Testing Socio-Economic Demographic Variables on building energy consumption scenarios at the urban scale in Italy

ID 19 - Testing Socio-Economic Demographic Variables on building energy consumption scenarios at the urban scale in Italy

Simona D'Oca[1], Chiara Delmastro[1], Valentina Fabi[1], Stefano Paolo Corgnati[1]

DENERG (Dipartimento di ENERgia), Politecnico di Torino

Corresponding e-mail: simona.doca@polito.it

SUMMARY

In order to meet EU 2020 energy efficiency aggressive goals, there is a need to scale up the achievement of real energy consumption savings of current high performing building to the urban level. However, current low-energy, zero carbon, and net zero energy buildings are not performing as designed. Within the building sector industry there is an increasing concern about the mismatch between expected and actual performance, typically addressed as the 'credibility gap'. The reasons for this discrepancy vary, but are mostly rooted in oversimplifying or ignoring occupant behavior in the building operation process. In this view, the presented study is testing the magnitude of socio-economic demographic variables on building energy consumption energy scenarios at the urban scale. Measured thermal energy consumption data for a block of building in Torino, Italy are coupled with Geographical System Information (GIS) public census socio-economic data related to building occupants and analyzed through data mining techniques. Hence, energy model scenarios will be drawn to represent the impact of different building occupant profiles over mid-long term building energy consumption at the urban scale.
Interpretation of the energy scenarios may allow energy designers and modelers, building operator and manager to develop energy efficiency measures and standards taking into account the leverage of socio-economic factors related to building occupants. Results may also support energy and urban planners to the release of energy policies and the development of robust energy urban planning tools aiming to bridge the "credibility gap" of targeted energy efficiency in building.

INTRODUCTION

Building occupants interact with the indoor environments in purposive and significant ways that contribute to both energy consumption and Indoor Environmental Quality (IEQ), and thus warrant significant attention in the building design and operation processes. For example, occupants' thermally adaptive behaviors (i.e. turning on fans/heaters, opening windows) are strongly tied to space heating and cooling loads, which make up 27% and 13% of global primary energy consumed in residential and office buildings in the Europe, respectively (Energy Efficiency trends in buildings in the EU, Enerdata). These behaviors also modify key thermal comfort determinants like air temperature, air speed and clothing insulation level [1]. Recent studies have begun to quantify the magnitude of occupant behavior's influence on energy use and comfort, reporting significant impacts that have intensified the focus on behavior as a key topic of built environment research [2, 3].
If the general importance of the human-building interaction is well established, however, the mechanisms behind this interaction are still being explored. Increasingly, this effort has involved the collection of longitudinal data, which allow one to observe occupant comfort and adaptive behavior as they evolve together across the day and season. Nevertheless, longitudinal studies are time-consuming and expensive to carry out, and existing comfort and behavior data are accordingly limited in their coverage of certain adaptive actions, building types and climates. In recent years, data mining has been used more and more in the building science area, to investigate occupant behavior patterns or to analyze building automation systems and building energy performance. In this context, data mining is a useful technique to elaborate huge samples of data extrapolating significant information. Data mining techniques lead the way to

automatically analyzing huge amounts of data. They can be used to extract interesting, useful, and previously unknown knowledge from data. The increasing usage and technology development of IT, large amounts of data will be available in the building sector. This presents an excellent opportunity for data mining applications to discover new science. The research has demonstrated the data mining method is capable of predicting occupancy patterns [4] and operations of both window opening and heating set point adjustment [5, 6] for energy scheduling purpose. Even if lot of work is done to better represent the occupant behavior patterns in building science research, further fields of studies still remain. For example, with only few exceptions, the role of occupant behavior at urban level is not yet investigated.
Sustainable urban development requires a long-term sector-integrative approach. Scenario-based analysis with respect to development paths determines the energy consumption of different city occupancy configurations. Long-term strategies are crucial especially with regard to the development of key projects, sites and locations, as the short-term realization of supposedly appropriate projects on specific locations might prohibit the future viability of sustainable projects in these locations.
This paper proposes a method of system analysis and partial simulation for urban structures for this purpose. The aim of this paper is to provide a methodology to develop strategies for complex situations while sustainable planning of urban settlements, based on modeling of urban energy and accounting for occupants' behavior defined by cluster analysis of an urban neighborhoods. The main purpose of this development is to derive a method to set up a systems model that supports decision processes in planning and designing the built environment and is tailored specifically for this purpose.

METHODS

Assessment of the reference blocks of buildings

The analysis is focused on 21 buildings sited in the municipality of Turin, Italy, that is located in the climatic E zone (temperate continental) (Fig.1). In this session, reference buildings characteristics and thermal performances are assessed in order to be able to investigate the refurbishment potential of the area.

Figure 1 – Reference Block of Buildings

Thanks to the availability of statistical information and supported by a Geographic Information Systems (GIS), physical characteristics have been associated to each building. It results that the stock is characterized by two construction periods – from 1971 to 1980 and 1881-1990 - and by three building types – multi family apartment buildings, towers and low rise buildings. According to the analysis performed by [7], space heating energy consumption of the 21 buildings has been evaluated by matching real consumption data provided by local utilities and the results of the thermal model described in [8]. This thermal model was built by analyzing 300 residential buildings and it depends on degree-days, on the envelope's thermal insulation (related to the period of construction) and on the shape factors of buildings. Depending on the relative difference between real consumption data and the thermal model results it is possible to evaluate if some buildings have been already refurbished.

Table 1. Assessment of building space heating energy consumption [7]

Building ID	Calc. consumptions ($kWh/m^2/y$)	Real consumption ($kWh/m^2/y$)
1	111.67	116.88
2	113.69	110.96
3	118.88	107.84
4	118.88	114.66
5	119.02	128.3
6	118.58	116.99
7	116.73	130.08
8	115.73	110.05
9	115.58	113.73
10	115.63	109.5
11	145.04	162.79
12	115.11	115.92
13	114.26	117.58
14	113.56	119.61
15	113.71	119.83
16	118.94	108.65
17	108.47	113.86
18	109.46	111.68
19	116.45	112.32
20	118.63	114.98
21	169.11	189.82

From Table 1 results that none of the buildings have been previously refurbished and that energy consumption are comprised in a range varying from about 105 and 190 $kWh/m^2/y$. For allowing the analysis proposed in the next sessions, buildings have been grouped into four main reference building blocks (RBBs accordingly to their census area). It results:

- RBB-A: census area 2745, composed by 2 buildings (ID: 1 and 2);
- RBB-B: census area 2746, composed by 9 buildings (ID: 3, 4, 5, 6, 12, 13, 14, 15 and 20);
- RBB-C: census area 2999, composed by 8 buildings (ID: 7, 8, 9, 10, 16, 17, 18 and 19);
- RBB-D: census area 3698, composed by 2 buildings (ID: 11 and 21).

Assessment of the Socio-Economic Demographic Variables

A preliminary methodological approach for including behavioral variables into the energy service demand analysis at an urban/district scale was introduced by Delmastro et al, 2015 [8]. Similarly, this study is incorporating socio-economic behavioral variables into energy model scenarios representing the impact of archetypal building occupant profiles over mid-long term building energy consumption at the district level. Socio- economic conditions of inhabitants such as education level, income and employment, as well as demographic variables, such as household composition and dwelling ownership affect their lifestyle and consequently the energy demand at the household level [9]. In Italy, socio-economic data related to population and building census are provided by the National Institute of Statistics [10]. Programs such as the IEA – Annex 66 "Definition and Simulation of Occupant Behavior in Buildings" [11] – have initiated advancements in the standardization of the quantitative descriptions and classification of archetypal occupant behavior profiles on building performance. Specifically Sub-Task C is investigating the most effective methodologies to represent population diversity – e.g., clustering, combining all data to make a model, combining data just from each occupant to make a model, etc. – when implanting the behavioral variables in building energy performance simulations and forecasts. Aim of the project is to address the simulation user towards

the choice of the best fit-for-purpose user type diversity profiles (stochastic profiles, agent/action-based model, standard schedules, archetypal user profiles) to be applied to different levels of modeling applications (building, block of building), aims of simulation (design, optimization, forecast, energy planning), building typologies (destination of use) as well as building related factors (envelope, location, etc.).
Considering the case of energy planning at the urban context, diversity profiles must be general enough to be valid and applicable to the huge variety of energy-related behaviors and user socio-demographic characteristics. In this view, archetypal user profiles should be particularly useful in areas in which statistics on census data is quite detailed. For instance, the city of Turin, Italy, allows via a statistic database (DB) to know extrapolate information on the socio-demographic variables for each census area. This permits to identify, for each of this census area, the probability distribution of the archetypes user profiles, which could be associated with the RBC distribution.
The KDD (Knowledge Discovery in Database) methodology [12] is employed in order to extrapolate valid, potential useful and understandable knowledge from the information related to population and building census of an urban district located in Torino, Italy.

RESULTS

Clustering archetypal user profiles based on Socio-Economic Demographic Variables

In this study, data on socio-economic and demographic variables such as education level, employment, ownership type and age (Table 2) are clustered by means a k-means algorithm, in order to find homogeneous archetypal user profiles among the four reference blocks of buildings (here named RBB-A, RBB-B, RBB-C, RBB-D).

Table 2. Socio-economic and demographic variables such as education level, employment, ownership type and age provided by the [10] regarding census statistical data collected among 2012 and 2013 over the population of four census areas corresponding to the reference blocks of buildings

Education level	**Employment level**	**Ownership type**	**Population Age**
Bachelor/master degree	Employed	Private dwellings	0-24 years
High school diploma	Unemployed	Rented dwellings	25-49 years
Middle school certificate			50-69 years
Primary school certificate			>74 years

Cluster analysis is the process of merging data into different clusters, so that instances in the same cluster have high similarity and instances in different clusters have low similarity [13]. The similarity between clusters was computed based on the k-means algorithm, given the simple but powerful nature of the algorithm, it is one of the widely used classification technique.
The performance of the cluster models was assessed by means the Davies–Bouldin index. The k=n algorithm that produced clusters with low intra-cluster distances (high intra-cluster similarity) and high inter-cluster distances (low inter-cluster similarity) had a low Davies–Bouldin index, and was considered the $k=n_{opt}$ cluster algorithm for the specific data set. An example of the cluster analysis performance index assessment is provided in Table 3, whereas a number of $3<k<8$ clusters of "population Age" among the census data was computed. A k=4 number of clusters was selected to assure the highest accuracy of the algorithm as follow: 1) 0-24 years; 2) 25-49 years; 3) 50-69 years; and 4) >74 years.

Table 3. Performance evaluation of the k-means algorithm for the "Population Age" Cluster Distribution, by using the Davis Bouldin Index

k- number of clusters	3	**4**	5	6	7	8

Davis Bouldin Index	-1.038	**-1.136**	-0.792	-0.604	-0.525	-0.415

Results of the cluster analysis showed a homogeneous distribution of the "Population Age" socio-demographic clusters, the four reference blocks of buildings. Specifically, the analyzed district emerged mainly populated by users aged 50-69 years old (Figure 2).
A consistent pattern also emerged among the socio-demographic clusters for "Education Level" and "Employment Level", while a wide dispersion arose with reference to the "Ownership type" (private owned or rented) of the dwellings located in the reference blocks of buildings (Figure 3).

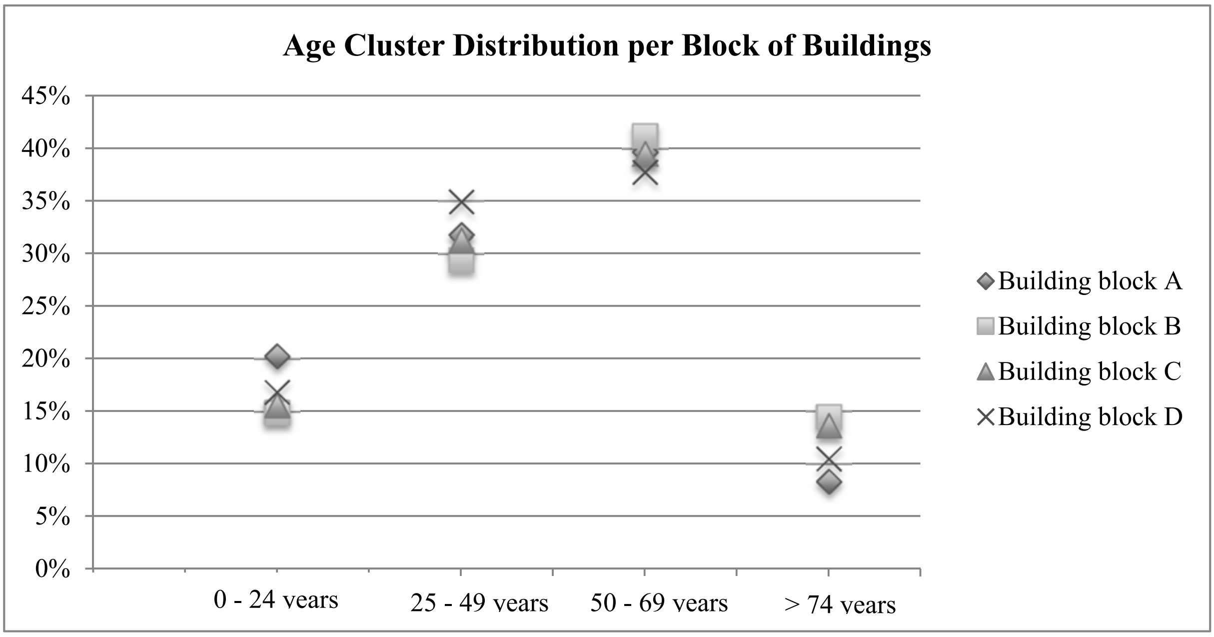

Figure 2. Age Cluster Distribution for each of the four case studies blocks of buildings

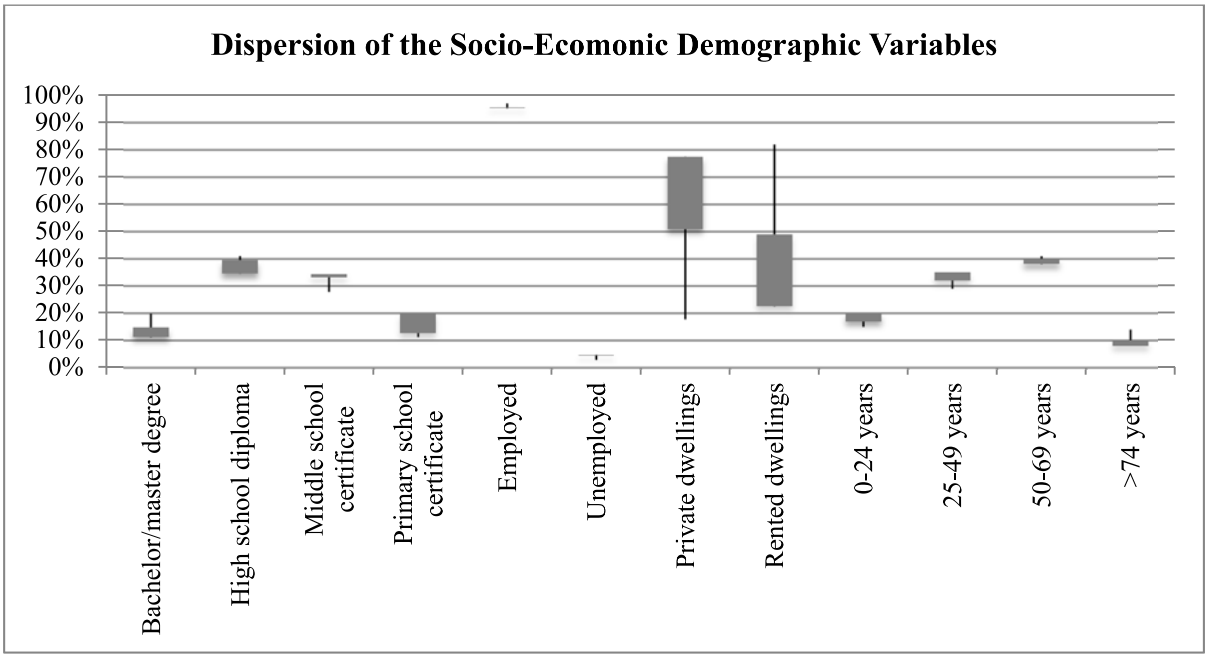

Figure 3. Dispersion of the socio-economic demographic variables regarding: education level, employment, ownership type and age among the 4 blocks of buildings.

Figure 4 illustrates the socio-economic demographic variables characterizing the archetypal user profiles emerged among the 4 blocks of buildings. Residents of the 4 blocks of buildings are pooled according to a similar and homogeneous: education level - high school diploma -, employment level

– occupied - and population age – from 50 to 69 years old. Hence, 3 different scenarios of ownership type – prevalence private owned dwellings, prevalence rented dwellings, balanced private owned/rented dwellings – are clustered respectively to the Building Block A, Building Block B, C and the Building Block D.

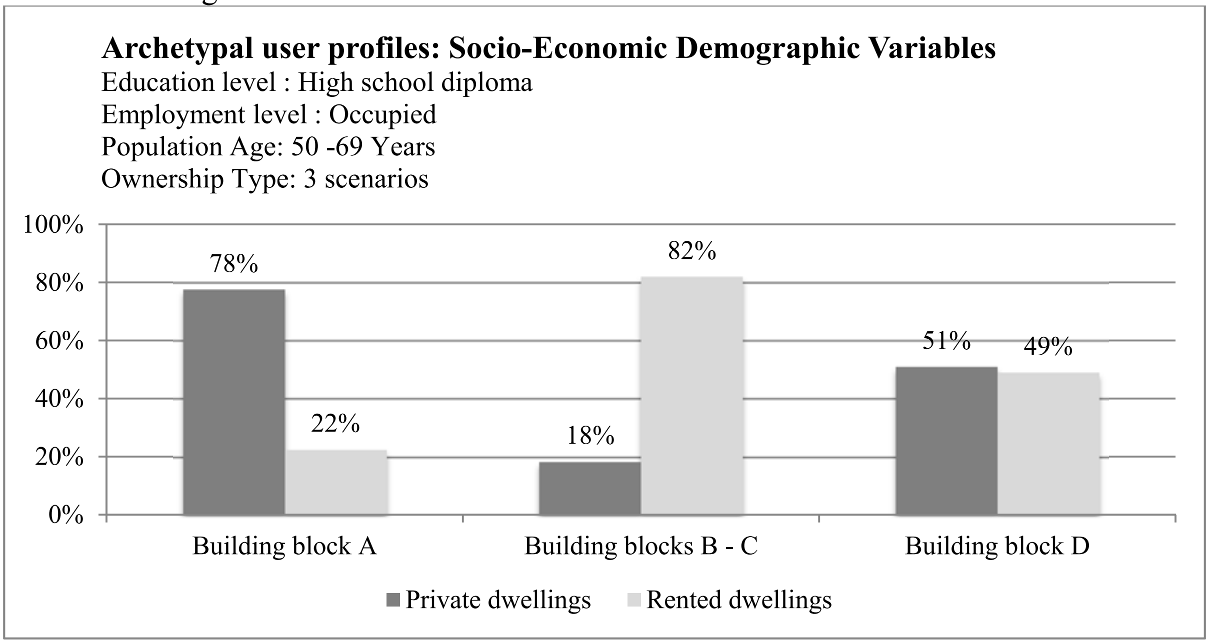

Figure 4. Socio-economic demographic variables characterizing the archetypal user profiles emerged among the 4 blocks of buildings.

Evaluation of the impact of the Socio-Economic Demographic Variables on the feasibility of retrofit measures at the urban scale

The goal of this session is to evaluate the renovation potential on the considered district. In a first step the global potential is evaluated (considering all the buildings to be refurbished), while in a second step the impact of cluster characteristics on the feasibility potential is quantified. As evaluated in the previous session, most the chosen socio-demographic variables (education level, employment level and population age) are homogeneously distributed within the district; moreover, their level do not impact negatively on the refurbishment potential since most of the population has a work occupancy, a medium education level and the age range is not to high to avoid the possibility to invest. The only variable that differs meaningfully from a census area to an other is the ownership type: in RBB-A only the 22% of the population lives in rented dwelling, in RBB-B and RBB-C most of the population (more than 80%) lives in rented dwelling while in RBB-D the distribution is balanced. This consideration has a weight in the feasibility of renovation works: the probability that population living in rented dwellings will invest in energy savings measures is very low. For that reason and considering that rental occupants are not all concentrated in the same buildings, in this preliminary approach it has been defined a feasibility index called “property factor (PF)” and dependent from users cluster’s that differs from a building block to another:

- PF-A = 1; the property factor of RBB-A is equal to one, all the inhabitants will refurbish their dwellings;
- PF-BC = 0.12; the property factor of RBB-B and RBB-C is equal to 0.12 that means that inhabitants have a very low probability to refurbish their buildings (only 2 on a total of 17);
- PF-D= 0.5; the property factor of RBB-D is equal to 0.5 that means that 1 building on a total of 2 has a very high probability to be refurbished;

Table 4. Energy savings potential of the buildings blocks

ID	RBB	Consumption (kWh/m2/y)	Global Refurbishment (kWh/m2/y)
1	A	116.88	85.55
2	A	110.96	81.22
3	C	107.84	55.08
4	C	114.66	57.21
5	C	128.30	55.36
6	C	116.99	65.43
7	B	130.08	64.53
8	B	110.05	57.84
9	B	113.73	54.25
10	B	109.50	58.85
11	D	162.79	57.67
12	C	115.92	54.65
13	C	117.58	57.27
14	C	119.61	56.18
15	C	119.83	56.50
16	B	108.65	58.31
17	B	113.86	59.14
18	B	111.68	60.16
19	B	112.32	60.27
20	C	114.98	119.2
21	D	189.82	138.9

In Table 4 the global energy savings potential (considering all the buildings to be refurbished) has been evaluated. The considered interventions are the following: replacement of all windows for the lower classes, thermal insulation of the roof and the floor below the building, insulation of the walls. The energy savings resulting from renovation works have been quantified by considering the existing literature [7-9, 15].

Table 5. Impact of the cluster analysis on total energy savings potential

RB	**PF**	**Energy consumption (MWh/y)**			**Δconsumption (kWh/y)**		**Clusters' impact**
		Existing situation (E)	*Global retrofit (R)*	*Retrofit with clusters (RC)*	*(E-R)*	*(E-RC)*	
A	1	1858	1360	1360	498	498	0%
B and C	0.12	15,822	7959	14,753	7864	1069	86%
D	0.5	878	643	677	235	201	15%
Total		18,558	9962	16,790	8597	1768	79%

From Table 5 it is possible to observe that the cluster analysis impacts for a total of +79% on the energy savings potential deriving from the retrofit of the 21 district's buildings. Of course, as for buildings B and C, in the areas characterized by low feasibility index the cluster's analysis impact is very high. This result highlights the importance of including archetypal socio-economic users profiles in the estimation of the feasibility and penetration of energy policy and measures. In particular, if scaled at higher scale (f.i the urban scale) it can support decision making in policy formulation by catching spatial differences in social conditions of citizens; it can allow the promotion of tailored policies from zone to zone and the estimation of not only the potential of an intervention, but the real attended perception of the policy by investors.

DISCUSSION AND CONCLUSION

In this paper, a method to develop energy consumption strategies at urban level considering occupancy features is proposed and applied to a case study. Cluster analysis of socio-demographic variables for a neighborhood in Turin – Italy – highlighted three main users' typologies. In particular, education level (high school diploma), employment level (occupied) and population age (50-69 years) resulted to be homogenous. As a matter of fact, a report of the national Census given by [10] indicates the average population age in Italy is 43 years. Moreover, it confirmed that the progressive aging of Italian population is evident through the analysis of the comparison between the number of old people (65 years and older) and children under 6 years (for every child there are more than 3 elderly). Furthermore, ISTAT [10] counts more the 64% of resident people in North Italy having an employment. A greater variation emerged instead among house ownership conditions. Based on the resulted classification of Archetypal user profiles, three main scenarios of renovation potential were evaluated. From the results, it emerged the cluster analysis have an high impact on the energy saving potential, highlighting the importance of considering Archetypal user profiles in the estimation of penetration of refurbishment actions. Further, having a more heterogeneous dataset of a larger area data could demonstrate a higher deviation representing the variability on energy consumption due to the occupancy features.

REFERENCES

1. Baker, N and Standeven, M. "Thermal Comfort in Free-Running Buildings" in Energy and Buildings, 23, 1996, pp. 175-182.
2. Hong, T and Lin, G. Occupant Behavior : Impact on Energy Use of Private Offices. Proceedings of Asim IBSPA Asia Conference; 2012.
3. Bourgeois, D, Reinhart, C, Macdonald, I A. Assessing the total energy impact of occupant behavioural response to manual and automated lighting systems. Ninth International IBPSA Conference, Montréal, Canada; 2005.
4. D'Oca, S and Hong, T. Occupancy schedules learning process through a data mining framework. Energy and Building 88, 2015, pp. 395-408.
5. D'Oca, S and Hong, T. A data-mining approach to discover patterns of window opening and closing behavior in offices. Building and Environment 88, 2014, pp. 726-739.
6. Ren, X, Yan, D, Hong T. Data mining of space heating system performance in affordable housing. Building and Environment 89, 2015, pp. 1-13.
7. Delmastro, C, Mutani, G, Schranz, L. The evaluation of buildings energy consumption and the optimization of district heating networks: a GIS-based model, International Journal of Energy and Environmental Engineering, Special Issue "8th AIGE Conference (Italian Association for Energy Management)", 2015
8. Mutani, G and Pairona, M. A model to evaluate the heating energy consumption for residential buildings in Turin, L'Ufficio Tecnico, Ed. Maggioli, pp. 21–36, (2014)
9. Delmastro, C, D'Oca S, Corgnati, SP. Advanced input modelling for energy planning: including archetypal user profiles in energy models. Proceeding of Sustainable Cities: Designing for People and the Planet Coimbra, 14-15 May, 2015
10. 15° Censimento generale della popolazione e delle abitazioni. 2011. Istituto nazionale di statistica – ISTAT. 11 agosto 2014.
11. IEA Annex 53 - Task Force. Total energy use in residential buildings – the modeling and simulation of occupant behavior. Final report 2012
12. Han, J and Kamber M. Data Mining: Concepts and Techniques. Morgan Kaufmann Publishers 2001.
13. Hand, D J, Mannila, H, Smyth, P. Principles of Data Mining. MIT Press 2001
14. IEA Annex 66 "Definition and Simulation of Occupant Behavior in Buildings" www.annex66.org
15. Mutani, G, Vicentini, G. The open data for the analysis of primary energy demand of the residential buildings and the potential energy savings, Smart City Exhibition 2013, Bologna pp. 15–37.