



**HAL**  
open science

# Spatial Evolution of Building Types: Methodological proposals

Joan Perez, Giovanni Fusco, Sadahiro Yukio

► **To cite this version:**

Joan Perez, Giovanni Fusco, Sadahiro Yukio. Spatial Evolution of Building Types: Methodological proposals. UMR 7300 ESPACE. SAGEO 2021 - 16th Spatial Analysis and Geomatics Conference. Actes de la conférence., 2021, 978-2-910545-12-1. hal-03477740

**HAL Id: hal-03477740**

**<https://hal.science/hal-03477740>**

Submitted on 13 Dec 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

---

## Spatial Evolution of Building Types: Methodological proposals

Perez Joan<sup>1</sup>, Fusco Giovanni<sup>1</sup>, Sadahiro Yukio<sup>2</sup>

1. Université Côte d'Azur, CNRS, ESPACE, Nice, France.

2. Department of Urban Engineering, University of Tokyo, Tokyo, Japan.

---

*ABSTRACT.* This paper presents different methods implemented with the aim of studying urban perforation and regeneration dynamics at the building level. The minimum requirement to perform these methods is a building footprint available for two different periods (time depth of at least 10 years) and heights (or floors) as attribute data. First, basic morphometric indicators are computed for each building: area, floor-area, number of neighbors, elongation, and convexity. Based on the availability of expert knowledge, different types of classifications and clustering are performed to obtain building types. A grid is superimposed on the test region of Osaka (Japan) and the number of building types per cell and for each period is computed, as well as the differences between each period. Mappings are then performed, showing that building types have specific locations and evolution trajectories. In some extreme cases, a specific building type can even gradually replace a type on a declining dynamic.

*KEYWORDS:* Spatial analysis, urban perforation, urban regeneration, building type, classification, clustering.

---

## 1. Introduction

Cities are in a perpetual flow of transformations with urban areas that constantly make, unmake and remake themselves. Different models are discussed in the literature, mostly focusing on the benefits and/or drawbacks of the increase (urban sprawl, compact cities, regeneration, etc.) or decrease (shrinking cities, perforation, etc.) of building densities. A factor that is often overlooked regarding the evolution of densities is the role of inherited urban forms. Occurrence of areas having difficulties in redefining themselves are often closely related in term of urban form. Of course, even if the form factor is often overlooked, this issue is nothing new. The numerous criticisms of modernist architecture (Jacobs 1961, Salingaros 2005), which brought large-sized and specialized buildings, unable to evolve with the needs of the cities, are a good illustration of the link between urban life and urban form. What has changed more recently however, are the increasing possibilities to explore this link through: (1) the existence of large-sized datasets related to urban form and available at different dates, such as the French BD TOPO ® or the Japanese Zenrin Maps ® and (2) an increase of computing power that allows running complex algorithms on the aforementioned datasets. Examples of such work in the literature include the classification of building footprints by a Random Forest Classifier (Hecht *et al.*, 2015), unsupervised Bayesian clustering for building types identification (Perez *et al.*, 2019a), etc. This paper seeks to emphasize on different needs and precautions that shall be taken while working on the link between urban form and space-time evolutions of urban areas. Suggestions and recommendations while using building footprints as the primary source of data are also provided throughout the paper.

Section 2 presents the data requirement, the test region and the computation of basic indicators of built-up form. Section 3 performs three different classifications, through a decision tree if reliable expert knowledge is available, a double k-means if only limited expert knowledge is available and finally, an unsupervised Self-Organizing Map if no expert knowledge is available. Section 4 creates a grid of a predetermined size over the test region in order to count the numerosity of each building type per cell and their evolution. Thematic maps are also displayed in this section. Section 5 concludes the paper.

## 2. Data Preparation

### 2.1. Minimum Data Requirement

The bare minimum dataset required to implement the different methods discussed in this paper is a GIS layer of building footprints with multiple years of consistency (at least two different periods). The gap in years between the layers is dependent upon the scope of the analysis, but also on specificities of the location under study. A gap of 10 years is enough in countries following a deconstruction/reconstruction model of buildings (some countries in Asia, Africa, etc.) while in countries where buildings are usually renovated or rehabilitated, such as in Europe, a minimum time depth of 20 to 30 years is needed to observe structural

## Regeneration and Perforation of Building Types

changes in the distribution of building types. Each building shall be digitized as a single unit. Corrections should thus be made to GIS layers which aggregate contiguous buildings (like row houses) into single units. In addition, one attribute is required: building height, and another one is valuable, but not mandatory: building specialization. Specialization attribute allows filtering non-residential buildings to solely focus on residential patterns and dynamics of regeneration and perforation. Yet, due to mixed-use buildings, the distinction between residential and non-residential buildings is hardly ever straightforward. Thus, according to the source of data, attention should be paid to how specialization is encoded.

## 2.2. Test Region and Data Presentation

The test region is an area of 15 by 18 km in Japan containing central Osaka and its surroundings (**Fig.1 a.**). The peculiarities of Japan are that houses and small collective residential complexes are easily demolished to be reconstructed or to make way to new urban projects (Shelton, 2012) and that urban areas are experiencing urban shrinkage phenomena (Buhnik, 2010). For example, single-family houses have a lifespan of only 30 years (MLIT, 2007), which is in sharp contrast with European countries where houses are renovated and passed down from generation to generation. A short lifespan for buildings is an undeniable advantage for the study of the evolution of building types. Indeed, even now, high-quality datasets of building footprints are uneasy to access and, if they are, historic data of equivalent quality are usually not. Regarding urban shrinkage, demolitions without reconstructions is another interesting phenomenon to monitor over time.

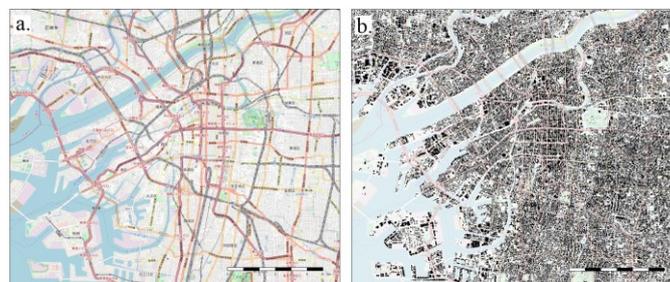


FIGURE 1. a. Osaka, OpenStreetMap. b. Buildings in Osaka in 2013-14 (Zmap TOWN II).

ZENRIN Residential Maps (Zmap TOWN II<sup>1</sup>), which are digital maps focusing on the building footprints throughout Japan, are extracted and compiled into a GeoPackage file for the extent of the test region in 2003/04 and in 2013/14 (**Fig 1.**

<sup>1</sup> ZENRIN is a private map information company that holds the top share in the Japanese market for local residential and car navigation maps.  
<https://www.zenrin.co.jp/product/category/gis/basemap/zmaptown/index.html>

b.). The original GeoPackage file only contains the two aforementioned building layers, respectively named BU0304 and BU1314. BU0304 is made of 760.067 inputs and BU1314 of 739.536, thus showing a decrease of 2.7% of the raw number of buildings in 10 years. ZENRIN Residential Maps possess several attribute data from which two are of interest in this research: the height of the buildings, expressed in number of floors, and building specialization. Building specialization is an encoded attribute, with codes as follows: 1200 for official and religious buildings (schools, administrative buildings, temple, etc.), 1363 for collective housing, 1364 for single-family homes, 1365 for private offices and mixed-use buildings, etc. In this paper, we focus on residential buildings, both collective and individual housings (1363 and 1364), which concern 469.519 buildings in 2003/04 and 428.875 in 2013/14. Each layer within the GeoPackage file originally contains 3 attribute data: an identifier, the number of floors and specialization.

### 2.3. Indicator Computation

The first step consists in computing a basic series of morphometric indicators for each building made of: the building footprint surface (area), the total amount of usable floor area, elongation, convexity and the number of adjoining neighbors. Elongation, convexity and the number of adjoining neighbors are detailed in Perez *et al.*, (2019a). They respectively provide a measure of how buildings are elongated compared to the most compact equivalent shape (a circle), how buildings have intricated or squared shapes, and finally, if buildings are free-standing detached structures or possess adjoining neighbors. The final set of variables is made of the newly calculated morphometric indicators plus number of floors and specialization. **Supplementary materials algorithm 1** shows how indicators are easily computed within R using the “sf” and “lwgeom” packages. This algorithm is applied to both BU0304 and BU1314. For the number of adjoining neighbors, we operate a small buffer (**algorithm 1**, # neighbors 1/2) with the aim of correcting buildings that shall be considered as adjoining, but are not due to low geolocation accuracy. Even if such occurrences are negligible within the ZENRIN Residential Maps, this precautionary step ought to be taken before computing the number of adjoining neighbors. Lastly, *range01* is a user defined function that allows normalizing elongation and convexity.

### 3. Residential Buildings: Classification of Types

Once a basic set of indicators has been calculated for each input layer, the next step is to perform a classification of building types that can be applied to different time points. Three options arise. First, reliable expert knowledge is available, and as such, the number and the characteristics of the relevant building types are known in advance. Second, limited expert knowledge is available, which leads to partial information known in advance, such as the number of types sought, but not their characteristics. Lastly, no expert knowledge is available, thus leading to a situation in which both the number and the content of the building types are unknown.

## Regeneration and Perforation of Building Types

## 3.1. Classification with Expert Knowledge

When expert knowledge is available, a simple and efficient method of classification is found in a manually built decision tree. Successive tests on attributes split the original dataset into different class labels. When a solution is found and deemed robust by the expert, the conditional control statements can be used as benchmarks for other datasets possessing similar attributes. **Supplementary materials algorithm 2** is a user defined function named *DTs* which allows classifying each row of a given dataset using the aforementioned morphometric indicators. Each test follows a *if condition then outcome* structure. Several conditions per test are possible, and the outcome is always a label attribution. Using the *apply* function, *DTs* can be used on any dataset possessing the relevant indicators, in our case BU0304 and BU1314.

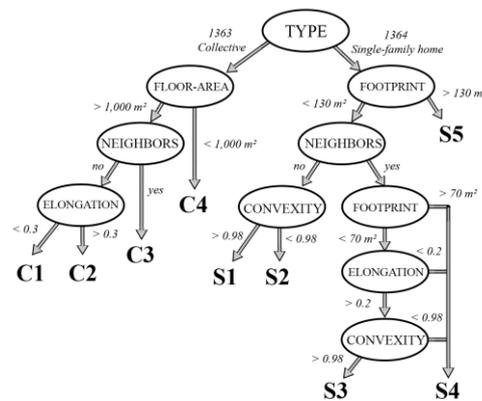


FIGURE 2. Visualization of the tests implemented in the *DTs* function

**Fig.2** is a flowchart visualization of the successive tests implemented in *DTs*. They are 9 different outcomes, in which 4 are for collective housings (C1 to C4) and 5 are for single-family homes (S1 to S5). Each label attribution is based on expert knowledge. For example, C1 are high rise residential buildings following a “tower in the park” model, C2 are elongated residential complexes representative, amongst other structures, of Danchi housing (団地) which are complexes of apartment buildings built following western models after the Second World War, S3 are small and elongated row houses, typical of traditional wooden Nagaya (長屋), etc. The relevance of these building types in Japanese urban areas is derived from literature (e.g. Shelton, 2012; Bonnin *et al.*, 2014; Perez *et al.*, 2019b).

## 3.2. Classification with limited Expert Knowledge

If expert knowledge is limited, but still available in some extent, it is possible to make several decisions prior to a series of unsupervised cluster analyses. For

example, we could know the number of required clusters, as well as the fact that some inputs must not be grouped together. This could for instance be the case for collective and single-family homes, for specialized and residential buildings, or for building possessing large and small footprints. In such cases, standard unsupervised clustering techniques that minimize intra-cluster distances and maximize inter-cluster distances can be separately used for each subset, such as K-Nearest Neighbors, hierarchical clustering, *k*-means, classification trees, random forests, neural networks, etc. To illustrate this point, one of the simplest technique, *k*-means clustering, which aims at partitioning inputs into clusters such that the sum of squares from inputs to the assigned cluster centers is minimized, is applied as follows: **Supplementary materials algorithm 3** is a R protocol performing two different *k*-means, one for collective (attribute 1663) and one for single-family homes (attribute 1664). Two random subsets are draw and merged from BU0304 and BU1314 for collective buildings only (sample size of 20% of the inputs), and, once again, from BU0304 and BU1314 for single-family homes only (20% again). Drawing and merging random samples made of both periods altogether allows avoiding giving more weight to one period over the other. A clustering analysis is then performed for each subset (collective housing and single-family homes), with a number of clusters per model chosen *a priori*. In **algorithm 3**, we select the same number of clusters as within the DTs model (section 3.1: 4 clusters for collective housing and 5 clusters for single-family homes). Data are scaled and, in order to avoid redundancy, floor and area are removed in favor of floor-area. The two clustering analyses are thus performed on convexity, elongation, number of direct neighbors and floor-area. Once the two models are trained, for both periods, and for both collective and single-family home inputs, the whole datasets are mapped to the nearest clusters. This mapping is made using the *k*-nearest neighbor searching algorithm of the “FNN” package.

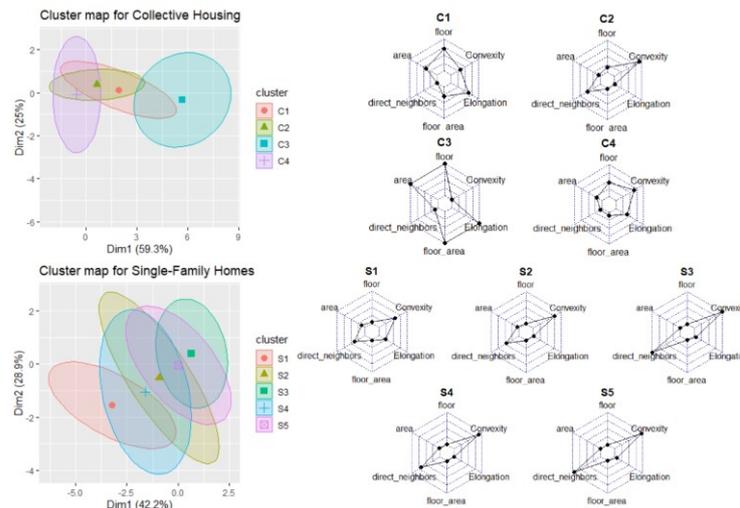


FIGURE 3. Cluster Map and Profiles for individual Clusters for Collective Housing (top) and Single-Family Homes (bottom)

## Regeneration and Perforation of Building Types

The left side of **Fig.3** is a possible representation of the two clustering within two-dimensional spaces made by extracting the first two principal components (package “factoextra”) along which the variation in the data is maximal. Quality metrics in these clustering analyses are the share of the total variance in the data set that is explained by the clustering, given by dividing the between sum of square (BSS) by the total sum of square (TSS). These values are 94.13% for 4-class collective housing clustering and 93.82% for the 5-class, respectively. The right side of **Fig.3** shows radar charts for each variable. It is possible to add variables in the radar charts that have not been used during the clustering processes, such as area and floor in **Fig.3**. These plots are made with the package “fmsb”. The charts point at highly differentiated profiles, especially for collective housing, with profiles that are partially matching the DTs model. C3 are for example regrouping both high rise residential buildings and Danchi housings (団地), C4 appears to be the equivalent of DTs C4: small-sized complexes, etc. Cluster concordances with the DTs model are less clear for single-family houses, except for S2: articulated detached houses (equivalent to DTs S2) or S1: large villas (DTs S5). S3 and S5 are row houses with close characteristics, also close in the cluster map. In fact, even if literature or partial expert knowledge say that we shall seek for 5 different profiles of single-family homes, it appears that 5 is not a robust number of clusters for the k-means algorithm. In this respect it is possible to either (1) add/remove variables, such as using floor and area, instead of floor-area, or add newly computed morphometric indicators, (2) re-run the procedure (k-means is sensitive to the location of starting centroids) or try another algorithm, such as one of the algorithm discussed at the beginning of this section (3) re-run the algorithm with a different number of sought clusters (4) manually merge the clusters that have close characteristics.

### 3.3. Classification with no Expert Knowledge

Finally, in the case of a total lack of expert knowledge, any widely known clustering algorithms can be used, such as k-means, neural networks, or Bayesian networks. The main issues with unsupervised clustering are the assessment of the quality of the partitions, as well as the number of “interesting” partitions (Haldiki *et al.*, 2001). Since there is no way to *a priori* determine the most suitable number of clusters, researchers tend to use heuristics, such as looking at a cutoff point regarding the improvement of the explained variance (or log-likelihood in the case of Bayesian clustering) for each additional partition (elbow method). An unsupervised Bayesian clustering on BU1314, based on the indicators detailed in section 2.3 (plus specialization) has been performed in Perez *et al.*, (2019a). This clustering used an expectation-maximization algorithm (Dempster *et al.*, 1977) to perform one thousand clustering analyses and a MDL score (combining log-likelihood and a penalization function for the growing number of clusters) to define the optimal number of partitions. This application yields seven interesting clusters, from detached small houses to isolated high-rise buildings. In what follows, we will propose an alternative approach to unsupervised clustering based on Self-Organizing

Maps (SOM) (Kohonen, 1982). Bayesian and SOM clustering have already been compared in unsupervised approaches (Fusco and Perez, 2019).

**Supplementary materials algorithm 4** is a R protocol performing SOM clustering on a sample made of BU0304 and BU1314 (sample size of 20% of the inputs). This time, we do not distinguish collective from single-family homes buildings. Once again, the sample is made with inputs draw from both periods altogether. The same indicators than in section 3.2 are used. SOM use a neighborhood function for each input to find the Best Matching unit (competitive training). Units, also called nodes, are distributed on a two-dimensional space (map). Once the best matching unit is found, a radius parameter allows updating the neighboring nodes, thus giving topological properties to the map that can be investigated through their proximity. There are two ways to identify clusters within a SOM. First, the number of nodes within the two-dimensional space can be set to a small value. Each node is then considered as a cluster. The second approach, which is the one used in **algorithm 4**, is to parameterize many nodes for the map (225, i.e. 15 by 15 in our application). The distances between couple of nodes form a distance matrix (U-Matrix). The advantage of this approach is that it allows visualizing each variable distribution across the map. Depicted by colors, these graphical representations are called heatmaps. Since similar values are aggregated in the same areas, heatmaps provide relevant information that can be used prior to the segmentation of the U-Matrix. The segmentation can then be performed using a simple hierarchical clustering, as in **algorithm 4**, in which the number of divisions is set manually.

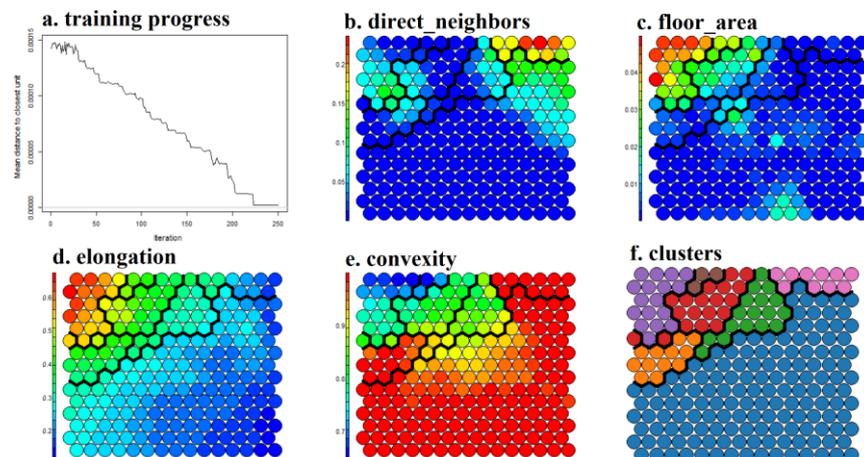


FIGURE 4. Self-Organizing Map outputs for a sample of BU0304 and BU1314 a. Training Progress b. Heatmap for direct neighbors c. Heatmap for floor-area d. Heatmap for Elongation e. Heatmap for Convexity f. Clustering results.

**Fig.4** displays the outputs of a SOM clustering performed on a sample made of BU0304 and BU1314 using the “kohonen” package. The first thing to control is that the training curve flattens, as in **Fig.4 a**. The training is not over if it doesn't, the

## Regeneration and Perforation of Building Types

*r*len parameter in **algorithm 4** must thus be set to a bigger value. **Fig.4 b. to e.** are the different heatmaps, showing that similar values are aggregating in the same areas. **f.** is the cluster map, where each color is a different building type. In this example, we set the number of clusters to seven. The black lines are the cluster boundaries, which have been retroactively mapped to the heatmaps. As in k-means, the ratio between BSS and TSS can be calculated for different numbers of clusters. For the 7-class solution, it reaches 84.01%. The “blue” cluster is made of different kinds of small sized single-family houses, the “pink” is for elongated row houses such as wooden Nagaya (長屋), the “purple” is for Danchi housing (団地), the “red” is representative of small collective residential complexes, etc. Once the SOM is trained with the sample, BU0304 and BU1314 data are mapped into the model using the *map.kohonen* function.

#### 4. Spatial Analysis of Perforation and Regeneration

##### 4.1. Automated Grid Creation and Evolution Count

In the previous section, we performed three different classifications that have been applied to the different periods associated with our datasets: 2003/04 and in 2013/14. To summarize, **algorithm 2, 3** and **4** are individually able to provide a new variable: a class label, that can be compared between two periods. This is the only input required in this section: a class label associated to at least two spatial datasets describing the same study area at different time points. **Algorithm 5** creates a grid with predetermined cell size, 250 meters, that are superimposed on the extent of the test region. Cells that are not intersecting any buildings are filtered out of the grid. Then, using a *tapply* structure, a recursive intersect is performed for each remaining cell to count the buildings grouped by labels. Once the *tapply* function has been ran for both datasets, a final line of code groups the results of both periods in a single dataframe. This whole process takes only a few seconds with an Intel® Core™ i7-8700 CPU @ 3.20GHz, 16gb RAM, while, on a traditional GIS software, for each and every class label, an intersect with a grid of this extent takes several minutes to run. For each cell, it is then possible to calculate the difference per building type, in order to map the temporal evolutions of each type.

##### 4.2. Mapping

To map the count and evolution of building types, we use the “tmap” package, which allows building thematic maps following a layer-based structure. **Supplementary materials algorithm 6** shows a layered structure leading to the map of the count of the number of building for the C3 profile of the DTs model (**Fig.5 a.**). Before setting the maps, we first import an OpenStreetMap background. C3 values are then discretized into several categories (natural breaks), sequentially added as new layers to the map. **Algorithm 6** is interesting on both counts. First, for each category, cells are dissolved into a single multipart polygon before being

plotted. It allows plotting only the geometries thus substantially reducing computation time. Second, for each category, polygons are smoothed in order to make localized aggregates stand out.

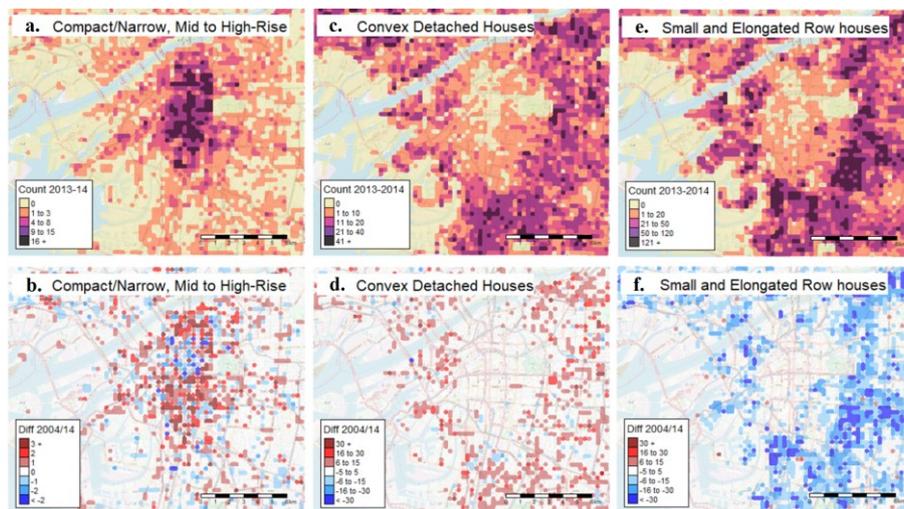


FIGURE 5. 2013-14 Count and 2003-4/2013-14 Evolution of building types for selected Profiles of the DTs model a. Count for C3 b. Evolution for C3 c. Count for S1 d. Evolution for S1 e. Count for S3 f. Evolution for S3

**Fig.5** shows the count and evolution of three building types from the DTs model. Mappings within the above figure can be realized for any building type and for any of the three models discussed in section 3 using the structure of **algorithm 6**. C3 are mid to high rise compact/narrow buildings, mostly located in the central part of Osaka. The evolution map shows different hot and cold spots within this central area, thus pointing at a building type which is self-regenerating. S1 are convex detached houses mostly located outside of the center of Osaka. The evolution map shows that this kind of single-family homes is increasing, which is consistent with the literature pointing at a spread of small-sized prefabricated homes in Japan (Buntrock, 2017). S3 are small and elongated row houses, typical of traditional wooden Nagaya (長屋). This type is also located outside central Osaka, but also highly concentrated in specific neighborhoods. The evolution map points at a gradual disappearance of this traditional building type, even perhaps at a replacement if we link this disappearance with the increase of modern prefabricated homes.

## 5. Conclusion and Discussion

This paper put forward several recommendations and suggestions with the aim of studying the spatial location and temporal evolution of building types in large metropolitan areas. The first and most important condition is that high quality GIS

## Regeneration and Perforation of Building Types

layers of building footprints with heights (or floors) as attribute data are available for at least two different time points. Another attribute is valuable, but not mandatory: building specialization. Attention should be paid to the time depth between the layers, as it must be sufficient to analyze structural changes in the spatial organization of building types. The temporality behind these structural changes is also dependent on cultural specificities regarding planning, building norms and urban development traditions (deconstruction, dismantlement, renovation, etc.). If the prerequisites are fulfilled, then it is possible to compute a basic set of indicators for each building, such as the one presented in section 2.3. Based on the kind of available expert knowledge, different types of classifications have been performed, from manually built decision trees to fully unsupervised clustering. Regardless of the method, quality metrics show that significantly different clusters can be found. It is possible to check the validity of each classification by drawing random samples and go on the field, or use a street-based urban imagery (such as Google Street View), then calculate the usual sensitivity and specificity metrics. The automated superimposition of a grid then allows counting occurrences of each type, as well as quantify the difference per cell between the different time points. For each category, mappings are performed, with cells that are discretized in the same category dissolved into a single multipart polygon, thus highlighting localized aggregates and, in the case of temporal evolution, hot and cold spots of perforation and regeneration of building types.

This paper focused on methodological suggestions, proposals and recommendations for studying building evolutions. Regarding the clustering applications, the methodological propositions could be improved, especially by focusing on maintaining output coherence over different settings or initializations. A matrix of similarity indexes can for example be calculated among the clustering results to evaluate the robustness of the protocols to pseudorandom number generation (seed), and cross-validation methods, such a k-fold or Jackknife resampling, can be used to evaluate the robustness of the outputs to different random samples. Yet, interesting preliminary results already stand out regarding specific locations and evolution trajectories of certain building types in Osaka, Japan. It shows that the perpetual mutation that characterize urban areas is linked, at least to some extent and to different degrees, to urban form. This research opens interesting perspectives, such as the study of the disappearance of some building types, or the gradual replacement of some types by other. Several methods, from standard correlation matrices to geographically weighted regressions, could allow exploring these dynamics in conjunction with other factors, such as population evolution, socio-demographic characteristics, planning policies, etc. These results must be analyzed together with literature and fieldwork related to the metropolitan area of Osaka.

## 6. References

Bonnin, P., Nishida, M. and Inaga, S. (2014) *Vocabulaire de la spatialité japonaise*. CNRS Editions, 560 p.

Buhnik, S. (2010) From Shrinking Cities to *Toshi no Shukushō*: Identifying Patterns of Urban Shrinkage in the Osaka Metropolitan Area. *Berkeley Planning Journal*, 23(1), pp. 132-155.

Buntrock, D. (2017) “Prefabricated housing in Japan”, in Smith, E. and Quale, J. (Eds.), *Offsite Architecture Constructing the Future*, Routledge: New York, NY, USA, pp. 190–213.

Dempster, A., Laird, N. and Rubin, D (1977) "Maximum likelihood from incomplete data via the EM algorithm", *Journal of the Royal Statistical Society. Series B*, 39(1), pp. 1-38.

Fusco G. and Perez J. (2019) Bayesian Network Clustering and Self-Organizing Maps under the Test of Indian Districts. A comparison, *Cybergeo: European Journal of Geography*, paper 887.

Hecht, R., Meinel, G. and Buchroithner, M. (2015) Automatic identification of building types based on topographic databases – a comparison of different data sources. *International Journal of Cartography*, 1(1), pp. 18-31.

Haldiki, M., Batistakis, Y. and Vazirgiannis, M. (2001) "On Clustering Validation Techniques", *Journal of Intelligent Information Systems*, Vol.17, pp. 107-145.

Jacobs J. (1961) *The Death and Life of Great American Cities*, Random House, New York, 458 p.

Kohonen, T. (1982) “Self-organized formation of topologically correct feature maps”, *Biological Cybernetics*, volume 43, pp. 59-69.

MLIT. (2007) “White Paper on Land, Infrastructure, Transport and Tourism in Japan”, *Ministry of Land, Infrastructure, Transport and Tourism Publication*, Ministry of Land: Tokyo, Japan, 68 p.

Perez, J., Fusco, G., Araldi, A. and Fuse, T. (2019a) “Identifying building typologies and their spatial patterns in the metropolitan areas of Marseille and Osaka”, *Asia-Pacific Journal of Regional Science*, 4, pp. 193-217.

Perez, J., Araldi, A., Fusco, G. and Fuse, T. (2019b) “The Character of Urban Japan: Overview of Osaka-Kobe’s Cityscapes”, *Urban Science*, 3(4), 105.

Salingeros, N. (2005) *Principles of Urban Structure*, Techne Press, Delft, 254 p.

Shelton, B. (2012) *Learning from the Japanese City: West Meets East in Urban Design*. Taylor & Francis, 224 p.

## 7. Supplementary Materials

<https://zenodo.org/record/4551462>

*Acknowledgment: This study is supported by Joint Research Program no. 774 at CSIS, UTokyo (Zmap TOWN II 2013/14 Shapefile Osaka prefecture).*