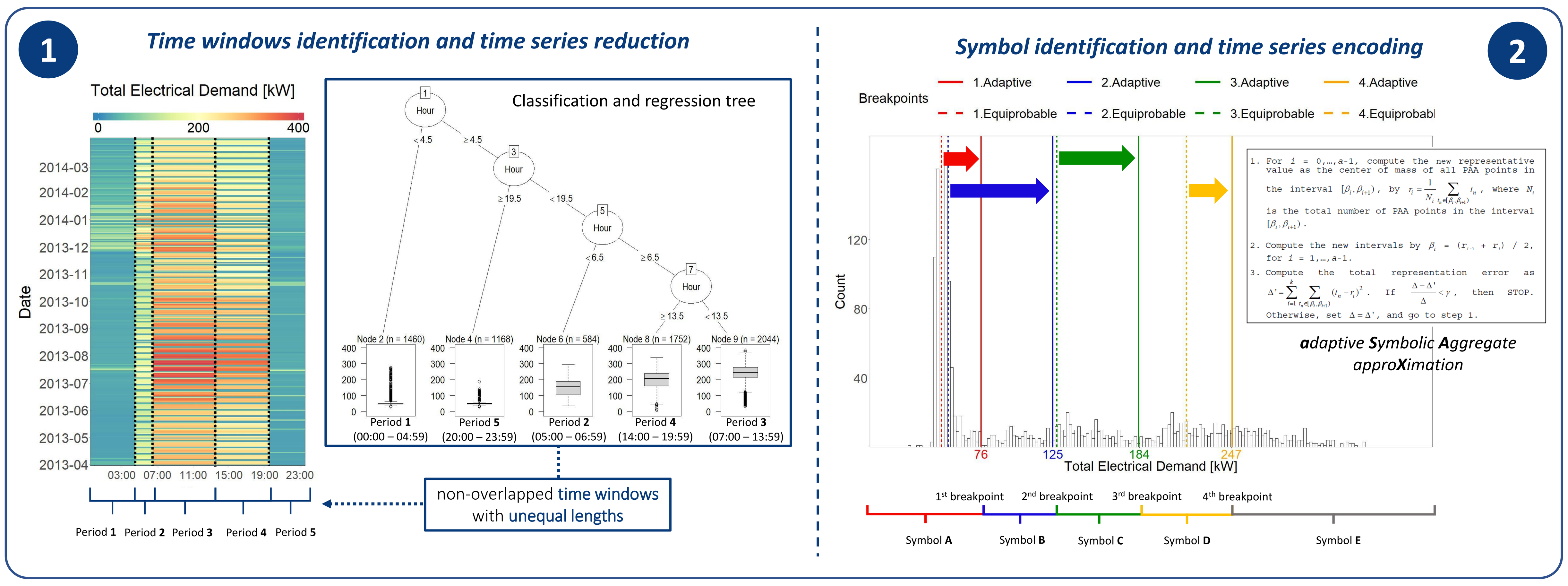
In last years, particular attention has been devoted to the branch of time-series analytics for describing and modeling energy dynamics in buildings and systems during their real operation. This kind of analyses, based on the ensemble of data mining and machine learning algorithms, are of a paramount importance in harvesting valuable information from data preserving their chronological relation. In fact, the temporal discovery process makes it possible to extract useful knowledge in the time domain on the actual relations between outdoor disturbances, indoor conditions and energy demand.
The great advantage of this data driven approach is related to its potential capability in learning and characterizing building energy dynamics features for unforeseen systems without a-priori knowledge of their configuration.
Algorithms related to (i) Significant and recurrent Pattern Mining (ii) Causality Analysis (iii) Time series similarity proved to be flexible in their combination and effective in addressing emerging issues in building energy management such as energy demand prediction, Fault Detection and Diagnosis, anomalous energy trend recognition.
Most of these techniques rely on temporal abstraction as a preprocessing stage for knowledge extraction. Temporal abstraction consists in transforming time series from numerical sequences to discrete state sequences by means of the reduction and the transformation of data.



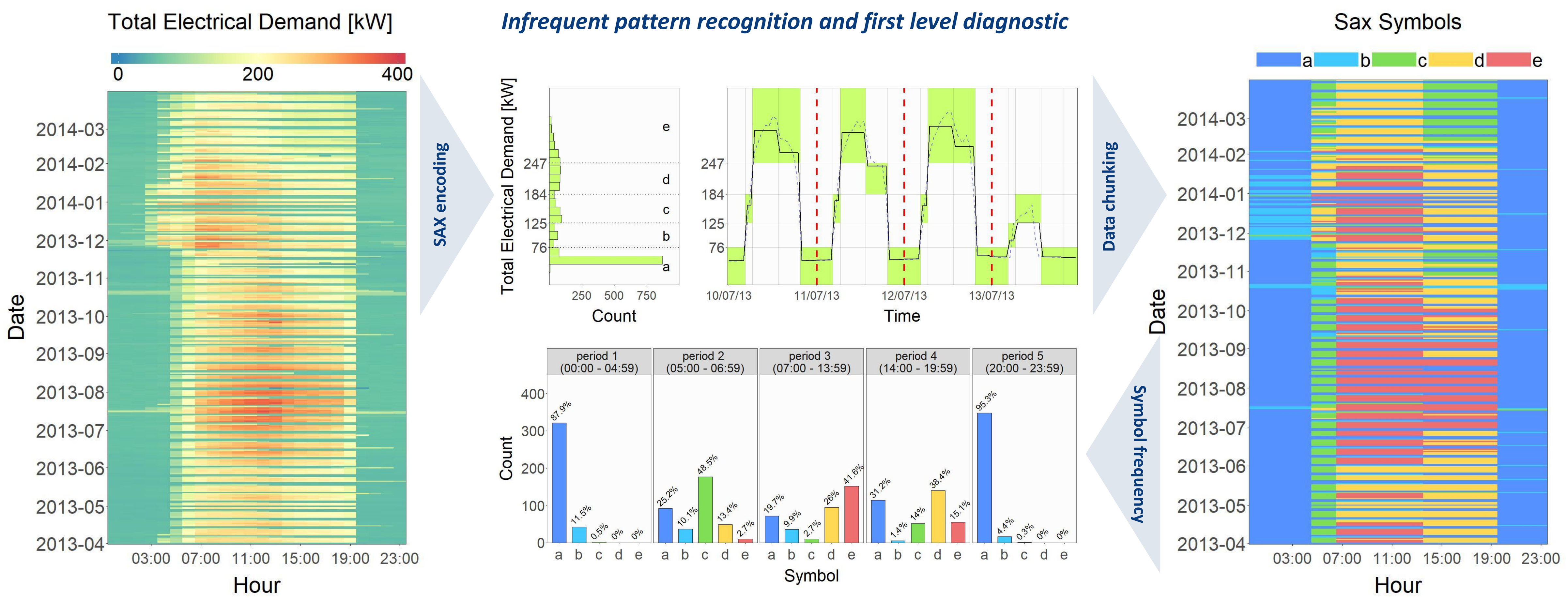
Symbolic Aggregate approXimation (i.e., SAX) is one of most promising technique available to discretize a time series, without losing temporal key information. It is based on the reduction of the time series through a piecewise technique and on its transformation into a symbolic string. The symbols are associated to discrete states of the time series in non-overlapped time windows of equal length.
The subsequent step consists of chunking the entire string into a set of N symbolic sub-strings, called SAX words. Those words represent then the abstraction of the sequence of events during a reference period.

Anomalous energy trend detection in buildings[1]: methodology

The first stage of the analysis is aimed at transforming the energy consumption time series by implementing an enhanced SAX process. In detail, two preliminary hypotheses are formulated in different ways from the classic SAX implementation:
(i) unequal time window lengths can be identified to conduct a finer discretization, (ii) rejection of equal probability of the symbols in order to encode each approximated constant segment on the vertical axis.



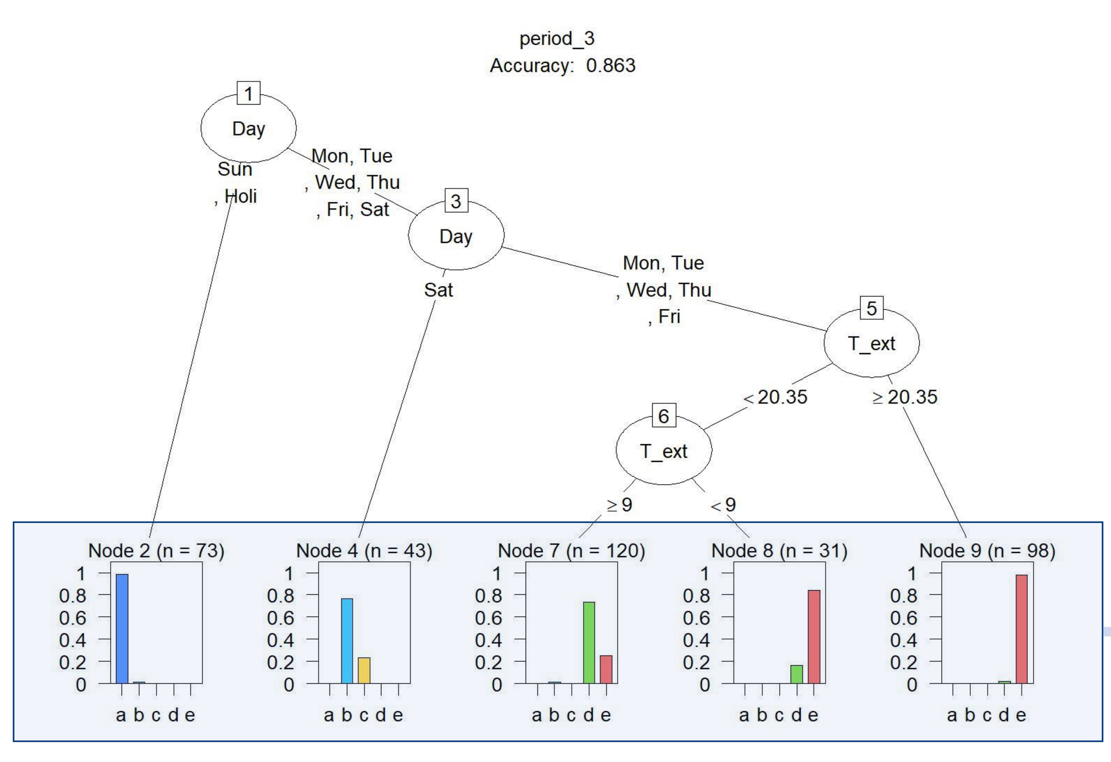
After the data transformation, the entire time series encoded in a unique string is chunked into N sub-strings of a daily length (i.e., T = 24 hours) in order to obtain constant time-scale based sequences. The N symbolic sub-strings are made up of a certain number, W, of time windows encoded in alphabetic symbols and organized in a NxW matrix. In this way, each daily load profile is represented by a SAX word that is then used as the input for the successive anomaly detection analysis.



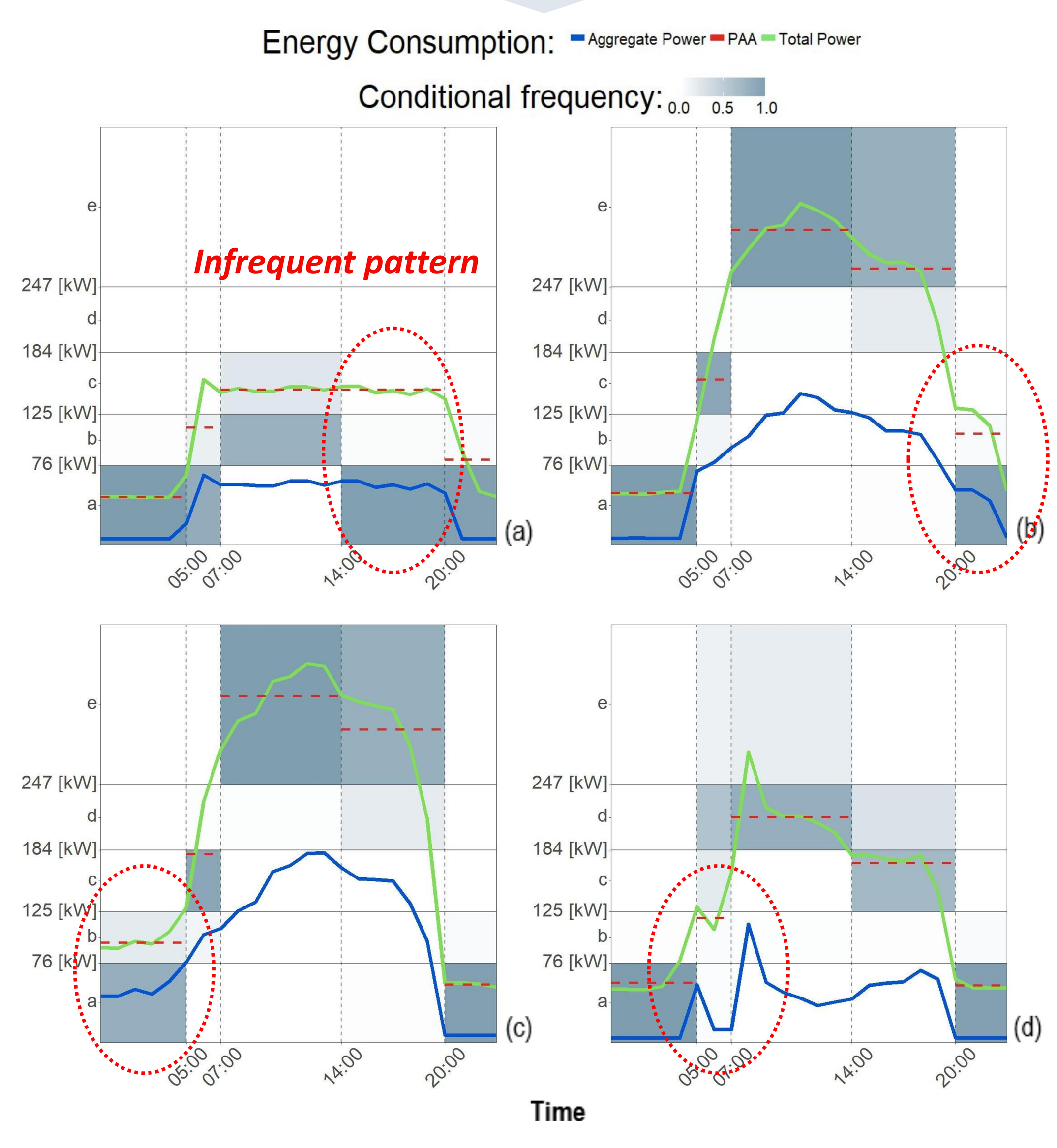
At this stage, the probability of each symbol occurring in each time window, under specific boundary conditions, is evaluated by means of a classification tree, which is based on additional explanatory variables (e.g., external temperature, internal temperature, day type, month). In this way, if the occurrence probability estimated with the classification tree and associated with a symbol is very low, it is likely that the energy consumption in the corresponding sub-daily time window is abnormal. Furthermore, the post-mining stage of the analysis is performed using additional datasets in order to further support the preliminary diagnosis of detected anomalies.
In this perspective, the proposed methodology represents a very useful tool that can be used to support the implementation of advanced targeted anomaly diagnosis in a specific time window of the entire time domain.

Time window	Decision rules	Symbol	Accuracy
Period 1 (00:00 – 04:59)	IF ***system_start*** = is turned OFF	→ ***a***	*98%*
	IF ***system_start*** = is turned ON at 04:00 a.m. AND ***T_int*** ≥ 23,43 °C	→ ***a***	*80%*
	IF ***system_start*** = is turned ON at 04:00 a.m. AND ***T_int*** < 23,43 °C	→ ***b***	*79%*
Period 2 (05:00 – 06:59)	IF ***Day*** = Holiday OR Sunday OR Saturday	→ ***a***	*83%*
	IF ***Day*** = Monday OR Tuesday OR Wednesday OR Thursday OR Friday AND ***T_int_pre (period 1)*** ≥ 23,55 °C	→ ***c***	*88%*
	IF ***Day*** = Monday OR Tuesday OR Wednesday OR Thursday OR Friday AND ***T_int_pre (period 1)*** < 23,55 °C	→ ***d***	*60%*
Period 3 (07:00 – 13:59)	IF ***Day*** = Holiday OR Sunday	→ ***a***	*99%*
	IF ***Day*** = Saturday	→ ***b***	*77%*
	IF ***Day*** = Monday OR Tuesday OR Wednesday OR Thursday OR Friday AND 9 °C ≤ ***T_ext*** < 20,35 °C	→ ***d***	*73%*
	IF ***Day*** = Monday OR Tuesday OR Wednesday OR Thursday OR Friday AND ***T_ext*** ≥ 20,35 °C	→ ***e***	*98%*
	IF ***Day*** = Monday OR Tuesday OR Wednesday OR Thursday OR Friday AND ***T_ext*** < 9 °C	→ ***e***	*84%*
Period 4 (14:00 – 19:59)	IF ***Sym_pre*** = a OR b OR c	→ ***a***	*96%*
	IF ***T_ext*** < 24,1 °C AND ***Sym_pre (period 3)*** = "*d*" AND ***T_int*** < 25,55 °C	→ ***c***	*69%*
	IF ***T_ext*** < 24,1 °C AND ***Sym_pre (period 3)*** = "*d*" AND ***T_int*** ≥ 25,55 °C	→ ***d***	*75%*
	IF ***T_ext*** < 24,1 °C AND ***Sym_pre (period 3)*** = "*e*"	→ ***d***	*94%*
	IF ***Sym_pre*** = "*d*" OR "*e*" AND ***T_ext (period 3)*** ≥ 24,1 °C	→ ***e***	*79%*
Period 5 (20:00 – 23:59)	-	→ ***a***	*95%*

Infrequent pattern recognition and first level diagnostic



Decision rules



Energetics PhD XXXII Cycle

PhD Day
Energy Center Auditorium
December 19th 2018

Phd student:
Marco Savino Piscitelli

Tutors:
Prof. Alfonso Capozzoli
Prof. Marco Perino

BAEDA Lab — Building Automation Energy Data Analytics

Publications

Capozzoli A, **Piscitelli M S**, Brandi S., Grassi D., Chicco G. ***Automated load pattern learning and anomaly detection for enhancing energy management in smart buildings***. Energy 2018; 157: 336-352
Capozzoli A, **Piscitelli M S**, Brandi S. ***Mining typical load profiles in buildings to support energy management in the smart city context***. Energy Procedia 2017; 134: 865–874.
Capozzoli A, **Piscitelli M S**, Gorrino A, Ballarini I, Corrado V. ***Data analytics for occupancy pattern learning to reduce the energy consumption of HVAC systems in office buildings.*** Sustain Cities Soc 2017; 35: 191–208.
Di Corso E, Cerquitelli T, **Piscitelli M S**, Capozzoli A. ***Exploring energy certificates of buildings through unsupervised data mining techniques***. IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), Exeter, UK, 21-23 June 2017. pp. 991-998
Capozzoli A, Serale G, **Piscitelli M S**, Grassi D. ***Data mining for energy investigation of large data set of flats***. (Proc Inst Civ Eng) Engineering Sustain 2017; 170:3-18.
Capozzoli A, **Piscitelli M S**, Neri F, Grassi D, Serale G. **A novel methodology for energy performance benchmarking of buildings by means of Linear Mixed Effect Model: The case of space and DHW heating of out-patient Healthcare Centres**. Appl Energy 2016; 171:592-607.
Capozzoli A, Cerquitelli T, **Piscitelli M S**. **Enhancing energy efficiency in buildings through innovative data analytics technologies** (book chapter). Pervasive Computing: Next Generation Platforms for Intelligent Data Collection. Elsevier 2016. Editors: Ciprian Dobre, Fatos Xhafa.
Capozzoli A, Grassi D, **Piscitelli M S**, Serale G. **Discovering Knowledge from a Residential Building Stock through Data Mining Analysis for Engineering Sustainability**. Energy Procedia 2015;83:370-9.

Collaborations

Department of Building Services Engineering
The Hong Kong Polytechnic University
ENEA
iren