TOWARDS A MORE ENERGY EFFICIENT EDUCATIONAL ARCHITECTURE IN CITIES: TYPOLOGIES OF BARCELONA AND QUITO PUBLIC SCHOOLS FOR ENERGY MODELLING

Ledesma, Gabriela [1]*; Nikolic, Jelena [2] y Pons-Valladares, Oriol [3]

Initial submission: 2019-06-04; **Definitive submission:** 2019-11-20; **Publication:** 2019-12-21

Citation: Ledesma, G. *et al.* (2019). Towards a More Energy Efficient Educational Architecture in Cities: Typologies of Barcelona and Quito Public Schools for Energy Modelling. En *XIII CTV 2019 Proceedings: XIII International Conference on Virtual City and Territory: "Challenges and paradigms of the contemporary city"*: UPC, Barcelona, October 2-4, 2019. Barcelona: CPSV, 2019, p. 8505. E-ISSN 2604-6512. DOI http://dx.doi.org/10.5821/ctv.8505

Resumen

La edificación es uno de los sectores con mayor consumo energético mundial, representando aproximadamente un tercio del consumo global de energía. El análisis del inventario edilicio de una ciudad requiere la aplicación de técnicas estadísticas para agrupar edificios con características homogéneas creando así tipologías arquitectónicas. En temas de eficiencia energética se ha utilizado el análisis de conglomerados para evaluar el rendimiento energético del total de un parque edificado usando edificios referenciales. Estos son edificios reales o teóricos cuyas características son representativas de toda la muestra y cuyos resultados pueden generalizarse sin incurrir en errores significativos.

Este estudio analiza el inventario total de escuelas existentes y establece un proceso de creación de edificios referenciales teóricos -basado en técnicas estadísticas- para las ciudades de Barcelona-España y Quito-Ecuador. El objetivo general es generar edificios arquetípicos para un análisis energético detallado, incluyendo su comportamiento térmico pasivo y su resiliencia al cambio climático. Este estudio es parte de una investigación más amplia sobre la rehabilitación de la envolvente arquitectónica a través de simulaciones energéticas usando algoritmos para un tratamiento pasivo óptimo de sus superficies individuales. Partiendo del análisis del parque edilicio de las ciudades fue posible obtener las características relevantes de la muestra y hacer una clasificación primaria basada en el año de construcción y el área bruta del edificio. El análisis de conglomerado se usó para dividir las muestras en subconjuntos de edificios similares con respecto a sus características geométricas y de comportamiento térmico. Finalmente, el cálculo matemático de los centros de cada subconjunto se utilizó para crear los edificios referenciales teóricos.

La muestra estuvo limitada a escuelas públicas en áreas urbanas consolidadas debido a la facilidad de acceso a la información, ya sea de dominio público o proporcionada por los departamentos educativos locales. La clasificación primaria se realizó siguiendo la metodología del proyecto europeo TABULA generando matrices bidimensionales entre el período de construcción y área total del edificio y, asignando cada centro escolar a una categoría específica. Se crearon fichas técnicas de cada edificación detallando sus características termo-físicas recopiladas de información catastral, planos, sistemas GIS y visitas de campo. Teniendo en cuenta el enfoque pasivo de esta investigación, las variables de caracterización son: compacidad, área en planta baja, área externa de paredes, valor U promedio de las paredes, valor U promedio de los techos y número de pisos. El método de conglomerados utilizado fue K-medias en combinación con una técnica jerárquica para subdividir la muestra en grupos no superpuestos mediante la asociación de edificios similares. El resultado del conglomerado es un cálculo matemático de sus centros expresados como los valores más representativos para cada una de las variables de datos.

Tanto en Barcelona como en Quito, el análisis de conglomerados arrojó resultados significativos para dos subgrupos demostrando que existen diferencias sustanciales en el comportamiento térmico de sus escuelas y, que los proyectos de rehabilitación deben adaptarse específicamente a cada tipología. Los resultados para Quito son: A) escuelas dispersas de 1 piso con techos inclinados y cubiertas metálicas y; B) Escuelas compactas de 2 plantas con losas planas de hormigón. Para Barcelona, los resultados son: A) Escuelas semi-compactas de 1 a 2 plantas con grandes

[1] Department of Architectural Technology, School of Architecture (ETSAB), Universitat Politècnica de Catalunya (UPC-Barcelona Tech), https://orcid.org/0000-0002-0625-2343; [2] Department of Physics, School of Architecture (ETSAB), Universitat Politècnica de Catalunya (UPC-Barcelona Tech), https://orcid.org/0000-0001-5949-2933; [3] Department of Architectural Technology, School of Architecture (ETSAB), Universitat Politècnica de Catalunya (UPC-Barcelona Tech), https://orcid.org/0000-0003-1747-8150; * Contact e-mail: maria.gabriela.ledesma@upc.edu

huellas y losas reticulares de hormigón y; B) Escuelas compactas, de más de 3 pisos y losas planas unidireccionales con entrevigado cerámico. Al momento, los edificios están siendo modelados en softwares especializados para evaluar su rendimiento energético. Los modelos virtuales incluyen los datos mencionados anteriormente sobre las características termo-físicas además de patrones de ocupación y operación recopilados de las escuelas reales más próximas a los centros estadísticos de sus subgrupos. Estas escuelas también están siendo monitoreadas durante un período de 3 meses para analizar su comportamiento térmico en operación pasiva. Los resultados del proyecto serán la base para avanzar en la eficiencia energética de la arquitectura educativa de estas ciudades y su futura renovación.

Abstract

The existent building stock is one of the most energy intensive sectors, accounting for approximately a third of the worldwide energy use. The analysis of a cities' built park requires the application of data mining in order to group buildings with homogenous properties thus creating architecture typologies. In the field of energy efficiency, clustering techniques have been used for evaluating the energy performance of an entire built stock by relying on reference buildings. These are real-life or theoretical edifices which characteristics are representative of the whole sample, and whose results can be generalized without incurring in significant errors.

This study analyses the educational built-park in order to establish a creation process based in statistical techniques for theoretical reference buildings in the cities of Barcelona-Spain and Quito-Ecuador. The overall aim is to generate archetype buildings for detailed energy analysis including their passive thermal behaviour and resilience to climate change. This study is part of a wider research on envelope rehabilitation of school's buildings through algorithm-based energy simulations for optimal passive treatment of their individual surfaces. Starting from the entire built-park of the cities it was possible to obtain the sample relevant characteristics and make a primary classification based on construction period and gross floor area. Clustering techniques were then used to subdivide the samples into subsets of similar buildings regarding their geometrical and thermal barrier properties. Finally, the mathematical calculation of the centroids of each subset of buildings were used to create the theoretical reference buildings.

The sample was limited to public schools in urban consolidated areas due to ease of access to information either on public domain or provided by the local Educational Departments. The primary classification was done following the methodology of the European TABULA project using 2-axis matrixes relating the construction period to the gross floor area (categories) and, allocating each school centre to a specific category. Information on the thermo-physical properties of the buildings was collected in standardized factsheets relying on information from cadastre, technical drawings and, GIS data. Considering the "passive architecture" focus of the research, the variables selected are: compactness, ground floor area, external wall area, average U-value of walls, average U-value of roofs and, number of floors. The statistical clustering method used was K-means in combination with a hierarchical technique as to subdivide the sample into non overlapping clusters by associating similar buildings together. The result of the clustering is a mathematical calculation of its centroids otherwise understood as the most representative values for each of the data variables. These values were used to create the reference building virtual models.

In both cases of Barcelona and Quito, the cluster analysis yielded significant results for two clusters showing that there are substantial differences in the thermal behaviour of their school built-park and that, retrofit interventions must be tailored specifically for each typology. The results in the case of Quito are: A) Disperse, 1-storey schools with pitch roofs and metal-cladding and; B) Compact, 2-storey schools with flat concrete slabs. In the case of Barcelona, the results are: A) Semi-compact, 1 up to 2-storey schools with large footprints and reticular slabs and; B) Compact, over 3-storey schools with relative small footprints and unidirectional slabs with ceramic interjoists. At present, the buildings are being modelled using dynamic simulation software as to assess their energy performance. The virtual models include the previous mentioned data on the thermo-physical properties of the buildings and, the occupation and operation patterns collected from the closest representative real-life school centre. These schools are also being monitored during a 3-month period on their thermal behaviour while in passive operation. The project results will be the foundation to move forward in the energy efficiency of these cities educational architecture and their future refurbishment.

Palabras Clave: edificios referenciales; análisis de conglomerados; rehabilitación de envolvente; simulación

Key words: reference buildings; clustering; envelope retrofit; simulation

1. Introduction

The existent building stock is one of the most energy intensive sectors -accounting for approximately a third of the worldwide energy use- fostering a large interest in the study of its energy performance and energy retrofit possibilities. The sustainable use of the built-park is one of the foundations of sustainable development and its main pillar is the energy refurbishment of existing buildings. On this note, the IEA in 2017 reported a rate of deep renovations below the 1% worldwide (International Energy Agency, 2017), which is far from reaching the UN Sustainable Development Scenario.

Comprehensive energy models of a cities built-park are needed to assess the effects of retrofit measures and to identify key improvement systems. The first step to develop a large building stock model is to classify the built park into typologies which renders a complete and in-depth knowledge of the sample set and its performance (Ballarini, Corgnati, Corrado, & Talá, 2011). Reference buildings (RBs) are models that act as building benchmarks (Li et al., 2018) of a significant percentage of the edifice stock and as such, permit the generalization of results. They are constructed based on statistically data and represent the average building characteristics of its typology considering functionality, geographic location and thermal exchange. Reference buildings have been in use for the last decade; the USA Department of Energy has defined RBs across several typologies, locations and construction periods; similarly, the EU back on 2010 required each member state to define RBs (Diario Oficial de la Unión Europea, 2010).

To date, most studies on building typologies for energy analysis deal with residential uses. In the EU, two recent projects have focus on creating residential typologies: 1) the TABULA (Typology Approach for Building Stock Energy Assessment) and, 2) the ASIEPI. As subsets of the TABULA project, studies have been made on cost-optimal analysis of energy retrofits (Corgnati, Fabrizio, Filippi, & Monetti, 2013), strategic comparisons against national energy statistics (Ballarini, Corgnati, & Corrado, 2014) and, development of methods for obtaining reference buildings of non-conditioned building stocks (Schaefer & Ghisi, 2016).

Despite the renovation of a 3% public building stock per year is one of the commitments of the EU (European Commission, 2018), there are no detailed research projects on their characteristics and retrofit possibilities. Public buildings have a key role as model examples of energy efficiency and sustainability; they have dissemination potential, have a defined function and operation patterns, and most importantly, they are under the control of local authorities facilitating the implementation of energy efficiency measures. Educational centres represent the 17% of the non-residential building stock in Europe (Österreicher & Geissler, 2016) and around 18% its energy consumption. Since educational centres are over 40 years old (on average), they do not comply with current building performance standards, which added-up to their non-stop consuming profiles due to inefficient active systems and management profiles makes them high energy consumers e.g. in Spain in 2014 educational buildings had a final energy consumption of 599Ktoe similar to that of hospitals (Liebana, Serrano, & Ortega, 2017).

Until now, the assessment of educational edifices has relied on particular case studies without an in-depth analysis of the built park. Arambula-Lara (2014) used cluster techniques and linear regressions to classify 59 schools in Treviso-Italy into groups based in optimal combinations of 6 thermo-physical parameters and; later on, selected real-life representative buildings as those closest to the cluster centroids (Arambula-Lara et al., 2015). Similarly, Marrone (2018) classified 80 Italian schools in Lazio into clusters using K-means method but relying on different parameters than the ones used in Treviso; and thus, rendering impossible a comparison of the results. More recently, Liebana (2017) made a first attempt at establishing educational typologies for the city of Valencia-Spain based on the methodology of the Tabula project and relying on the surface-to-volume ratio and building form; though, the reference buildings were selected on experts choice rather than on statistically data. On 2018, the University of Belgrade published a comprehensive study in the national typology of school buildings in Serbia (University of Belgrade, 2018) identifying 13 school types and calculating their current energy demands and its possible improvement.

Prior to a wider research on envelope rehabilitation of school's buildings through algorithm-based energy simulations for optimal passive treatment of their individual surfaces, this study aims to identify school typologies in two different socio-economic contexts using one unique approach and testing its validity on widely different samples. Here, we analysed the educational built-park of the cities of Barcelona-Spain and Quito-Ecuador and, create building typologies using statistical data of the thermo-physical features of the sample. The typologies will later be used to make a detailed energy analysis of the reference buildings, monitoring of their indoor conditions and to test the performance of passive retrofit measures. The structure of the next parts of this paper are as follows: the methodology explaining the workflow of the research, the results and discussion divided into two main parts and, the conclusions. The results section is divided in: 1) Data analysis for typological classification and, 2) Extrapolation of reference buildings; each of which is subdivided according to the methodology steps.

2. Methodology

The analysis of a cities' built-park requires the application of data mining and machine learning techniques in order to group buildings with homogenous characteristics into architectural typologies. In energy studies, clustering techniques are used to investigate the energy saving potentials of large stocks and to classify them according to their performance. This study is built-upon the methodology of the TABULA project (Ballarini et al., 2014) which makes a preliminary typological classification of the built-park based on the construction period and gross floor area of each element and then, extrapolates reference buildings for each category using statistical clustering. As stated before, RBs are real-life or theoretical edifices which properties are representative of the whole sample. As such, the use of reference buildings in energy modelling is both resource and time efficient.

Two main stages were developed: 1) Classification of the educational built park into typological categories and, 2) Creation of theoretical reference buildings for each prior-obtained category. For achieving the first stage three steps were required: a) Delimitation of the sample and database compilation, b) Analysis of the sample using descriptive statistics and, c) Typological classification. The second stage was composed of two steps: a) Clustering analysis using hierarchical techniques and, b) Selection of reference buildings for energy modelling. As to ease the reading, the specific methods used are explained in each subsection of the results chapter.

3. Results

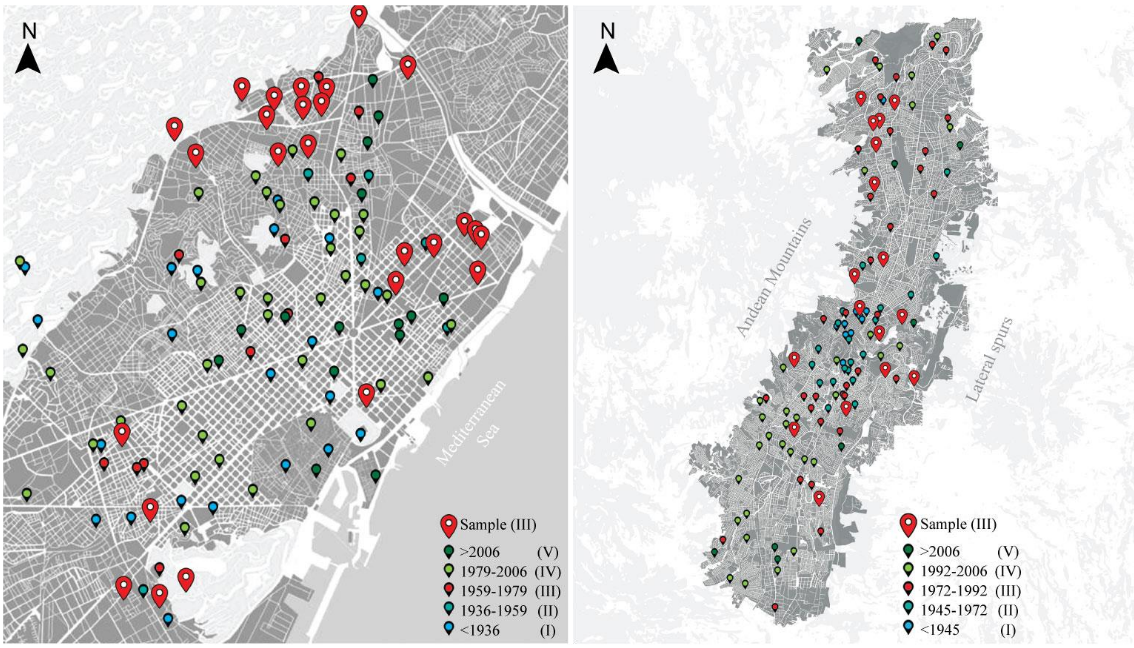
3.1 Data analysis for typological classification

a) Delimitation of the sample and database compilation

Educational buildings worldwide have been erected over a large time span and as such, have very different construction technologies, envelope materiality and internal distributions related to the pedagogic philosophy of their time. What is more, some education centres built at the beginning of the XX century and before are still in use and have been subject to punctual retrofit interventions on specific decayed building components. Considering the heterogeneity of the educational built set, the departure point was the construction of a reliable database with sufficient number of entries. The sample was limited to public schools in urban consolidated areas due to ease of access to information either on public domain or provided by the local Educational Departments (Ajuntament de Barcelona, 2017; Ministerio de Educación, 2018).

The cities of Barcelona and Quito were selected because of their significant differences in geographical and socio-economic contexts. Barcelona is the capital of Catalonia and it is the second most populated city in Spain. It has an area of 102.16 km^2 and a density of 16135 hab/km^2 (INE 2014); its climate is Mediterranean with relative mild winters and warm summers. Quito is the capital of Ecuador, a South-American country located on the Andean mountains. It is also the second most populated area of the country with a surface of roughly 372 km^2 and a density of 5312 hab/km^2 (INEC 2010). Its climate is equatorial highland with year-round average temperatures of 17°C and no seasonal variations. In both cases, the need of retrofitting its educational stock has been identified; according to Cuchí-Burgos (2014) schools in Spain have a potential energy saving of 10-30%, while in 2018 Quito's local government identify the need to make deep holistic renovations of its entire educational stock.

Figure 1. **a) Schools geographical distribution in Barcelona and, b) Schools geographical distribution in Quito**



(a) (b)

Source: own elaboration. Notes: Colours represent the construction time-periods. Sample schools correspond to the ones selected for detailed analysis.

The school's datasets created include ID number, contact information, number of students, construction year, gross floor area, executed retrofit interventions (year and component intervened), degradation state of the infrastructure and, historic listings and assigned category (if any). Additional information on the building structure type, external wall and roof materiality, number of floors, geometrical shape and other relevant information is also included in the dataset. The later parameters usually refer to the main building if the educational centre has complementary buildings and/or building additions; or, to the general morphology of the complex if the centre is comprised of modular units. Figure 1 shows the location and primary classification of the building sets in both cities regarding construction period; the "sample" schools correspond to the ones selected for detailed analysis regarding the predominant period.

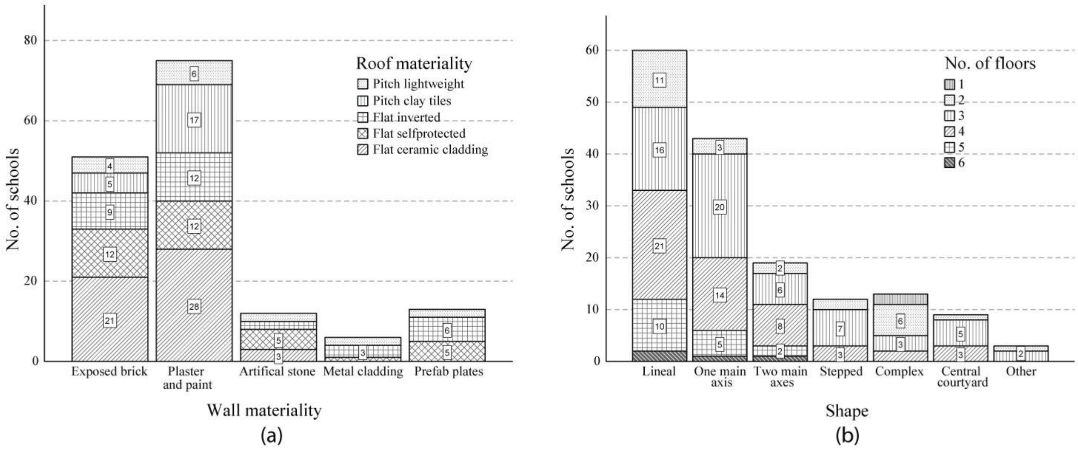
The dataset for Barcelona is formed of 165 entries of which 6 were discarded for being outside the scope of this research and 1 was discarded due to incomplete information. The evolution of the educational park in Barcelona is characterized by a uniform distribution of the equipment during the different time frames in exception for the period between 1959 and 1979 where the majority of schools were built in the periphery of the city; this situation responds to the migration movements and overpopulation of the industrial zones resulting in a lack of school places. The dataset in Quito is comprised of 268 schools. As most buildings house two or three different institutions working on different schedules, the final dataset is comprised of 130 building complexes of which, 7 were excluded due to incomplete data. As can be seen in Figure 1, until the 1970's schools were stacked on the historical district of the city due to an economic segregation by which the upper classes settled on the north part of the city fostering the creation of private education centres. During the dubbed "oil-boom period" (1972-1992) schools were constructed in all the territory alike experimenting its highest growth in number. In the following decades and up to this date, new facilities are constructed regarding the increase in population and the economic class of the districts, favouring low-income neighbourhoods.

b) Analysis of the sample

The collected data on shape, height and envelope materiality of the schools was analysed using descriptive statics. Stacked bars were constructed correlating wall to roof materiality and shape to height (number of floors above ground). Figure 2 and 3 show the data of the schools in both cities. Barcelona's building population has a wider array of envelope materiality and volumetric characteristics than buildings in Quito; however, there is a clear tendency for lineal low-rise (3 to 4 storeys) buildings with brick walls and flat ceramic cladding roofs. Regarding the wall materiality, the differences are based mostly on the external rendering of the envelope than in its principal component and as such a more detailed analysis regarding construction parameters was deemed necessary; this analysis was carried on in a later stage. The roofs are mainly flat unidirectional slabs with a concrete compression layer, the existence of thermal insulation is the main difference between roof constructions. Lastly, school buildings are shaped in lineal simple models following one main construction axis (rectangular, L, T, Z, C shapes); the other shapes can be categorized as more complex structures that respond specifically to the characteristics of the urban surroundings and architectural design considerations.

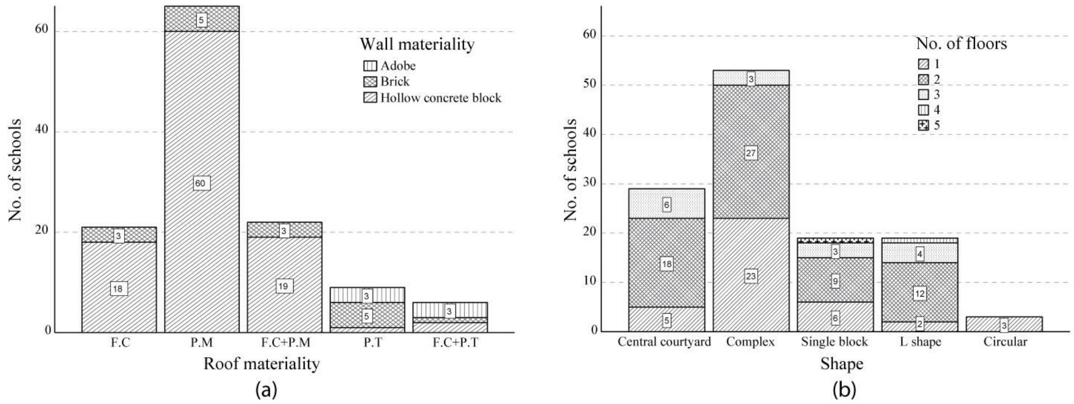
The sample in Quito is more homogeneous than the one in Barcelona responding to the one-leaf envelope construction tradition and the relative low-to-none use of thermal insulation materials. The low number of categories and the predominance of some features makes it easy to establish a first school typology based in small 1 up-to 2 storeys modular units with envelopes formed of hollow concrete block walls and lightweight pitch roofs with metal cladding. Larger school complexes are usually the combination of one main building (2 up-to 3 storeys) and small modular units added up according to increments in the number of students; in this cases the main building can be rectangular or follow a lineal configuration around a central courtyard. Additionally, the main building is of more stern construction than the modular units and their structure is generally a concrete frame with bidirectional reinforced concrete slabs.

Figure 2. **Descriptive data of the schools in Barcelona**



Notes: a) Correlation of wall to roof materiality; b) Correlation of the building shape to the number of floors
Source: Own elaboration.

Figure 3. **Descriptive data of the schools in Quito**



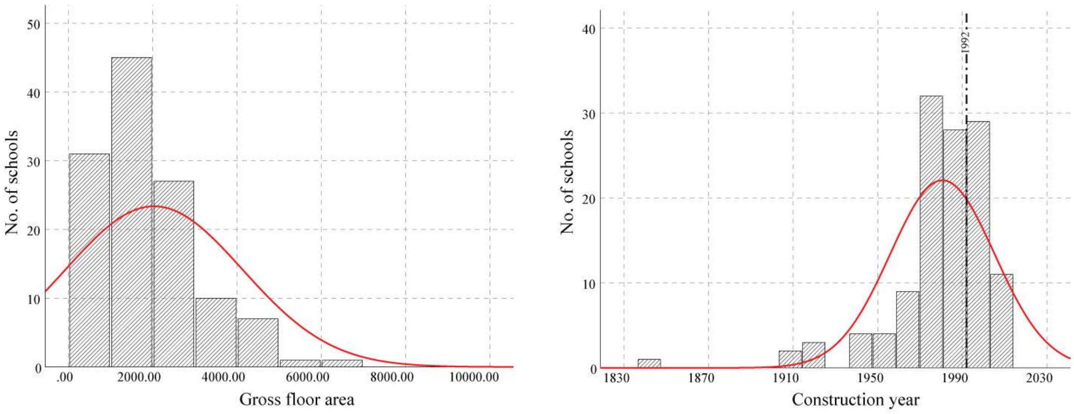
Notes: a) Correlation of the wall to roof materiality; b) Correlation of the building shape to the number of floors. FC (flat concrete slab), PM (Lightweight pitch roof), FC+PM (flat slab on main building and pitch roofs on secondary buildings), PT (pitch roof with clay tiles), FC+PT (pitch roof with clay tiles on main building and flat slabs on secondary buildings).
Source: Own elaboration.

c) Typological classification

According to the variables used for primary classification in the TABULA project the gross floor area and construction periods were analysed using histograms. For gross floor area, intervals of one thousand square meters were used; while for construction year, intervals of 20 and 10 years were used for Barcelona and Quito respectively. The difference in time intervals responds to the larger time span between building constructions in Barcelona with examples dating to the second half on the XIX century.

There is a predominance of medium size schools in Barcelona with gross floor areas ranging from 2000 to 4000 m^2 (see Figure 4a). There are few examples of edifices with areas above the 5000m^2 and none under the 1000m^2; this responds to the obligatory presence of complimentary and specialized spaces. Additionally, larger edifices are associated with institutes (school and middle-school) more than with schools. Figure 4b shows the distribution of schools regarding their creation time line; the largest construction campaign takes place during the 60's and 70's. The year 1979 has been set as benchmark due to the implementation of the first Spanish energy code which utterly changed the traditional architectural and construction style of the school buildings. Though a significant part of the built park dates after 1979, these buildings are assumed to fulfil the thermal requirements of their time; and hence, perform better than those prior to them.

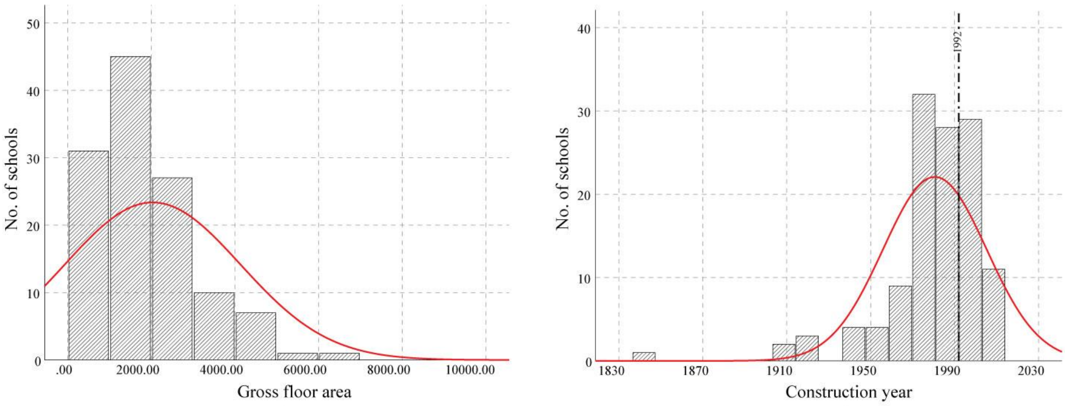
Figure 4. **Distribution of school buildings in Barcelona**



(a) (b)

Note: according to a) Gross floor area and, b) Construction year
Source: Own elaboration.

Figure 5. **Distribution of school buildings in Quito**



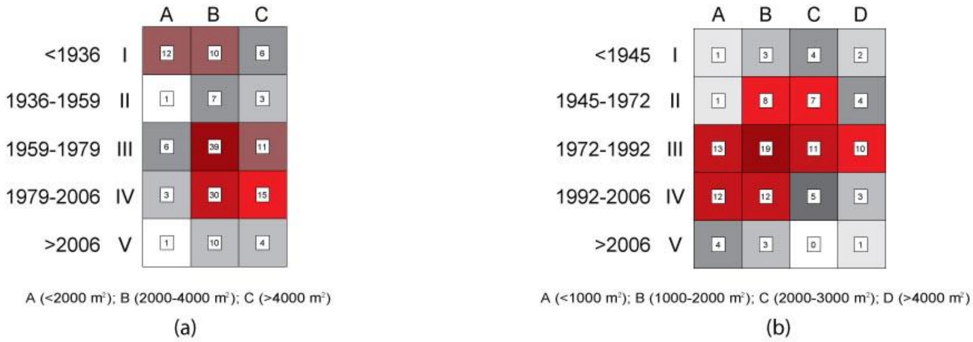
(a) (b)

Note: according to a) Gross floor area and, b) Construction year
Source: Own elaboration.

Contrary to Barcelona, schools in Quito are smaller in size due to architectural programs only considering classroom and service spaces without specialized areas. Additionally, many edifices designed for housing kinder gardens have been adapted to act as schools despite space limitations. Larger schools are usually the result of adding modular classrooms units to assist their main buildings. Though there are concrete examples of schools dating to the 1840's, most current working centres were constructed from the beginning of the century. The largest growth in number occurred during a 30-year period from the 1970's as a result of the continuous increment in population due to accelerated rural migration.

The histograms served as basis for the typological classification using both variables distribution and frequencies as subdivision categories. Milestones for the time periods are economic or political relevant events for each country and the appearance of thermal regulations (Crespo, 1999; Liebana, 2017). The results are 2-axis matrixes that relate construction period to gross floor area and allocate each school centre to a specific category (see Figure 6). Categories with representativeness below 2% of the sample were disregarded. Clustering techniques are then applied to each category as to obtain representative buildings. This typological classification is needed due to the large influence of both variables on clusters formation.

Figure 6. **Typological classification of school buildings according to gross floor area and construction date**



Note: a) Schools in Barcelona and b) Schools in Quito
Source: Own elaboration.

3.2 *Extrapolation of reference buildings*

Mining techniques were used for extrapolating reference buildings in each of the typological categories by finding significant subgroups within the sample. These techniques are exploratory and thus, the results depend on the data variables selected and its correlation to the study objectives. Considering the passive architecture focus of the research, the variables selected reflect the properties of the building envelope as thermal barrier. Based on (Arambula Lara et al., 2014), whom found that combinations up-to 6 variables are optimum for clustering analysis and, in studies of school building's typologies (Arambula Lara et al., 2015; Marrone et al., 2018) the characterization variables selected were: compactness (SV), ground floor area (Agf), external wall area (Aw), U-value of walls (Uw), U-value of roofs (Ur) and, number of storeys (F). Variables like building orientation, urban microclimate and urban typology were not taken into account for the clustering since the main objective of the research is creating building archetypes. In a later stage, the archetypes will be tested regarding the aforementioned parameters to analyse their adequacy to the various urban conditions of each city.

The technical properties of the buildings were obtained from cadastre data, technical drawings, scientific publications, photographs, on site visits and data from local authorities. This information was collected in individual factsheets for each school. The thermal transmittance (U-values) of the building envelope was calculated according to ISO 6946 and ISO 13370, the hygro-thermic properties of the materials were obtained from ISO 10456 and from the catalogue of construction elements of the Spanish Technical Building Code. When the available information on the building was insufficient to calculate its thermal properties, the buildings were assumed to follow the construction characteristics of their time period.

The correlation between the characterization variables was reviewed as to not generate a bias in the results. It was expected that the compactness – defined as the surface-to-volume ratio- would be the only significant correlation to the other variables. The Pearson correlation with a significance level of ρ=0.01 was used. For both study cases the compactness had significant values in relation to the ground floor area (Agf) and the number of storeys (F); this is due to the surface area being one of the composing terms of the compactness ratio; however, this ratio is of great importance in the potential heat balance of the building and thus could not be excluded from the analysis. A significant relation was also found between the area and the number of storeys as lower ground areas resulted in higher rise-buildings. This relation, however, served for the post association of atypical buildings.

As the studied variables have multiple units, statistical standardization using Z-scores was used (see Equation 1). This process removes the influence of the degree of dispersion of each variable by expressing the data in terms of the number of standard deviations from their means. Following this, the data was checked for the presence of multivariate outliers, considered as anomalous or incoherent observations in regard to the entire data set. The cumulative probability in a chi-square distribution of the Mahalanobis distance was used as the outlier measure. The Mahalanobis distance is a measure of the distance between each data in a variable to their mean distribution and, it accounts for the variance of each variable and the covariance between variables (see Equation 2). The probability (ρ-value) then is calculated as the cumulative distribution function for 6 degrees of freedom –the number of variables used- (see Equation 3). Values of ρ<0.005 were considered outliers.

$$Z_X = \frac{(x_i - \bar{x})}{\sigma_x} \qquad \text{Equation 1}$$

Where Z_X is the standardized value; x_i is the data variable; $\bar{x}$ is the variable mean and, $\sigma_x$is the variable standard deviation.

$$X_i = [Z_{SV_i}\, Z_{Agf_i}\, Z_{Aw_i}\, Z_{Uw_i}\, Z_{Ur_i}\, Z_{F_i}]^T \qquad \text{Equation 2}$$

$$d(X,\bar{X})_i = \sqrt{(X_i - \bar{X}_i)^T S^{-1}{}_i (X_i - \bar{X}_i)}\ \forall X_i \in \mathbb{R}^{6\times 1}, S \in \mathbb{R}^{6\times 6}$$

$$\rho_i = 1 - CFD[d(X,\bar{X})_i; DF] \qquad \text{Equation 3}$$

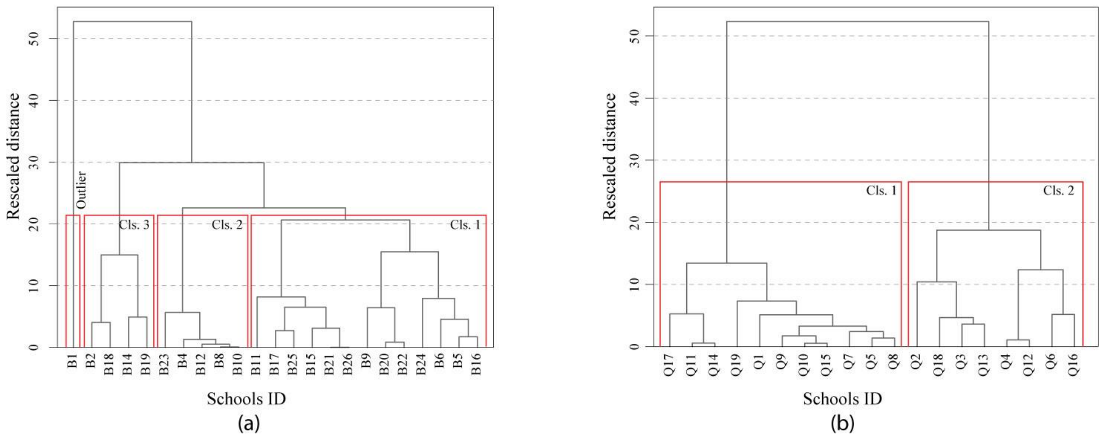
Where X_i is the variables vector for each school i; $d(X,\bar{X})_i$ is the Mahalanobis distance; $\bar{X}_i$ is the variables mean vector; S_i is the covariance matrix; ρ_i is the probability value; CFD is the cumulative function distribution, and DF are the degrees of freedom.

Clustering is a multivariate data analysis which groups closely related individuals together and then, characterizes each cluster based on its centroid value. After the data had been processed, a hierarchical clustering using Ward's method and the squared Euclidean distance was performed to obtain the number of clusters *k*. Hierarchical clustering uses a tree like structure as to facilitate the graphical understanding of the clusters formation; in each stage the closest individuals are paired together until all elements are grouped as a whole. The number of clusters *k* served as input in the k-means algorithm which relies in mathematical calculations to determine the cluster centroids.

For exemplification of the creation of reference buildings, the most representative typological categories of school buildings in both cities were selected: Cat B.III with 39 schools for Barcelona and, Cat B.III with 19 schools for Quito. The detailed information of these education centres was collected on the previous mentioned factsheets and, considering this additional information some schools in Barcelona were discarded from the sample. Buildings which had been integrally refurbished or, which envelope had been updated to current energy saving standards were removed from the sample; similarly, buildings flagged for demolition by the local authorities were also removed. The final sample was of 23 schools for Barcelona; the number of schools in Quito remained unaltered as none matched the previous characteristics. As to maintain confidentiality all school names have been replaced for identification codes.

None outliers were detected in the samples; however, one school in Barcelona had a ρ=0.009 greatly differing from the next ρ value of 0.124 and above. As this school cannot be removed as a statistical outlier it was expected that the clustering analysis would signal it as an independent cluster. The dendrograms (hierarchical clustering) were constructed using R software (see Figure 7). The number of clusters were obtained from the dendrograms by cutting a line through the largest rescaled distance, representing the largest variability between clusters centroids. As expected, in the Barcelona sample, school B1 is assigned to an independent cluster, other three clusters are formed being cluster 1 the largest with 13 elements and; cluster 3 the more compact one with only 4 elements.

Figure 7. **Hierarchical clustering dendrograms using Ward's method**



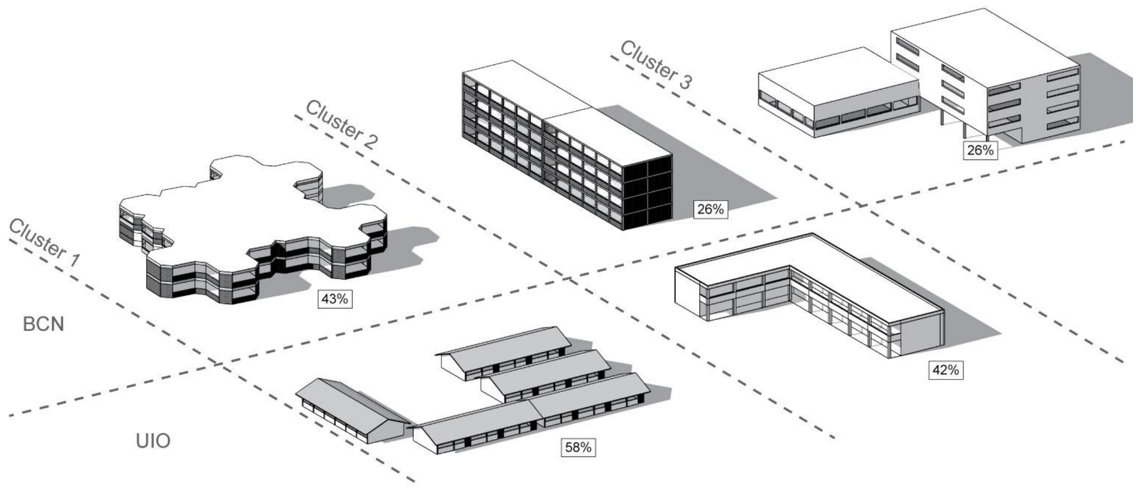
Note: a) Barcelona and, b) Quito
Source: Own elaboration.

In Quito the minimum ρ-value was of 0.09 showing that none of the schools are substantially different from the sample mean. Figure 7.b shows that two clusters are significantly different from each other and have also a good level of similarity between their composing elements. The clusters are formed of 11 and 8 elements respectively making both clusters significant for the studied sample. In the case of Barcelona only cluster 1 is significant for the sample with a representativeness of 56.52% since clusters below 40% of the sample are not considered representative.

K-mean algorithm was also implemented in R software with inputs k=4 and k=2 for Barcelona and Quito respectively. The advantage of this algorithm over the hierarchical clustering is that the members of the clusters are constantly changing until the largest difference between centroids occurs; while, in the hierarchical method objects assigned together at the first stage will remain together along the process. The calculations were performed iteratively until no variations in the cluster's membership occurred. The number of elements in each cluster differed from the results of the hierarchical clustering due to this iterative process. The centroids were then calculated as the variables mean values for each of the clusters. These values were used for generating theoretical virtual models of the buildings which exemplify the properties of their clusters. The fictional models were based on the closest real life school building to the cluster's centroid and altered accordingly to best fit the centroid values. Figure 8 displays the axonometric views of the extrapolated reference buildings and the representativeness of the clusters to the entire sample.

In particular, the case of Barcelona shows cluster 2 and cluster 3 with the same representativeness; however, an analysis of the distance of the elements of cluster 3 to its centroid showed that all elements are distant from it. This reflects the heterogeneity of cluster 3 and thus, suggests that no significant characteristics from the elements can be inferred. In consideration of this cluster 3 is not considered representative of schools in Barcelona.

Figure 8. **Theoretical reference school buildings for the cities of Barcelona and Quito**



Source: Own elaboration.

4. Conclusions

The combination of typological classification and clustering techniques are adequate to deal with large samples of different features. Considering the exploratory character of the techniques employed, special attention must be placed in the selection of the characterization variables. In view of the passive focus of this research, the variables selected deal with the thermo-physical properties of the buildings; and therefore, the resulting reference buildings are constructed based on them. In both cases of Barcelona and Quito, the cluster analysis yielded significant results for two clusters showing that there are substantial differences in the thermal behaviour of their school built-park and that, retrofit interventions must be tailored specifically for each typology.

The results in the case of Quito are: A) Disperse, 1-storey schools with pitch roofs and metal-cladding and; B) Compact, 2-storey schools with flat concrete slabs. In the case of Barcelona, the results are: A) Semi-compact, 1 up to 2-storey schools with large footprints and reticular slabs and; B) Compact, over 3-storey schools with relative small footprints and unidirectional slabs with ceramic interjoists. At present, the buildings are being modelled using dynamic simulation software as to assess their energy performance. It is not expected that the reference buildings have the same performance as all the buildings in the cluster but that its thermal performance be similar to those buildings when exposed to the same external conditions.

The final aim is the creation of archetype building libraries for detailed energy analyses. The virtual models include the previous mentioned data on the thermo-physical properties of the buildings and, the occupation and operation patterns collected from the closest representative real-life school centre. These schools are also being monitored during a 3-month period on their thermal behaviour while in passive operation. The project results will be the foundation to move forward in the energy efficiency of these cities educational architecture and their future refurbishment.

Acknowledgements: The authors would like to thank the Ecuadorian National Secretary of Higher Education, Science and Technology (SENESCYT) for awarding a research scholarship to Gabriela Ledesma (CZ02-000592-2018). The co-author Oriol Pons Valladares is a Serra Húnter Fellow.

Author contributions: The first author performed the research, analysed the data and wrote the paper with inputs from the other authors. The second and third author review, edited and approved the final manuscript.

Conflict of interest: The authors declare no conflict of interests

References

Ajuntament de Barcelona. (2017). Enseñanza secundaria. Centros y profesores. Cursos 2013-2017. Retrieved November 8, 2018, from http://www.bcn.cat/estadistica/castella/dades/anuari/cap05/C0502060.htm

Arambula Lara, R., Cappelletti, F., Romagnoni, P., & Gasparella, A. (2014). Selection of Representative Buildings through Preliminary Cluster Analysis. In *International High Performance Buildings Conference*. Retrieved from http://docs.lib.purdue.edu/ihpbc/137

Arambula Lara, R., Pernigotto, G., Cappelletti, F., Romagnoni, P., & Gasparella, A. (2015). Energy audit of schools by means of cluster analysis. *Energy and Buildings*, *95*, 160–171. https://doi.org/10.1016/j.enbuild.2015.03.036

Ballarini, I., Corgnati, S., Corrado, V., & Talá, N. (2011). Improving energy modelling of large building stock through the development of archetype buildings. In B. Soebarto (Ed.), *Building Simulation* (pp. 14–16). Sidney: IBPSA Australasia.

Ballarini, I., Corgnati, S. P., & Corrado, V. (2014). Use of reference buildings to assess the energy saving potentials of the residential building stock: The experience of TABULA project. *Energy Policy*, *68*, 273–284. https://doi.org/10.1016/j.enpol.2014.01.027

Corgnati, S. P., Fabrizio, E., Filippi, M., & Monetti, V. (2013). Reference buildings for cost optimal analysis: Method of definition and application. *Applied Energy*, *102*, 983–993. https://doi.org/10.1016/j.apenergy.2012.06.001

Crespo, P., & Ortiz, C. (1999). Aportes para una historia de la educación municipal en Quito. *ProcesoS, Revista Ecuatoriana de Historia*, *I*(13).

Cuchí Burgos, A., & Sweatman, P. (2014). *Informe Gtr 2014. Estrategia para la rehabilitación*.

Diario Oficial de la Unión Europea. (2010). *Directiva 2010/31/UE del Parlamento Europeo y del Consejo, de 19 de mayo de 2010, relativa a la eficiencia energética de los edificios*. Retrieved from https://www.boe.es/doue/2010/153/L00013-00035.pdf

European Commission. (2018). Energy Efficiency. Retrieved November 8, 2018, from https://ec.europa.eu/energy/en/topics/energy-efficiency

International Energy Agency. (2017). Energy Technology Perspectives 2017, 443. https://doi.org/10.1787/energy_tech-2017-en

Li, X., Yao, R., Liu, M., Costanzo, V., Yu, W., Wang, W., … Li, B. (2018). Developing urban residential reference buildings using clustering analysis of satellite images. *Energy and Buildings*, *169*, 417–429. https://doi.org/10.1016/j.enbuild.2018.03.064

Liebana, E., Serrano, B., & Ortega, L. (2017). Typological analysis of school centres to characterize the energy consumptions . The case of the city of Valencia . In *International Congress on Sustainable Construction and Eco-Efficient Solutions* (pp. 415–426). Sevilla: Universidad de Sevilla. Retrieved from https://idus.us.es/xmlui/handle/11441/59088

Marrone, P., Gori, P., Asdrubali, F., Evangelisti, L., Calcagnini, L., & Grazieschi, G. (2018). Energy benchmarking in educational buildings through cluster analysis of energy retrofitting. *Energies*, *11*(3), 1–20. https://doi.org/10.3390/en11030649

Ministerio de Educación. (2018). AMIE (Estadísticas educativas a partir de 2009-2010). Retrieved November 13, 2018, from https://educacion.gob.ec/amie/

Österreicher, D., & Geissler, S. (2016). Refurbishment in Educational Buildings - Methodological Approach for High Performance Integrated School Refurbishment Actions. *Energy Procedia*, *96*(October), 375–385. https://doi.org/10.1016/j.egypro.2016.09.163

Schaefer, A., & Ghisi, E. (2016). Method for obtaining reference buildings. *Energy and Buildings*, *128*, 660–672. https://doi.org/10.1016/j.enbuild.2016.07.001

University of Belgrade. (2018). *National Typology of School Buildings in Serbia*. (M. Jovanović Popović & D. Ignjatović, Eds.), *Deutsche Gesellschaft fur Internationale Zusammenarbeit (GIZ)* (Vol. 1). Belgrade. Retrieved from http://eeplatforma.arh.bg.ac.rs/publikacije?tab=0