

Clustering of educational building load data for defining healthy and energy-efficient management solutions of integrated HVAC systems

*Daniele Testi*¹, *Alessandro Franco*^{1,*}, *Paolo Conti*¹, and *Carlo Bartoli*¹

¹Department of Energy, Systems, Territory and Constructions Engineering (DESTEC), University of Pisa, 56122 Pisa, Italy

Abstract. The COVID-19 pandemic is changing the way individuals, worldwide, feel about staying in public indoor spaces. A strict control of indoor air quality and of people's presence in buildings will be the new normal, to ensure a healthy and safe environment. Higher ventilation rates with fresh air are expected to be a requirement, especially in educational buildings, due to their high crowding index and social importance. Yet, in this framework, an increased use of primary energy may be overlooked. This paper offers a methodology to efficiently manage complex HVAC systems in educational buildings, concurrently considering the fundamental goals of occupants' health and energy sustainability. The proposed four-step procedure includes: dynamic simulation of the building, to generate synthetic energy loads; clustering of the energy data, to identify and predict typical building use profiles; day-ahead planning of energy dispatch, to optimize energy efficiency; dynamic adjustment of air changes, to guarantee a safe indoor air quality. Clustering and forecasting energy needs are expected to become particularly effective in a highly regulated context. The technique has been tested on two university classroom buildings, considering pre-lockdown attendance. This notwithstanding, quality and significance of the obtained thermal energy clusters push towards a benchmark post-pandemic application.

1 Introduction

The outbreak of COVID-19 pandemic has brought global attention to the need of a strict indoor air quality control, to prevent airborne virus transmission and avoid unsustainable lockdown of activities in crowded indoor spaces. Educational buildings have been shut in an early phase of the emergency, due to the epidemic risk associated with the high density of young, socially interacting, and often asymptomatic individuals. Post-lockdown strategies for reopening these spaces involve strict regimentation of people's presence and implementation of operating-room-like air changes. Albeit effective in times of health

* Corresponding author: alessandro.franco@ing.unipi.it

emergency, these solutions are highly energy-consuming and, on the long term, contrast to the equally important climate emergency. If on one hand it is not likely that people's presence in indoor spaces will return to be loosely regulated as in the pre-pandemic era, on the other hand, there is plenty of room to optimize the energy management of HVAC systems and of the integrated energy-generation and storage devices, without compromising on the air quality levels required by health authorities. In this respect, correctly forecasting the profiles of building energy needs, including ventilation, becomes paramount, so that daily energy generation and dispatch can be appropriately planned to dynamically meet those demands in the most energy-efficient and environmental-friendly way. With this aim, the adoption of clustering techniques to process energy data is a promising way of dealing with buildings having a somehow predictable occupation, such as educational buildings, and particularly in the aftermath of a respiratory virus pandemic, when occupancy is expected to be strictly controlled.

Worldwide, in 2018, buildings and construction sector accounted for 36% of final energy use and 39% of energy and process-related carbon dioxide (CO₂) emissions [1]. The energy breakdown by different end-uses in 2015 for the EU-28 countries [2] shows that heating and cooling accounted for almost 50% of the total final energy demand. Official data by the Italian Ministry of Education, University and Research show that 55% of educational buildings have been built before 1976. Several national studies on the state of these buildings report the urgent need of deep refurbishment for energetic purposes, together with anti-fire and anti-seismic reasons [3-4]. In this context, educational buildings, being mainly public, highly visible, and socially strategic, must set the example not only keeping a high standard in occupational health (health risk factors are thoroughly reviewed in [5]), but also in terms of high energy efficiency and low environmental impact.

In contrast to the above objectives, schools and universities are places where students and teachers spend most of their time in crowded indoor spaces, with subsequently critical energy needs, particularly for ventilation and air conditioning purposes. Correctly sized HVAC systems with appropriate controls guarantee that the required quality of indoor air is maintained at any time. However, especially in times of pandemic, when these quality levels are inevitably high for public health reasons, an increased effort must be made to accomplish the air-quality goals with reduced energy expenses.

New normal solutions, as mentioned, will concurrently be respectful of occupants' health, of energy, and of the environment. As a novel contribution to this aim, the present paper proposes a general procedure that is based on dynamic simulation of the building, clustering of energy demand data, and day-ahead planning of energy systems operation for minimization of non-renewable primary energy uses, followed by a variable air volume (VAV) ventilation, based on actual people's presence, to keep concentration of pollutants and airborne microorganisms below risk thresholds.

The paper is structured as follows: in the next section, clustering approaches applied to the energy needs of buildings will be reviewed, showing the usefulness of this data-processing technique; then, in Section 3, the proposed energy-management procedure will be illustrated, starting from the dynamic simulation of the building, up to the identification of the optimal energy-dispatch sequence in each cluster of forecast daily loads and the real-time control of air changes necessary to maintain a healthy indoor environment; finally, in Section 4, dynamic simulation and data clustering will be applied to the thermal energy demand profiles of two classroom buildings of the University of Pisa, implementing actual pre-pandemic attendance, which is a more challenging problem than the one posed by future paradigms that involve a programmed presence and space distribution of students in classrooms.

2 Clustering of building thermal load data

Clustering of time-series signals is an important area of research in several fields [6], such as economics, engineering, finance, medicine, biology, physics, geology, and many others. Clustering refers to the unsupervised ability to aggregate similar objects together, and the basic clustering operation corresponds to take a set of N objects and group them into a reduced number of K clusters. Apart from the employed algorithm, chosen among the large variety of available ones [7], there are three main motivations for doing so. First, a good clustering has predictive power; clustering leads to a more efficient description of the available data and helps in planning better actions. Secondly, a cluster allows to compress a large amount of data into a single information, corresponding to the centre of the cluster, the so-called centroid. The third reason for clustering is to identify the outliers, i.e., some particular conditions in which clusters fail to accurately represent the original data.

The application of clustering techniques to building energy-use data is quite recent. Thanks to the increasing availability of reliable and economic metering devices, data acquisition and data storage systems, long time series of energy loads can be collected in databases and then processed by automated software. Different algorithms have been effectively applied to building monitoring data for defining representative energy-use profiles [8]. From the analysis of the energy patterns corresponding to each of the obtained clusters, conditions of energy misuse can be exposed, and energy-saving measures can be identified, as in [9].

Focusing on educational buildings, specific studies have applied cluster analysis to a large stock for the purpose of energy benchmarking [10] or energy rating [11]. In [12], hourly monitored data of electricity use in several university buildings have been grouped in clusters, to guide possible retrofit measures. As for hourly or sub-hourly thermal load data, obtaining meaningful clusters is a much more difficult task, because of the scarce repeatability or similarity of heating, cooling and ventilation energy-demand profiles, which are significantly dependent on the evolution of the external climate. In addition, heat and cooling metering devices generally produce noisy data, due to the measurement errors typically introduced by small temperature differences and to the slow dynamics and lags especially associated to hydronic systems.

Yet, clustering thermal load patterns would be extremely helpful in handling complex HVAC systems, typical of modern educational buildings. The problem of a high variability in energy loads caused by their dependence on the external temperature can be mitigated or partially solved by introducing a load-temperature correlation, the so-called energy signature, and then clustering the data as energy deviations from the values expected by the analytical application of the correlation. However, a possibly long monitoring period is necessary to obtain a reliable building energy signature. This latter issue, together with the previously mentioned metering difficulties, can be addressed by the creation of a thermodynamic model of the building and the subsequent computational dynamic simulation of its thermal and ventilation needs, as in [13]. In this way, synthetic load data can be generated, at a time discretization compatible with the physics and the spatial granularity of the adopted building model (typical time scales range between 15 minutes and 2 hours).

The energy requirements predicted by the dynamic simulation of the building constitute the time series datasets to be processed by a suitable clustering algorithm, such as k -means or k -medoids. To guide the choice of the number of clusters in which daily energy profiles should be grouped, some performance metrics evaluating the consistency of the obtained classification ought to be calculated: examples are intra-cluster cohesion or compactness (how similar each element is to the other elements of its own cluster) and inter-cluster separability (how different an element is to the elements of the other clusters). The so-

called silhouette is a common technique to be applied for this purpose [14], also providing a handy graphical representation.

3 Description of the methodology

A step-by-step methodology for obtaining healthy and energy-efficient management solutions of HVAC systems in educational buildings is illustrated hereafter. The methodology is targeted to systems that have a significant level of complexity and integration, otherwise, their management would be trivial, with no space for improvement. Conversely, in systems with storage devices, renewable energy technologies, hybrid generators, or different energy carriers, there is plenty of room for planning the energy-dispatch sequence of each unit and efficiently provide the required heating, cooling, and ventilation services to the end users.

Clearly, in the aftermath of the COVID-19 pandemic, HVAC systems are going to become more and more advanced, in response to the urgency of accurate indoor air quality control. This tendency is not expected to cease in the next years, whether SARS-CoV-2 becomes endemic on the global scale or not. As a matter of fact, at least 3 new threatening respiratory viruses have spilled over from animals to infect humans in the last 20 years, namely SARS-CoV, identified in 2002, MERS-CoV, identified in 2012, and the new coronavirus identified in late 2019. In addition, other respiratory viruses with inter-human transmission such as S-OIV, responsible for swine influenza, are still endemic in many areas and may pose periodical threats to the public health. This outlook suggests as highly unlikely that ventilation rates and air quality levels will be allowed to return to pre-pandemic values. Consequently, HVAC systems will have to be upgraded or newly designed to comply with these higher standards.

The increased investment on components and controls of HVAC and energy systems must be paid off by a correct management, not only to guarantee a healthy indoor environment, but also to minimize non-renewable primary energy uses. One of the keys to an optimal management of the system, as it will be explained, is the ability to forecast energy loads and ventilation rates.

The proposed methodology involves 4 main phases, to be performed consecutively:

- 1) building modelling and simulation;
- 2) energy-load clustering and forecasting;
- 3) energy-dispatch optimization and implementation;
- 4) post-strategy for a healthy environment.

As a result of the application of these steps to an educational building, indoor air quality is kept under control and energy usage is minimized. A detailed description of the 4 phases will follow, together with possible case-dependent alternatives.

3.1 Phase 1: building modelling and simulation

The objective of this phase is to obtain a long time series (e.g., hourly data for one year of building use) of heating, cooling, and ventilation loads, without the need of actually monitoring the building energy systems.

The first step is creating a sufficiently accurate physical model of the building, which correctly takes into account all the internal and external energy flows and exchanges. Once all information about building geometry, materials, and use are collected, the model can be implemented in available dynamic simulation software (there exist several software, such as EnergyPlus, DOE-2, and TRNSYS; each of these tools has specific features that are reviewed and compared in [15]), in-house tools (as the one developed in [16]), or mixed solutions, with user-built subroutines integrated in a commercial software (as done in [17]).

In any case, validation of the simulation results is needed, and it can be performed by comparing the actual energy billings with the calculated annual energy uses, reproducing the climatic conditions of the year to which the billings are referred.

Databases of statistical climatic years are available in most building simulation software for the main cities around the globe, based on at least 10-year long local measurements. If a specific meteorological year is needed, a solution is to download the necessary data by the online meteorological database MERRA-2 (Modern-Era Retrospective analysis for Research and Applications, version 2). MERRA-2 provides weather variables generated by a numerical model that employs satellite observations by NASA. The database covers, on an hourly basis, a timespan beginning in January 1980, with a constant update up to one month prior the current day. The model has a spatial resolution of approximately 50 km [18].

The output of the dynamic simulation is a complete synthetic dataset of energy uses for the examined educational building, which can be conveniently processed in the next phase of data clustering.

3.2 Phase 2: energy-load clustering and forecasting

The objective of this phase is to process the data generated in Phase 1, in order to identify typical days of energy loads and perform day-ahead predictions of energy use profiles. The employed technique for compressing and classifying the data in groups of typical days is clustering, with its related algorithms, as discussed in Section 2.

As above explained, prior to clustering thermal energy data, it is useful to find the building energy signature, which provides a first estimate of the thermal load as a function of the external temperature (or, especially for estimating the cooling need, of the sol-air temperature [19-20]). The load-versus-temperature correlation can be simply obtained by a best fit of its coefficients. Then, data clustering can be performed on the energy deviations from the correlated values, to produce more compact and temperature-independent clusters. Instead, these deviations will mainly depend on occupants' presence and behaviour. Indeed, people play an important role on the energy balance of crowded buildings, due to sensible and latent heat produced by their bodies, internal gains associated to the use of electronic and lighting devices, solar radiation screening by management of shutters, curtains, or blinds, switch-on and switch-off schedules of energy systems, and, most important of all in post-pandemic times, variable air changes, based on the number of occupants. Since people's presence is somehow aleatory for many building uses, dynamic simulation loads profiles and subsequent clusters of typical days may carry a significant degree of uncertainty. However, this is not the case of educational buildings, where the number of occupants is quite predictable, thanks to a known calendar of lessons in classes. In a new normal, post-pandemic situation, ventilation rates per person will be increased, emphasizing the energy impact of attendance level and people's behaviour, but their prediction will also be much more accurate, having to comply with social-distancing rules and with a predetermined number of available (or even reserved) seats. As a consequence, not only the dynamic simulation results, but also the outcomes of the clustering procedure are going to be more reliable, favouring the application of the present methodology, also in view of the next phases.

A proper clustering of energy-load data allows to classify the most relevant types of requests and accordingly plan the energy dispatch, optimizing the operation of the available generation and storage devices, thus minimizing the use of non-renewable primary energy in the identified groups of energy profiles.

By a close observation of the obtained clusters, it is possible to tag them, referring to known and replicable conditions: e.g., workdays and weekends, holiday periods, lesson

weeks and exam weeks, and so on. This classification is going to be more accurate in buildings with predictable usage and with energy requirements that are highly dependent on people's presence and behaviour, a situation that is typical of present and, even more, of future educational buildings, as previously explained.

Associating a cluster of data to a repeatable condition means that the energy requests for that case can be periodically forecast, arranging the energy management of the systems, in advance, to comply with those nominal loads.

3.3 Phase 3: energy-dispatch optimization and implementation

The objective of this phase is to identify and implement, for each group of data obtained in Phase 2 (i.e. typical days), the best operative solutions of energy management.

As reasoned above, occupants' behaviour in indoor spaces has a prominent influence on energy needs. In addition, the use of public buildings cannot be left to chance anymore; on the contrary, in the future, it will be strictly planned, especially in educational buildings. Indeed, if compared to commercial buildings, schools and universities have the advantage of knowing a priori, with adequate confidence, the number of attendees in a given room and at a given time, based on the students enrolled to each course.

This is undoubtedly an aspect capable of improving the quality of the performed clustering, thus limiting the number of predetermined conditions in which the energy-management choices must be optimized. The acquired predicting power allows to plan, for the day ahead, the optimal use of the existing energy systems. Possible energy-saving actions in integrated HVAC systems are:

- choice of the ventilation rate according to the estimated number of occupants [21], rather than running the mechanical ventilation or air conditioning system at its maximum flow rate, in force of an exaggeratedly precautionary approach;
- choice of the energy generator to be switched on in a hybrid system, based on the expected performance of each device [22-23];
- choice of the storage system to be charged or discharged [24], based on the optimization rules calculated in relation to the forecast energy requirements and to the expected energy production by the available renewable energy systems [25];
- activation of free-cooling or free-heating options on the air system;
- use of solar screens;
- rational occupation of rooms and spaces, when applicable.

Obviously, the above list is non-exhaustive, and many other energy-management and energy-dispatch solutions can be implemented, depending on the existing energy system.

3.4 Phase 4: post-strategy for a healthy environment

The objective of this phase is to always guarantee healthy conditions in the indoor environment, controlling in real time the operation of the HVAC system, and correcting the unavoidable forecasting errors.

In spite of a high-quality clustering and day-ahead forecasting of energy needs, the actual energy flows will inevitably be different from the predicted ones. There are various reasons for this, and some of them will be mentioned as examples:

- unexpected weather conditions, altering the thermal response of the building and the performance of devices such as heat pumps, photovoltaics, etc.;
- unexpected variation of people's presence and behaviour;
- inaccuracies in the modelling of the building;
- inaccuracies in the modelling of the HVAC and energy systems;
- errors or delays in the implementation of the desired energy-dispatch sequence.

These discrepancies between expected and actual conditions might cause occupants' thermal discomfort or, in the worst case, could lower the air quality to unacceptable levels from the hygienic-sanitary standpoint. This risk calls for the introduction of a so-called post-strategy, consisting in the adjustment of indoor air conditions by a real-time controlled VAV ventilation, capable of continuously guaranteeing a healthy environment for the occupants by means of fresh air.

Briefly, after an optimized planning of energy dispatch, real-time adjustment of air changes is a practical way for achieving the necessary goal of a healthy indoor environment, in addition to energy efficiency. The overall methodology, as explained in detail, applies particularly well to educational buildings and, notably, for the needs of a respiratory virus post-pandemic era. As a matter of fact, social needs for health and comfort in living spaces (which can include not only air quality and thermal comfort, but also acoustic and visual comfort) are technically related aspects that can be handled by a proper design and management of engineering systems.

4 Application to two university buildings and clustering results

To show the potential of the proposed methodology, Phase 1 and Phase 2 are applied to real test cases, namely two classroom buildings of the School of Engineering of the University of Pisa. Students' attendance is based on pre-pandemic schedules of classes, with implemented random uncertainties that are certainly conservative with respect to the new normal of access and presence control.

The imposed meteorological conditions are the ones obtained in the last year by an in-house weather station, located within the Engineering School perimeter, and already used for scientific research purposes [26].

The main characteristics of the two classroom buildings under analysis (Building C and Building F) are reported in Table 1.

Table 1. Geometrical, structural, and functional characteristics of the examined buildings.

Property	Building C	Building F
Net floor area	1561 m ²	2400 m ²
Gross volume	6198 m ³	10080 m ³
Number of floors	5	1
Number of classrooms	12	9
Total number of seats	876	1424
Surface-area-to-volume ratio	0.33 m ⁻¹	0.56 m ⁻¹
Window-to-floor area ratio	0.25	0.15
Type of structure	Masonry wall with face bricks and reinforced concrete structure	Brick masonry wall with plaster finish
Average vertical-wall thickness	30 cm	35 cm
Year of construction	1993	2004

Figures 1 and 2 show the number of occupants implemented in the dynamic simulations of Building C and Building F, respectively. The adopted thermodynamic model for the buildings is of thermal-network type. A detailed description of the validated model and related equations can be found in [13].

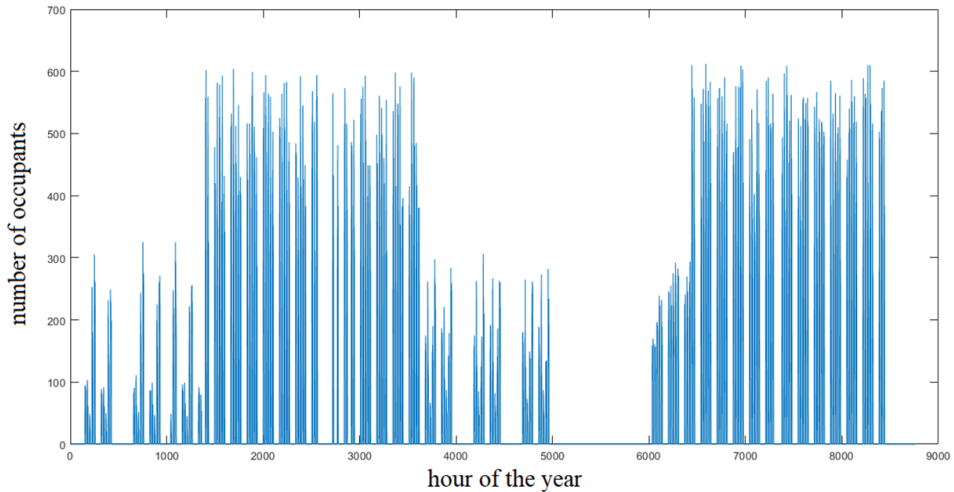


Fig. 1. Profile of students' presence implemented in the annual dynamic simulation of Building C.

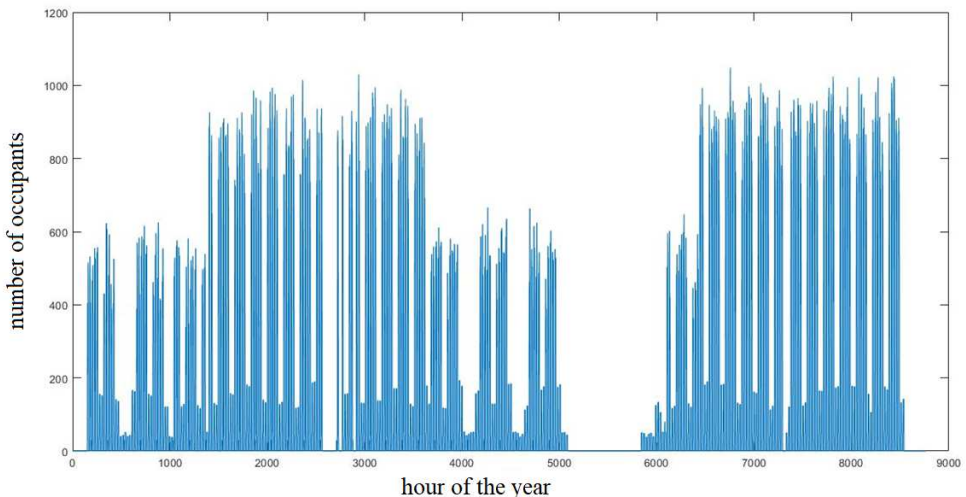


Fig. 2. Profile of students' presence implemented in the annual dynamic simulation of Building F.

As a result of the annual dynamic simulation of Building C, monthly averaged daily profiles of heating loads can be obtained, as illustrated in Figure 3. It can be observed that in January and February (period of the winter-session exams) the energy needs are significantly higher than in the other heating months (during which classes are given). Analysing the weekly profiles of heating load of Building C, higher demands on Mondays and shorter demands on Saturdays can be reported. Daily, a morning peak can always be noted (see again Figure 3), at the opening hour of the building, and a second smaller peak is often present during the afternoon.

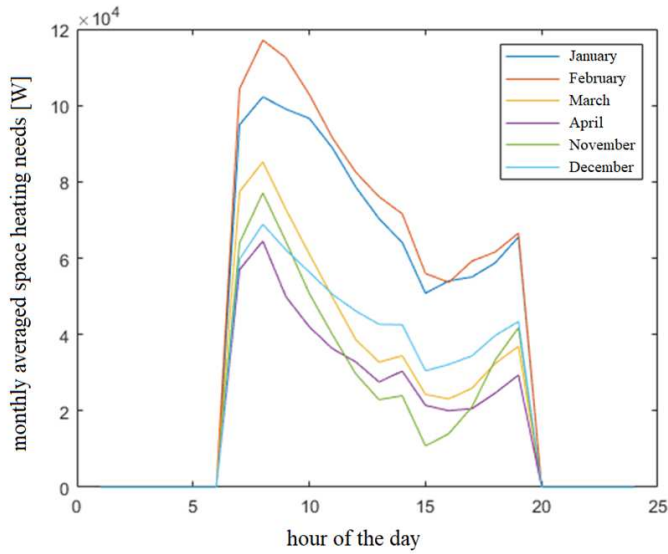


Fig. 3. Monthly averaged daily heating loads obtained by annual dynamic simulation of Building C.

Figure 4 shows the energy signature of Building F, based on its annual dynamic simulation. As illustrated, conditions with a higher number of students are responsible for larger energy uses, due to higher ventilation requirements. The black line in Figure 4 represents the nominal energy signature, based on design external and internal temperatures and design thermal load.

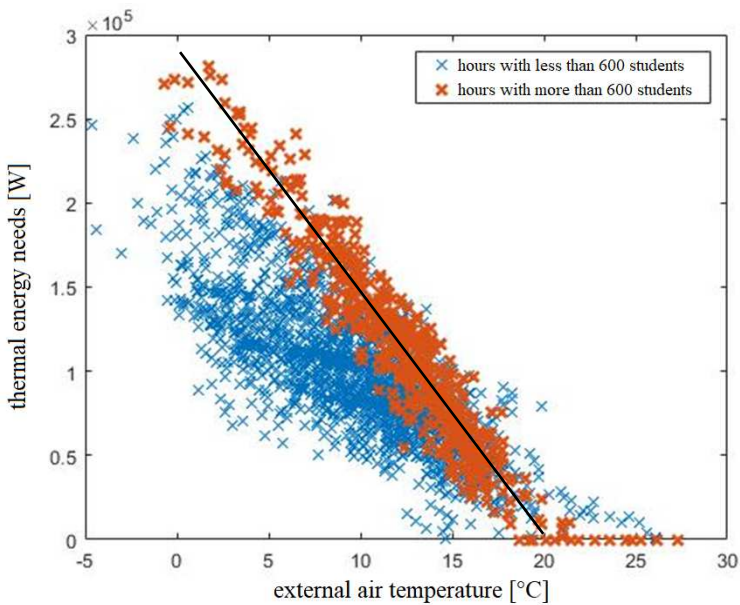


Fig. 4. Thermal energy load versus temperature (energy signature) for Building F.

The raw data of thermal energy needs of Building F have been grouped by the k-means algorithm into 6 clusters of daily profiles. Figure 5 clearly shows that clusters are mainly classified based on their mean daily temperature, hiding the dependence of energy requirements on people's presence.

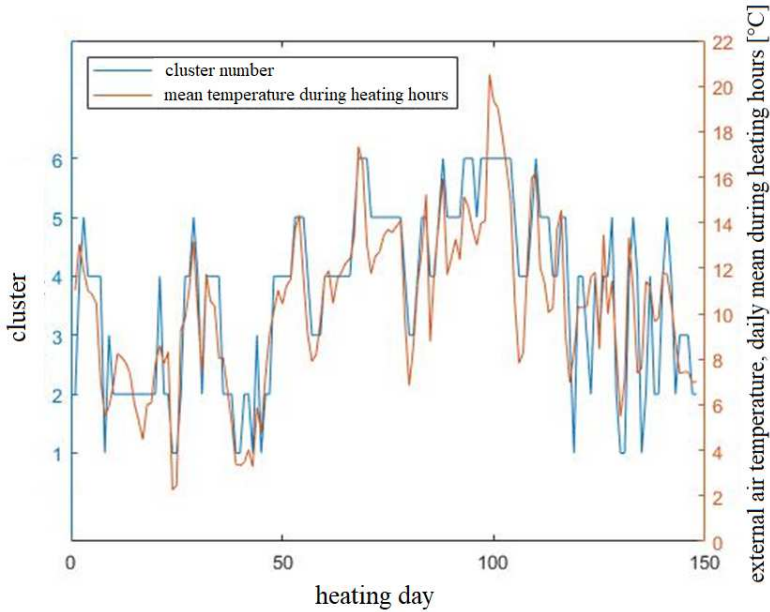


Fig. 5. Clustering of thermal energy need daily profiles and correlation to the mean external temperature for Building F.

However, as previously explained, the thermal energy data can be appropriately processed, and clustering can be performed on the deviations from the energy-signature correlation, to obtain a lower number of clusters, which are also temperature independent. Figure 6 depicts the silhouette values [14] of the two clusters in which the data, processed as stated, are classified. A silhouette value close to one means that the point has been correctly classified in that cluster, while a positive or negative value close to zero means that the point is at the border between two clusters (a value close to negative one implies that the point should have been included in the neighbouring cluster).

Cluster 1 contains most of the days during which exams or classes are given, and Cluster 2 includes almost all weekends and other days in which rooms are occupied by students mainly for studying.

Cluster 2 has a better quality than Cluster 1, but the overall result is encouraging, especially considering the introduced random term of uncertainty on people's presence (with reduction of attendance implemented in the model as high as 50%). As already remarked, this latter term is coherent with pre-pandemic routines, but the repeatability and predictability of future occupation schemes can certainly improve these figures.

In conclusion, Figure 7 illustrates the results of VAV ventilation control in Building F. In this case, the concentration of CO₂, as airborne pollutant, is used as indicator of room air quality, and air changes are adjusted to keep CO₂ concentration below the threshold of 1000 ppm, which is a value accepted for maintaining a healthy indoor environment in educational buildings [27-28]. Figure 7 also compares VAV to constant air volume (CAV) ventilation, implementing a nominal value of almost 4 air changes per hour during the building opening hours, in accordance with the Italian current standard [29]. It is evident that a precautionary approach to ventilation rates can be extremely energy consuming in this type of buildings.

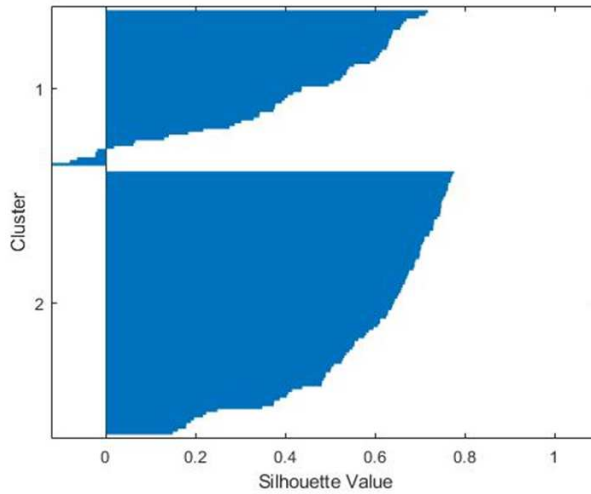


Fig. 6. Silhouette values of the two clusters obtained for Building F, as deviations from the energy-signature correlation.

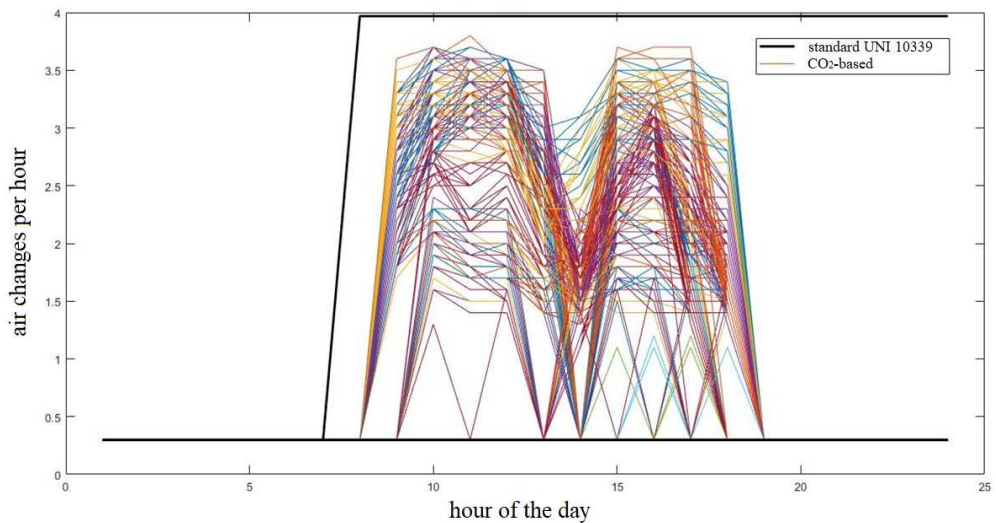


Fig. 7. Daily air changes per hour for Building F: CO₂-based VAV versus standard-based CAV ventilation.

5 Conclusions

Among the effects of the COVID-19 pandemic, a new social perception of indoor environments is emerging, with immediate consequences on the design, operation, and maintenance of HVAC systems, to limit the risk of transmission and contagion from airborne respiratory viruses. Schools and universities, shut almost worldwide during the epidemic emergency, should be promptly reopened, to avoid a long-term impact on students, not only from the educational point of view, but also from the social and psychological standpoints [30-31]. Being the crowding index of educational buildings typically high, to comply with new hygienic-sanitary rules, the air changes with fresh (uncirculated) air are going to be increasingly demanding in terms of energy usage. This

issue calls for new energy-management solutions, to face the health emergency without disregarding the equally important climate emergency.

In this framework, the present paper has proposed a novel methodology to manage integrated HVAC systems in educational buildings, to obtain a healthy environment and a high energy efficiency. The outlined procedure has been divided in four consecutive phases, which include a dynamic simulation of the building, a clustering of energy demand data, a day-ahead planning of energy systems operation and energy dispatch for minimization of non-renewable primary energy uses, and a post-strategy, to correct the forecasting errors, with the adjustment of ventilation rates so to keep a sufficiently high level of indoor air quality.

With respect to other works employing clustering techniques on energy data of buildings, the proposed use of synthetic loads obtained by dynamic simulation solves issues associated with energy-flow monitoring. Besides, a specific attention is devoted to the new needs of indoor air quality arisen during the pandemic, concurrently with advanced energy-saving strategies.

The methodology has been presented in detail, and subsequently applied to two classroom buildings of the University of Pisa, showing that a good-quality of energy-load data clustering can be obtained, even with pre-pandemic unregulated attendance.

Further advancements include the application of indoor air quality metrics, similar to the ones defined in [32], specifically tailored to educational buildings and to the updated requirements for limiting airborne transmission of respiratory viruses. Moreover, multi-objective optimization techniques could be introduced for building management and decision making, weighting indicators of a healthy indoor environment and of an energy-efficient HVAC system [33], even including data obtained via social interaction of students through tailored ICT platforms [34]. Finally, a detailed analysis conducted on a classroom building with post-lockdown attendance could provide an important benchmark for future studies.

The authors gratefully acknowledge the financial support of the Italian Ministry of Education, University and Research (MIUR), in the framework of the Research Project of Relevant National Interest (PRIN) “The energy FLEXibility of enhanced HEAT pumps for the next generation of sustainable buildings (FLEXHEAT)” (PRIN 2017, Sector PE8, Line A, Grant n. 33). We also would like to thank the M.Sc. student Mr. Lorenzo Miserocchi for his contribution to setting, launching, and processing the dynamic simulations of the two university buildings.

References

1. International Energy Agency (IEA), *Global status report for buildings and construction 2019*, IEA, Paris (2019), available at: <https://www.iea.org/reports/global-status-report-for-buildings-and-construction-2019> [accessed on September 2, 2020]
2. T. Fleiter, R. Elsland, M. Rehfeldt, J. Steinbach, U. Reiter, G. Catenazzi, M. Jakob, C. Rutten, R. Harmsen, F. Dittmann, P. Rivière, P. Stabat, *Profile of heating and cooling demand in 2015*, D3.1 report, Heat Roadmap Europe 2050, A low-carbon heating and cooling strategy (2017), available at: https://heatroadmap.eu/wp-content/uploads/2018/11/HRE4_D3.3andD3.4.pdf [accessed on September 2, 2020]
3. D. Testi, M. Rocca, E. Menchetti, S. Comelato, *Criticalities in the NZEB retrofit of scholastic buildings: analysis of a secondary school in Centre Italy*, Energy Procedia, **140**, 252-264 (2017)
4. S. Della Torre, M. Bocciarelli, L. Daglio, R. Neri, *Buildings for education – a multidisciplinary overview of the design of school buildings*, Springer (2020)
5. M. Dovjak, A. Kucek, *Creating healthy and sustainable buildings*, Springer (2019)

6. L. Kaufman, P.J. Rousseeuw, *Finding groups in data – an introduction to cluster analysis*, John Wiley & Sons (2005)
7. G. Gan, C. Ma, J. Wu, *Data clustering: theory, algorithms, and applications*, ASA-SIAM Series on Statistics and Applied Probability (2007)
8. J. Yang, C. Ning, C. Deb, F. Zhang, D. Cheong, S.E. Lee, C. Sekhar, K. Tham, *k-shape clustering algorithm for building energy usage patterns analysis and forecasting model accuracy improvement*, Energy and Buildings, **146**, 27-37 (2017)
9. A. Zakovorotnyi, A. Seerig, *Building energy data analysis by clustering measured daily profiles*, Energy Procedia, **122**, 583-588 (2017)
10. P. Marrone, P. Gori, F. Asdrubali, L. Evangelisti, L. Calcagnini, G. Grazieschi, *Energy benchmarking in educational buildings through cluster analysis of energy retrofitting*, Energies, **11**, 649, 20 pp. (2018)
11. M. Santamouris, G. Mihalakakou, P. Patargias, N. Gaitani, K. Sfakianaki, M. Papaglastra, C. Pavlou, P. Doukas, E. Primikiri, V. Geros, M.N. Assimakopoulos, R. Mitoula, S. Zerefos, *Using intelligent clustering techniques to classify the energy performance of school buildings*, Energy and Buildings, **39**, 45-51 (2007)
12. K. Li, Z. Ma, D. Robinson, J. Ma, *A two-level clustering strategy for energy performance evaluation of university buildings*, Proceedings of the 4th International Conference on Building Energy and Environment (COBEE 2018), Melbourne, Australia, Paper 061, 168-173 (2018)
13. E. Schito, P. Conti, L. Urbanucci, D. Testi, *Multi-objective optimization of HVAC control in museum environment for artwork preservation, visitors' thermal comfort and energy efficiency*, Building and Environment, **180**, 107018, 15 pp. (2020).
14. P.J. Rousseeuw, *Silhouettes: a graphical aid to the interpretation and validation of cluster analysis*, Computational and Applied Mathematics, **20**, 53-65 (1987)
15. D.B. Crawley, J.W. Hand, M. Kummert, B.T. Griffith, *Contrasting the capabilities of building energy performance simulation programs*, Building and Environment, **43**, 661-73 (2008)
16. D. Testi, E. Schito, E. Tiberi, P. Conti, W. Grassi, *Building energy simulation by an in-house full transient model for radiant systems coupled to a modulating heat pump*, Energy Procedia, **78**, 1135-1140 (2015)
17. E. Schito, D. Testi, *Integrated maps of risk assessment and minimization of multiple risks for artworks in museum environments based on microclimate control*, Building and Environment, **123**, 585-600 (2017)
18. R. Gelaro and other 30 authors, *The Modern-Era Retrospective analysis for Research and Applications, version 2 (MERRA-2)*, Journal of Climate, **30**, 5419-5454 (2017)
19. P.W. O'Callaghan, S.D. Probert, *Technical note – sol-air temperature*, Applied Energy, **3**, 307-311 (1977)
20. D. Testi, E. Schito, P. Conti, *Cost-optimal sizing of solar thermal and photovoltaic systems for the heating and cooling needs of a Nearly Zero-Energy Building: design methodology and model description*, Energy Procedia, **91**, 517-527 (2016)
21. A. Franco, E. Schito, *Definition of optimal ventilation rates for balancing comfort and energy use in indoor spaces using CO₂ concentration data*, Buildings, **10**, 135, 19 pp. (2020)
22. F. D'Ettorre, P. Conti, E. Schito, D. Testi, *Model predictive control of a hybrid heat pump system and impact of the prediction horizon on cost-saving potential and optimal storage capacity*, Applied Thermal Engineering, **148**, 524-535 (2019)
23. F. D'Ettorre, M. De Rosa, P. Conti, D. Testi, D. Finn, *Mapping the energy flexibility potential of single buildings equipped with optimally-controlled heat pump, gas boilers and thermal storage*, Sustainable Cities and Society, **50**, 101689, 13 pp. (2019)

24. L. Urbanucci, F. D’Ettorre, D. Testi, *A comprehensive methodology for the integrated optimal sizing and operation of cogeneration systems with thermal energy storage*, *Energies*, **12**, 875, 17 pp. (2019)
25. L. Gigoni, A. Betti, E. Crisostomi, A. Franco, M. Tucci, F. Bizzarri, D. Mucci, *Day-ahead hourly forecasting of power generation from photovoltaic plants*, *IEEE Transactions on Sustainable Energy*, **9**, 831-842 (2017)
26. C. Bartoli, P. Conti, A. Franco, D. Testi, *Experimental analysis of an air heat pump for heating service using a “hardware-in-the-loop” system*, *Energies*, **13**, 4498, 18 pp. (2020)
27. L. Schibuola, M. Scarpa, C. Tambani, *CO₂-based ventilation control in energy retrofit: an experimental assessment*, *Energy*, **143**, 606-614 (2018)
28. A. Franco, F. Leccese, *Measurement of CO₂ concentration for occupancy estimation in educational buildings with energy efficiency purposes*, *Journal of Building Engineering*, 101714 (2020), available online (<https://doi.org/10.1016/j.jobe.2020.101714>)
29. UNI, *Air-conditioning systems for thermal comfort in buildings – General, classification and requirements – Offer, order and supply specifications*, Standard UNI 10339 (in Italian), Italian National Agency for Unification (1995)
30. X. Xie, Q. Xue, Y. Zhou, K. Zhu, Q. Liu, J. Zhang, R. Song, *Mental health status among children in home confinement during the coronavirus disease 2019 outbreak in Hubei province, China*, *JAMA Pediatrics*, Research Letter, April 24, E1-E3 (2020)
31. M. Poletti, A. Raballo, *Evidence on school closure and children’s social contact: useful for coronavirus disease (COVID-19)?*, *Eurosurveillance*, **25**, Letter to the editor, April 30, 1-2 (2020)
32. E. Schito, D. Testi, W. Grassi, *A proposal for new microclimate indexes for the evaluation of indoor air quality in museums*, *Buildings*, **6**, 41, 15 pp. (2016)
33. E. Schito, P. Conti, D. Testi, *Multi-objective optimization of microclimate in museums for concurrent reduction of energy needs, visitors’ discomfort and artwork preservation risks*, *Applied Energy*, **224**, 147-159 (2018)
34. A. Franco, *Balancing user comfort and energy efficiency in public buildings through social interaction by ICT systems*, *Systems*, **8**, 29, 16 pp. (2020)