Building stock categorization for energy retrofitting of historic districts based on a 3D city model



Categorización de edificios basada en un modelo 3D de la ciudad para abordar la reconversión energética de distritos históricos

Iñaki Prieto, Jose Luis Izkara, Aitziber Egusquiza

TECNALIA. División de Construcción. Parque Científico y Tecnológico de Bizkaia, Geldo Auzoa, 700 Edificio - 48160 Derio (España)

DOI: http://dx.doi.org/10.6036/8147 | Recibido: 08/09/2016 • Evaluado: 12/09/2016 • Aceptado: 23/12/2016

RESUMEN

- La mejora de la eficiencia energética de los edificios es uno de los aspectos clave para la reducción de emisiones de contaminantes y la mejora de la sostenibilidad a escala mundial. La rehabilitación energética de edificios existentes está siendo una de las apuestas clave para la revitalización del sector de la construcción en Europa. Abordar dicha rehabilitación a escala urbana supone importantes beneficios frente a la rehabilitación individual de edificios. Sin embargo también lo convierte en una tarea más compleja, para lo cual se hace necesario el uso de herramientas que faciliten la identificación de las mejores estrategias de rehabilitación sobre datos objetivos. Cuando hablamos de entornos históricos es necesario abordar el problema desde planteamientos específicos que tengan en cuenta las peculiaridades de los edificios contenidos en ellos. En este contexto la solución presentada en este artículo propone la identificación de las principales tipologías de edificios de un entorno urbano y la selección de un edificio representativo de cada tipología para simplificar la recogida de datos y utilizar los datos de los edificios representativos como base para los cálculos totales teniendo en cuenta la representatividad de dicha tipología en el distrito. El proceso de categorización del parque edificado se detalla paso a paso en el artículo, así como su implementación en una herramienta software que automatiza algunas de las actividades de cada paso. La validación se realiza mediante la implementación en el distrito histórico de Santiago de Compostela.
- Palabras Clave: Eficiencia energética, distritos históricos, ciudades 3D, tipologías de edificios, software de categorización.

ABSTRACT

Improving energy efficiency of buildings is one of the key areas for reducing emissions of pollutants and improving global sustainability. Energy rehabilitation of existing buildings is currently one of the key challenges for the revitalization of the construction sector in Europe. Addressing such rehabilitation at urban scale versus individual rehabilitation of buildings has important benefits. However it makes the task much more complex, for which the use of tools to facilitate the identification of the best strategies for rehabilitation based on evidences is necessary. When we talk about historic districts it is necessary to address the problem from specific approach taking into account the peculiarities of the buildings contained therein. In this context the solution presented in this paper proposes the identification

of the main typologies of buildings in the district and the selection of a representative building of each typology to simplify data collection process and use the data from the representative buildings as a basis for total calculations taking into account the representativeness of each typology in the district. The building stock categorization process is detailed step by step, and its implementation in a software tool that automates some of the activities of each step. Validation is carried out through the implementation in the historic district of Santiago de Compostela.

Key Words: energy efficiency, historic districts, 3D cities, building typology, categorization tool.

1. INTRODUCTION

Buildings are responsible for 40% of energy consumption and 36% of CO2 emissions in the European Union¹. The construction industry is continuously searching for new retrofitting solutions and technologies in order to increase the energy efficiency of buildings. However, the majority of current development regarding energy efficiency focuses on new building work, ignoring the specific issues of existing buildings in general, and historic building in particular. Moreover, addressing the problem at district scale provides significant benefits from the building scale approach (e.g. economy of scale allows reaching cost-effective improvements, overcome barriers and restrictions applied to individual buildings and optimization of energy consumption and CO2 emissions associated with urban infrastructures). Current approaches for urban retrofitting projects are based on the use of energy simulation software (e.g. Energyplus, TRNSYS, IES VE). These tools require a huge amount of data in order to get precise and accurate results [1]. A good balance between the accuracy of the results and the availability of data needs to be addressed.

The work described in this publication is focused on the identification of building typologies as a previous step to the selection of energy retrofitting strategies. The identification of building typologies allows estimating the energy consumption of a set of buildings from the results of a representative building of this typology. This approach provides a drastically reduction of the need for collecting information. At the same time, it allows identifying the impact of the intervention in a specific typology within the whole district taking into account the representativeness of the typology.

¹ https://ec.europa.eu/energy/en/topics/energy-efficiency/buildings

The article is structured as following: First the building stock categorization process is detailed. In order to assist on the identification of building typologies, a web tool based on an urban 3D model has been developed and described here. Finally, the article presents the results of the application to the case of the historic district of Santiago de Compostela. The work described in this paper has been developed within the European project EFFESUS².

2. MATERIALS AND METHOD

2.1. MODELLING STRATEGIES FOR ENERGY RETROFITTING IN URBAN DISTRICTS

Urban modelling defined by [2] as the representation of functions and processes which generate urban spatial structure embodied in computer programs to enable location theories, is a necessary tool for planning and development of policies [3] as they can support from the design of policies to the implementation of specific strategies. In order to be effective and feasible, the urban modelling has to fall between the search for simplicity in articulating the urban structure and the need to translate their complexity [2]. Different modelling techniques have been considered for energy assessment:

- Large scale energy assessment: Nouvel et al. in [4] describes an automatic large-scale assessment of building energy behaviour based on a 3D city model using data available from public sources. It is based on the hypothesis that strong correlations exist between information about residential buildings (e.g. location, use, age, size) and their energy consumption [5].
- Identification of priority areas: This approach first identifies priority areas for energy efficiency interventions and, then focuses data acquisition on the buildings within this area [6]. Based on a low level of information, blocks or groups of buildings can be prioritized according to their conservation state, a rough estimation of their energy behaviour and the concerns of the citizens.
- Modelling through representative buildings: One procedure used to analyse large building stocks is to examine and categorise the stock according to some parameters (e.g. climatic zone, dimension, age). The representative buildings of each category could be real buildings (sample buildings) or ideal buildings constructed with the statistical characteristics identified as representative (archetype buildings). Parekh et al. in [7] describes the process of developing archetypes for energy simulation and Mata in [8] applies the sample building modelling strategy for energy, carbon, and cost assessments.

Creating models of whole districts based on the identification of sample buildings has been selected as the modelling strategy for energy retrofitting of historic districts in the EFFESUS project. The sample building strategy allows starting with basic information about all the buildings of the district in order to identify homogeneous groups of buildings. But, after performing the clustering process, more detailed semantic information (e.g. heritage significance of building elements or U-value of building envelope) can be acquired just for those representative buildings and extrapolate the results.

In order to cluster the buildings of a historic district for energy

² http://www.effesus.eu/

retrofitting, two relevant aspects have to be combined; on the one hand the energy profile of the buildings, and on the other hand, the constraints imposed by the historic value of the building and its elements (e.g. windows, façades, and doors). Previous attempts for clustering of buildings have been focused on the energy side of the buildings (e.g TABULA project³), but merging both aspects is what allows for the adequate clustering in order to model the historic districts for energy retrofitting purposes.

2.2. 3D CITY MODEL

A 3D city model is a georeferenced digital representation of objects, structures and phenomena that correspond to a real city [9]. CityGML4 is an open data model defined by the Open Geospatial Consortium (OGC) for the storage and exchange of a 3D city model. It has been designed to store semantic and 3D multiscale geometric information, considering urban and building scales. Semantic 3D city models have demonstrated being of great value for the prediction of energy demand of buildings, increasing accuracy and easiness compared to the 2D GIS models [10]. The implementation of the modelling strategy through sample buildings defined in this paper is based on a semantic 3D city model.

A 3D city model of the historic district of Santiago de Compostela is used as a material for the implementation of the work described in this paper. The city model includes: (1) a textured digital elevation model of the urban area of Santiago de Compostela and its surroundings; (2) geometry of buildings in the historic district in LoD2 (represented by independent facades and roofs); (3) CityGML data model extended with required parameters for the domain of energy retrofitting of historic buildings. The list of the data sources used for the generation of the 3D city model is:

- Building Footprints from Spanish Cadaster
- Digital Terrain Model (DTM) from Spanish National Geographic Institute (SNGI)
- Orthophotographs from SNGI
- LiDAR from SNGI

Some of the required attributes are already included as part of the CityGML core data (e.g. main use, year of construction, number of floors), while some others need to be added to the data model (e.g. area, level of protection).

3. BUILDING STOCK CATEGORIZATION PROCESS

This section describes the building stock categorization process defined in the EFFESUS project. The final objective of this process is to select a limited number of sample buildings that reflect almost the entire building stock of the historic district within the constraints of the data available. The number of selected typologies must be the minimum, but covering the maximum representativeness of the buildings stock.

For the identification of the sample buildings a set of relevant parameters have to be selected, taking into account the two relevant aspects previously mentioned (i.e. energy and heritage). The selected parameters need to be easy to acquire for all the buildings in the district. For each parameter a group of intervals should be defined in order to identify categories. The combination of different parameters and intervals will generate the typologies. Only those typologies with a significant number of buildings will be selected as the relevant ones. For each of the relevant typologies

³ http://episcope.eu/building-typology/

⁴ CityGML: OGC City Geography Markup Language (CityGML) Encoding Standard 12-019 - http://www.opengeospatial.org/standards/citygml

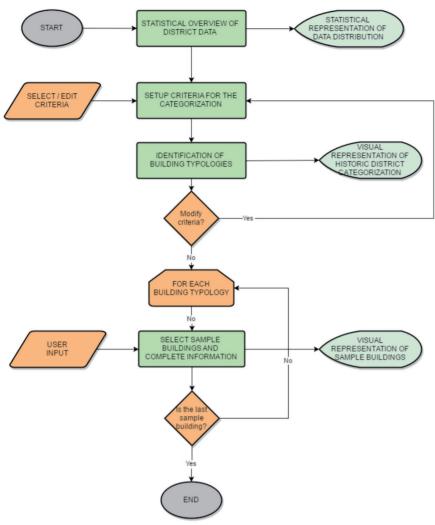


Figure 1: Workflow of the building stock categorization

a building will be identified as the representative of the whole typology.

The process is divided into 4 main steps:

- **1. Statistical Overview of the district data:** Assessment of the data included in the 3D City Model for the identification of those parameters which are relevant for the categorization.
- **2. Set-up Criteria for Categorization:** Selection of the parameters, intervals and threshold that will be used for the generation of typologies.
- **3. Identification of buildings typologies:** Identification of the most representative building typologies according to the selected criteria. If the number of typologies and their representativeness are not the expected ones, it is possible to go back to the previous step and modify the criteria.
- **4. Select Sample Buildings and complete information:** Selection of one building as representative of each typology and introduction of detailed information in the 3D City Model, for each representative typology.

Figure 1 shows the workflow that summarizes the different steps of the building stock categorization process. This workflow has been used for the implementation of the process in a software tool, which is described in Section 4.

3.1. STATISTICAL OVERVIEW OF THE DISTRICT DATA

The set of appropriate parameters and intervals for the identification of the building typologies will depend strongly on the his-

tory and location of the historic district. An expert judgement will be necessary to analyze the key parameters that influence the clustering process. As support for this process, several statistical methods and tools can be used for the representation of the distribution of the different parameters in building stock analysis. Frequency histograms have been previously used in the literature as in [11], in order to identify concentrations of particular values. Box-plot is used to graphically depict groups of numerical data through their quartiles. A histogram combined with a box-plot provides a clear understanding of the normality, tendency and dispersion of the data [12].

3.2. SETUP CRITERIA FOR THE CATEGORIZATION

Once the user has a clear vision of the distribution of each parameter, the criteria for the categorization will be selected. Criteria for categorization are represented by a set of parameters, the intervals for each parameter and the threshold. The threshold is the minimum percentage of buildings that allows a building typology to be considered representative.

For the categorization process it is important to consider that: (1) Typologies have to group the buildings regarding basically their energy characteristics and historic significance. (2) The clustering criteria have to be coherent with the criteria and constraints that later will be used for the identification of retrofitting strategies. (3) The required data have to be easily obtained. Following these considerations, the proposed list

of parameters for the categorization of historic districts regarding their energy performance is:

- Main use: mainly if it is residential or not, since their final energy consumption changes significantly (different use patterns and operating systems).
- **Number or facades**: it shows the "exposed area" of the building that is very important for the energy performance.
- Year of construction: buildings built in the same period have similar construction techniques. It can be used also, along with the protection degree, to assume homogenous heritage significance of the buildings.
- Protection degree: a direct indicator of the heritage significance and of the measures that can be applied.
- Volume: it represents the "heated volume" of the building that is also an indicator with a direct link to the energy performance.

The proposed method intends to be flexible regarding the used parameters; so parameters can be added or discarded if considered. Information provided by the volume and the number of façades are very much correlated, and both are direct indicators of the energy performance of the building. In order to reduce the number of typologies one of them could be removed from the final list of parameters.

3.3. IDENTIFICATION OF BUILDING TYPOLOGIES

Generation of the building typologies is the key step of the whole process. The number of parameters selected, the number of

intervals and the defined threshold will determine the success or failure of the categorization process. As an example: Five different parameters with three intervals for each can produce a huge number of different typologies making the analysis unmanageable ($3^5 = 243$ different typologies). Consequently, it is necessary to select the proper parameters, intervals for each parameter and also to discard the less representative groups through an adequate threshold.

The process starts selecting a first parameter and then proceeds to subdivide the typologies with respect to other parameters. By adjusting the number of intervals for each parameter the building typologies will be constructed (See Figure 2).

4 or 5 can be defined as the most suitable number of parameters to obtain a diversity of typologies but keeping them into a manageable number. The definition of the intervals is a bit more complex, due to the diversity of data type of each parameter (e.g. the use of the building will be residential or non-residential, the number of facades will vary from 1 to 4 and the year of construction will be a number between 1298 and 1996). Number of intervals for the categorization will vary from: 1, if only one of the possible values want to be included in the categorization; 2 to 4 if different intervals are defined for the possible values; or discrete if all the possible values will be considered for the categorization. The combination of the selected set of parameters and the corresponding intervals (number and limits) will determine the number of typologies generated.

When all the possible typologies have been identified the most representative ones have to be selected through the establishment of a threshold of minimum representativeness. A good balance between the number of typologies and the representativeness of them depends on the homogeneity of the building stock, the size of the district or the expert judgement of the user. However, a set of no more than 10 different typologies with a representativeness of at least 80% (usually obtained with a minimum threshold between 2% and 5%) could be considered as appropriate balance in most cases.

3.4. SELECT SAMPLE BUILDINGS AND COMPLETE INFORMATION

Once the building typologies have been identified and their statistical representation assessed, the next step involves selecting one representative building for each typology. This is a crucial decision as the results obtained with this building are going to

be extrapolated to the entire typology. The sample buildings can be selected by analyzing the building stock to find out a building showing characteristics similar to the mean value of geometrical or construction features of the statistical sample or by expert judgement. It is also important to consider that detailed information for the selected buildings needs to be available.

Once the sample buildings have been selected, the 3D city model will be completed with detailed information for the selected buildings. The geometric information contained in the model is already sufficient but the semantic information has to be completed regarding the two main topics: heritage significance and energy characteristics.

Regarding heritage significance, the necessary information is the historic value of the building components. This information could be obtained using the public data bases that documents the historic values of the district or by expert judgment if this data bases are not available or not of enough quality.

Regarding the energy characteristics (ratio of openings in the envelope, U values of building elements, glazing characteristics...), different methods are usually used to get this semantic information: public data bases, expert judgment, European databases (e.g. the one created by the TABULA project or others) or even Google street map as in [13].

4. BUILDING STOCK CATEGORIZATION TOOL

The process described above is implemented in a software tool that partially automatizes it and provides the user an easy and intuitive way for the identification of building typologies. The tool is based on data included in the city model described in Section 2.2. User can interact with the tool editing parameters, intervals and thresholds for categorization as well as selecting representative buildings and completing their properties.

4.1. USER INTERFACE

Graphical user interface of the categorization tool is divided into four main areas or panels (see *Figure 3*). Each panel includes several tabs which are enabled or disabled depending on the phase of the categorization process.

Categorization tool is accessible at the URL: http://3dcity.tecnalia.com/DynaCategorization/faces/DynaCategorization.xhtml?city=effesussantiago

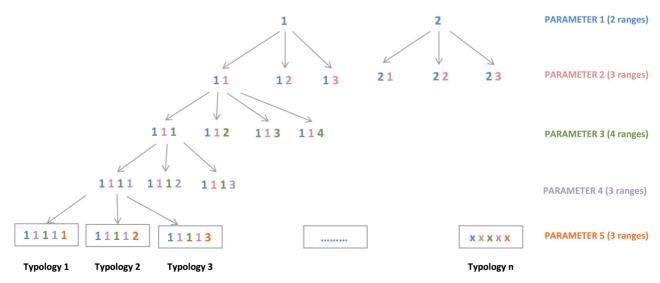


Figure 2: Generation of typologies

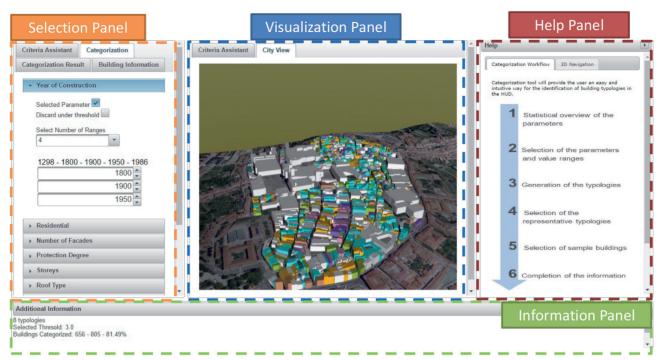


Figure 3: Graphical user interface of the categorization tool

4.2. CRITERIA ASSISTANT

For the statistical overview the categorization tool provides the criteria assistant functionality which aims to easily give an idea of the building data distribution in the whole historic district. Frequency histograms and box-plots for each parameter are automatically generated by the categorization tool (see Figure 4).

4.3. CATEGORIZATION

In order to setup parameters, intervals and threshold for the categorization, the categorization tool provides the user an automatic identification of suitable number of intervals and the upper and lower limits of the intervals for each parameter. The automatic setup of the number of intervals as well as the limits is based on an adaptation of the k-means algorithm widely used for clustering [14]. The complete algorithm is presented in Figure 5. The selection of parameters to be taken into account and the threshold value need to be manually provided by the user. Default intervals

identified by the tool can be manually edited by the user, in order to adapt the results to the user preferences and knowledge.

Intervals can be defined only for parameters with numeric values. In the case of parameters such as building use (e.g. residential, office or commercial) or roof type (e.g. flat, gable or hip) no intervals will be defined and all possible values will be considered in the categorization tool.

After all the possible typologies are generated according to the number of parameters and intervals, the threshold is applied in order to select the most representative typologies.

The automatic setup of the parameters and intervals for the categorization process analyses each parameter and performs the following steps. First, the data distribution of the parameter is analysed (line 8). If the parameter has less than 5 different values, all the values will be used in the categorization (line 10). Otherwise, different intervals are created depending data distribution. If the data distribution is homogeneous (each value has less than

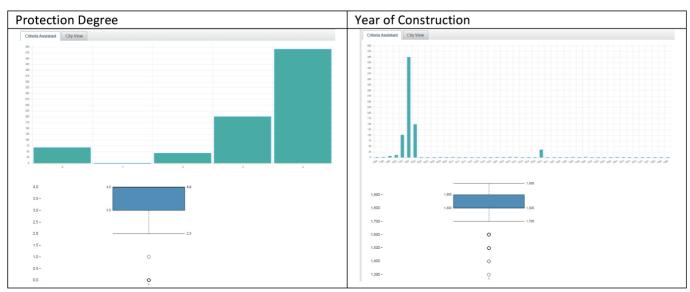


Figure 4: Visualization of parameter distribution in the categorization tool

```
1
     Algorithm Automatic Setup of the parameters and intervals for the categorization
     Input Categorization parameters
3
     Output List with intervals setup values for each categorization parameters
4
5
     var listOfIntervals
6
7
     for each parameter do
8
       Analyse data distribution of the parameter
9
10
       if (the number of different values of the parameter < 5) then
11
        Set interval number = All
12
       else
13
         Get the value that has more elements
14
         var intervalType = (Number of elements of the value that has more elements / Total number
15
     of elements in the parameter)*100
16
         if (intervalType< 20) then
17
          Set interval number = 4
18
19
         else if (intervalType>80) then
20
          Set interval number = 1
21
22
          Apply k-means clustering algorithm with 2 and 3 clusters
23
          Set interval number = The k-means result with less total distance between each value and
24
     its cluster centroid
25
        end if
26
       end if
27
28
       if (interval number = = 1) then
29
         Set listOfIntervals[1] = Select maximum value
30
       else if (interval number = = 2) then
        Set listOfIntervals[2] = Higher value of the first cluster
31
32
       else if (interval number = = 3) then
33
         Set listOfIntervals[31] = Set first interval with the higher value of the first cluster
34
         Set listOfIntervals[32] = Set second interval with the higher value of the second cluster
35
       else if (interval number = = 4) then
         Set listOfIntervals[41] = Set first interval after 25% number of values are reached
36
37
        Set listOfIntervals[42] = Set first interval after 50% number of values are reached
38
        Set listOfIntervals[43] = Set first interval after 75% number of values are reached
39
       else if (interval number = = All) then
40
         Set listOfIntervals[5] = All values are used
41
42
43
     return listOfIntervals
44
     end for
45
```

Figure 5: Algorithm Automatic Setup of the parameters and intervals for the categorization

the 20% of elements), 4 intervals are set (line 17), and each interval will contain the 25% of the values (lines 35-38). If there is a value that contains the 80% of the elements, only 1 interval is set (line 19), which will be the value that has more elements (line 29). If 1 or 4 intervals are not set, the k-means algorithm is used to decide whether using 2 or 3 intervals (line 22). The criterion for the decision is the total minimum distance between the elements and their centroid (line 23). If 2 intervals are set, the limit between intervals will be set to the higher value of the first cluster (line 31), and if 3 intervals are set, limits between intervals will be set to the higher values of the first and second clusters (lines 33-34).

4.4. CATEGORIZATION RESULTS

As a result of the previous steps, the different typologies are presented giving statistical information for each of them such as: intervals of each parameter, total number of building, total number of buildings included and representativeness of the category in the whole set of buildings. Different colours are used to visually identify different typologies. Also a 3D visualization of the city

model is used in order to identify which buildings belong to which typology and their geographical distribution.

If the categorization does not provide the user with satisfactory enough results, it is possible for the user to come back to the previous stage and modify the selected criteria.

4.5. BUILDING INFORMATION

This process is manually done in the categorization tool by selecting the sample building in the city viewer of the visualization panel. The software tool provides support in the identification, first providing the information of the selected building (e.g. address); second by the colour of the typology; and third by including the selected building with the statistical information of the typology when the sample building is selected.

5. IMPLEMENTATION: THE CASE STUDY OF SANTIAGO DE COMPOSTELA

The historic district of Santiago de Compostela has been selected as the case study for the implementation of the described building stock categorization process. Santiago de Compostela is located in the north-west of Spain with around 100.000 inhabitants, whose historic centre was declared World Heritage Site by UNESCO. The area selected as case study is the area that was histori-

cally inside the walls of the city. 805 buildings have been considered within this area (699 residential and 106 non-residential).

3D city model described in Section 2.2 has been used for the implementation of the described process. The model has been completed with the information of the key parameters for the categorization identified in Section 3.2. For the completion of the information, both the Spanish Cadastre and the Masterplan of the historic city of Santiago have been used. According to the masterplan of the city, 5 different levels of protection are defined for the buildings in the historic city: Monumental buildings of outstanding value (Level 1), buildings of singular features and major value (Level 2), buildings with special features regarding architecture and environment (Level 3), Interesting buildings in the urban context (Level 4) and not listed (Level 0). None of the residential buildings located in the historic district is catalogued as Level 1.

The categorization process has been performed in the categorization tool with the following configuration:

Year of construction: 2 intervals. 1298 – 1800 and 1800
 –1986

| IDENTIFICATION OF BUILDINGS TYPOLOGIES | | | | | SELECT REPRESENTATIVE TYPOLOGIES | | | |
|--|----------------------|-------------------------|------------------------|-----------------|----------------------------------|--------------------|---|--------------------------------------|
| Use | Number of Facades | Year of Construction | Level of Protection | Typlogy Code | Number of Buildings | Total Buildings | Historical and Residential Buildings | Selected Tipologie (3% Threshold) |
| Residential | 1 | <1800 | 4 | 1111 | 9 | 1,12% | 1,29% | |
| | | | 3 | 1112 | 2 | 0,25% | 0,29% | |
| | | | 2 | 1113 | 0 | 0,00% | 0,00% | |
| | | | 0 | 1114 | 1 | 0,12% | 0,14% | |
| | | >=1800 | 4 | 1121 | 73 | 9,07% | 12,07% | 5 |
| | | | 3 | 1122 | 7 | 0,87% | 1,00% | |
| | | | 2 | 1123 | 1 | 0,12% | 0,14% | |
| | | | 0 | 1124 | 7 | 0,87% | 1,00% | |
| | 2 | <1800 | 4 | 1211 | 22 | 2,73% | 3,64% | 6 |
| | | | 3 | 1212 | 30 | 3,73% | 4,96% | 2 |
| | | | 2 | 1213 | 1 | 0,12% | 0,14% | |
| | | | 0 | 1214 | 0 | 0,00% | 0,00% | |
| | | >=1800 | 4 | 1221 | 264 | 32,80% | 43,64% | 7 |
| | | | 3 | 1222 | 85 | 10,56% | 14,05% | 3 |
| | | | 2 | 1223 | 3 | 0,37% | 0,43% | |
| | | | 0 | 1224 | 38 | 4,72% | 6,28% | 1 |
| | 3 | <1800 | 4 | 1311 | 8 | 0,99% | 1,14% | |
| | | | 3 | 1312 | 10 | 1,24% | 1,43% | |
| | | | 2 | 1313 | 10 | 1,24% | 1,43% | |
| | | | 0 | 1314 | 1 | 0,12% | 0,14% | |
| | | >=1800 | 4 | 1321 | 60 | 7,45% | 9,92% | 8 |
| | | | 3 | 1232 | 33 | 4,10% | 5,45% | 4 |
| | | | 2 | 1233 | 4 | 0,50% | 0,57% | |
| | | | 0 | 1234 | 10 | 1,24% | 1,43% | |
| | 4 | <1800 | 4 | 1411 | 0 | 0,00% | 0,00% | |
| | | | 3 | 1412 | 3 5 | 0,37% | 0,43% | |
| | | | 0 | 1413 | - | 0,62% | 0,72% | |
| | | >=1800 | 4 | 1414 | 7 | 0,00% | 0,00% | |
| | | | 3 | 1421 1422 | 3 | 0,87% 0,37% | 1,00% 0,43% | |
| | | | 2 | 1422 | 2 | 0,37% | 0,43% | |
| | | | 0 | 1423 | 0 | 0,23% | 0,00% | |
| Non residential | | | | 2424 | 106 | 13,17% | | |
| | | | | | | | 15,16% Selected Tipologies | 8 |
| | | | | | | | Selected Tipologies Selected Buildings | 605 |
| | | | | | | | Representativity | 86,55% |

Table 1: Generation of typologies

- Main use: 1 interval. Only residential →The automatic categorization set up process has set to All values; however, for the case study purposes only residential buildings are considered.
- Number of facades. All values.
- Protection degree. All values.

The selected configuration for the case study provides the results presented in Table 1. 32 typologies have been generated for residential buildings in the historic district. Only 8 of them are over the selected threshold (3%) of representativeness. Within the selected typologies a total of 605 buildings have been considered, representing the 86,55% of the total number of residential buildings allocated in the historic district of Santiago de Compostela.

Considering only residential buildings, the majority of them have 2 façades, due to the urban configuration of the historic districts. Also it can be noticed that most of the buildings in the historic district has been reconstructed after 1800. The level of protection of the residential buildings is low (3 or 4). The most common typology is represented by residential buildings of 2 façades, reconstructed after 1800 and with a protection level of 4.

Figure 6 shows the geographical distribution of the typologies within a 3D visualization in the categorization tool. Note that only residential buildings were of interest for the categorization and the 3D city model includes also the non-residential ones. The whole set of resulting typologies, including their representative parameters and statistical information, as they are shown in the categorization tool, are also presented.

The selection of sample buildings is based on the availability of detailed data of buildings in the masterplan of the city of

Santiago. Detailed information has been completed for each sample buildings using the web forms provided by the categorization tool. The detailed information about each sample building is shared with other application developed in the EFFESUS project in order to select the hest retrofitting strategies for each typology.

6. DISCUSSION

The four modelling strategies identified in section 2.1 can be useful depending on the circumstances, the objectives and the data availability. All of them try to represent a simplified version of a com-

plex reality based on correlations and estimations from a small number of basic building data. The accuracy of the results obtained will be directly proportional to the number and reliability of the data used. In the case of representative buildings, although the method generates uncertainty in the extrapolation to the entire district of the results for the selected buildings, accurate results can be obtained from a very basic dataset. Furthermore, in a decision making process regarding historic districts, the constraints regarding heritage significance are mandatory. The identification of representative buildings allows taking into account the historical value of buildings during the modelling process. The modelling technique based on sample buildings, versus the use of archetypes, offers the advantage of being based on real data of the buildings. In addition, when a virtual city model is used, it can be completed when more information becomes available, thus achieving greater accuracy in the obtained results.

7. CONCLUSIONS AND FUTURE WORK

The main objective of the work presented in this article is the development of a software tool for building stock categorization of historic districts for energy retrofitting, based on data available or requiring little field work and that provides high representativeness results with a reduced number of typologies

The buildings stock categorization is based on a modelling strategy based on sample buildings. The described method for building stock categorization proposes a reduced an easily acquired set of parameters (use, number of facades, year of construction and level of protection) that gives optimal balance between the number of typologies and the represented percentage



Figure 6: Geographical distribution of typologies and categorization results

of the building stock (in the case of Santiago de Compostela 8 typologies represent the 86,55% of the residential buildings).

The method described is feasible and affordable for the implementation in a web-based tool, which partially automatizes the process and provides the user an easy and intuitive way for the identification of building typologies. The categorization tool is based on a 3D city model which provides some of the required geometric data (i.e. volume, number of façades) while other required data can be obtained from public sources and can be added to the model almost automatically. The k-means clustering algorithm has been adapted in order to optimize the identification of categorization criteria (intervals and limits) which provides good results in the implemented scenario.

The result of modelling based on sample buildings facilitates the estimation of the energy performance of a district with a good balance between the amount of information needed and the accuracy of the results obtained. The modelling process based on sample buildings and an urban 3D model is a good starting point for initial references and estimates that can be incrementally completed for more accurate results in the future.

The described categorization method is applicable to other application domains (e.g. vulnerability to climate change) requiring minor adaptations, especially with regard to criteria for the selection of categorization parameters. A potential future work area is also the improvement of the method and automation of sample buildings selection.

BIBLIOGRAPHY

- [1] A. R. Amorrortu, E. B. Sánchez, and V. S. Zabala, "Edificios de balance energético positivo," Dyna, vol. 85, no. 9, pp. 746–759, 2010.
- [2] M. Batty, "Urban modeling," in International Encyclopedia of Human Geography, R. KitchinNigel, Ed. Oxford: Elsevier, 2009, pp. 51–58.
- [3] M. Kavgic, A. Mavrogianni, D. Mumovic, A. Summerfield, Z. Stevanovic, and M. Djurovic-Petrovic, "A review of bottom-up building stock models for energy consumption in the residential sector," Build. Environ., vol. 45, no. 7, pp. 1683–1697, 2010.
- [4] R. Nouvel, M. Zirak, H. Dastageeri, V. Coors, and U. Eicker, "Urban Energy Analysis Based on 3D City Model for National Scale Applications," in Presented at the IBPSA Germany Conference, 2014, vol. 8.

- [5] T. Krüger and T. Kolbe, "Building analysis for urban energy planning using key indicators on virtual 3D city models—the energy atlas of Berlin," Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. Vol. XXXIX-B2, 2012. XXII ISPRS Congr. 25 August – 01 Sept. 2012, Melbourne, Aust., vol. XXXIX, pp. 145–150, 2012.
- [6] A. Egusquiza, A. Gandini, J. L. Izkara, and I. Prieto, "Management and decision-making tools for the sustainable refurbishment of historic cities," V Congr. Latinoam. REHABEND 2014 sobre Patol. la Construcción, Tecnol. la Rehabil. y Gestión del Patrim., 2014.
- [7] A. Parekh and P. Eng, "Development of archetypes of building characteristics libraries for simplified energy use evaluation of houses," in In Proceedings of the Ninth International Building Performance Simulation Association Conference (IBPSA, 2005).
- [8] Mata, É.; Sasic Kalagasidis, A.; Johnsson, F. "A modelling strategy for energy, carbon, and cost assessments of building stocks.". Energy and Buildings, vol. 56 pp. 100- 108, 2013
- [9] L. Ross, J. Bolling, J. Döllner, and B. Kleinschmit, "Enhancing 3d city models with heterogeneous spatial information: Towards 3d land information systems," Adv. GlScience, pp. 113–133, 2009.
- [10] F. Biljecki, J. Stoter, H. Ledoux, S. Zlatanova, and A. Çöltekin, "Applications of 3D City Models: State of the Art Review," ISPRS Int. J. Geo-Information, vol. 4, no. 4, pp. 2842–2889, 2015.
- [11] A. A. Famuyibo, A. Duffy, and P. Strachan, "Developing archetypes for domestic dwellings - An Irish case study," Energy Build., vol. 50, pp. 150– 157, 2012.
- [12] O. Barnett and A. Cohen, "The Histogram and Boxplot for the Display of Lifetime Data," J. Comput. Graph. Stat., vol. 9, no. 4, pp. 759–778, 2000.
- [13] M. Carpenter, T. Crosbie, M. Crilly, N. Dawood, and Mhalas Amit, "Implications of open access data for low cost KPIs measuring energy efficiency," ICT Sustain. places Int. Conf., 2014.
- [14] J. A. Hartigan and M. A. Wong, "Algorithm AS 136: A k-means clustering algorithm," J. R. Stat. Soc. Ser. C (Applied Stat., vol. 28, no. 1, pp. 100–108, 1979.

ACKNOWLEDGMENT

The work described in this article is partially funded by the "Energy Efficiency for EU Historic Districts' Sustainability" (EFFESUS) project, Grant Agreement Number 314678, 2012–2016, as part of the EU's FP7.