Developing Zonal House Vacancy Estimation Models with a Resolution of 250 m: A Case Study of Residential Zones in Kyoto City

Ziad Abdelhalim*, Ryoji Matsunaka**, Tetsuharu Oba**

Abstract
As Japan is expected to enter a long period of population decline, the house vacancy ratio in Japanese cities is expected to increase. Therefore, efficient house vacancy surveying methods are required to accelerate research on housing vacancies in Japan. In this study, a standardized vacancy surveying method is developed. This method is fast and independent of the subjectivity of the surveying personnel, allowing the surveying process to be easily streamlined. This surveying method—which provides a vacancy ratio at a resolution of 250 m—was applied to selected parts of Kyoto city. The results of the survey were analyzed to examine the zonal variation of the residential characteristics of the study areas. The insights obtained from that analysis were then used to develop dwelling units estimation models for two of the residential zoning classes in Japan. The newly developed zonal models provide an accurate estimation of the volume of dwelling units and consequently vacant dwelling units in target areas. Hence, they provide researchers and municipalities with a tool to identify areas with a high concentration of vacant dwelling units in the exclusively low-rise and mid/high-rise residential zoning classes in Japan.

Keywords:
House vacancy surveying, vacant dwelling units estimation, urban shrinkage

[*] École Polytechnique
[**] Graduate School of Engineering, Kyoto University
E-Mail: ziadabdelhalim12@gmail.com
1. Introduction

Projections of the Japanese population based on the 2015 census predict that Japan is expected to enter a long period of population decline from 127.09 million in 2015 to 99.24 million in 2053 based on medium fertility estimations (National Institute of Population and Social Security Research in Japan, 2017). This trend implies that vacant houses will emerge more frequently in Japan, suggesting that there will be an oversupply of vacant houses well over the current (2013) vacancy ratio of 13.5 % [figure 1] (Nomura Research Institute, 2016).

In light of the experience of North-American and European cities that faced urban shrinkage in the late 20th century, the negative impacts of vacant houses are well documented (Wiechmann & Pallagst, 2012). A high vacancy ratio is associated with an increased cost of social services and infrastructure maintenance; moreover, vacant houses are associated with waste dumping and illegal activities (Rybczynski & Linneman, 1999). A study on the negative externalities of vacant houses in Toshima, Tokyo found that a vacant house decreases the values of nearby rental prices by 1~2 % on average (Sadayuki, et al., 2019). Furthermore, in a survey by the Ministry of Land, Infrastructure, Transport and Tourism (MLIT) on the public awareness of familiar land issues, 41 % of the respondents were aware of vacant land and empty houses in their environments with their biggest concern being “Intrusion of suspicious persons and arson” (Ministry of Land, Infrastructure, Transport and Tourism, 2017).

In Japan, urban shrinkage is not a phenomenon expected in the future, but one already manifested since the burst of the bubble economy in the 1990s. Although the Japanese population only started declining from 2008, 26.7 % of the Japanese cities with populations over 100,000 persons experienced population loss during the period from 1990 to 2007 (Martinez-Fernandez, et al., 2016). Currently, the housing rental market vacancy ratio in Tokyo [15.7 %] is well above the optimal vacancy ratio for the rental housing market [1.96 %] (GU & Asami, 2016).

As more urban shrinkage is expected in the future, the motivation behind this study is to tackle the problems preventing fine-scale house vacancy research in Japan. Hence, the first objective of this study is to develop a standardized house vacancy surveying method—which avoids the usual problems associated with house vacancy surveying—for Japanese cities. The second objective of this study is to develop dwelling units estimation models for residential areas. The estimation models—which are developed to be employed in situations when surveying is not possible—will use insights from the zonal variations of the residential characteristics of the study pixels and therefore will be zonal models. Aside from the informative value of identifying the spatial distribution of vacant housing, the models developed are the first step in conducting fine scale economic studies to identify the dynamics of housing vacancy on the neighborhood level.

This study consists of 6 chapters including the current one. In chapter 2, a literature review of urban shrinkage in the Japanese context and existing house vacancy estimation methods will be introduced. Chapter 3 represents the data collection section of this study. The sources of the data and the data collection methodology—including a newly proposed house vacancy surveying method—employed in this study are established. The zonal variations of the study areas’ residential characteristics are analyzed in chapter 4. While in chapter 5, the insights obtained from the latter analysis are used to develop the dwelling units, estimation models. Finally, this study’s conclusion and recommendations for future research are presented in chapter 6.

2. Literature Review

The research field on urban shrinkage and its subsidiaries—including housing vacancies—is well established in Japan. Generally, the topic has been addressed through a multitude of facets. For example, the influence of socio-economic and built environment factors at the city level on housing vacancies has been evaluated by Baba & Asami (2017). While, urban shrinkage’s manifestation of land vacancies was extensively studied by Sakamoto, et al. (2017). In the latter
study, the factors that led to the persistence of land vacancies have been identified using the city of Tottori as a case study. Similarly, Abe, et al. (2014) evaluated the same topic on a much broader scale. Their study concerned the factors that lead to the transformation of underused land in 37 core cities in Japan using historical data from 1985 to 2005.

Figure 1. Past (1978-2013) and Projected (2018-2033) House Vacancy Ratio in Japan (Nomura Research Institute, 2016).

Another set of studies focused on identifying the impacts of the proliferation of housing vacancies in Japanese cities. This research sub-field uses case studies to identify the symptoms of urban shrinkage’s manifestation of housing vacancies. This had the form of identifying the impact of vacant houses on nearby rental prices in Sadayuki, et al. (2019) study on Toshima, Tokyo. Their study helped identify the factors that decrease rental prices the most in the context of housing vacancies. On the other hand, the optimal rental housing vacancy rate [1.96%] of the Tokyo metropolitan area was formulated and compared to the existing one [15.7%] as of 2015 (GU & Asami, 2016). This set the social cost of increased housing vacancies in the Tokyo rental market to be around 442.3 billion JPY, highlighting the need for policy responses.

Although the above studies have tackled diverse problems within the housing vacancy spectrum, they all shared the difficulties of procuring fine-resolution house vacancy data. Given that housing vacancy figures are only provided on the municipal level in Japan, micro-scale studies often rely on visual inspection techniques to provide data for research. Moreover, there are several surveying techniques employed to provide housing vacancy figures in Japan (e.g., The Housing and Land’s survey and the Ministry of Land, Infrastructure, Transport and Tourism’s vacancy survey). This problem was highlighted by So (2017) who compared these house vacancy surveying techniques using data of Tokyo’s 23 wards. He confirmed that vacancy data that relied on visual inspection (i.e., The Housing and Land’s survey) were likely to be exaggerated and emphasized the need for a simple and accurate method for house vacancy surveying.

Few studies have been made on house vacancy estimation. Chen, et al. (2015) developed an empirical model that uses NPP-VIIRS Nighttime Light Composite and Land Cover Data to estimate vacancies in 15 metropolitan areas in the United States with a spatial resolution of 500 m. However, more testing is required to validate the application of the model in different topographic and urban settings like those of Japan. Additionally, some studies developed models that employ
utility data to estimate housing vacancies (e.g., Kumagai, et al. (2016) and Akiyama, et al. (2018)). Nonetheless, their methods relied on slow and subjective surveying techniques for house vacancy classification which significantly decrease the replicability of these methods and confine them within the contexts of their case studies. Furthermore, these methods exclusively target detached houses and their application on apartment and condominium buildings is questionable.

The need for a simple and unbiased surveying method for the provision of fine-resolution house vacancy data is well established in the literature. Nevertheless, to the knowledge of the author, no such method has yet to be developed. In the current study, the problem of house vacancy surveying is addressed on the macro-level, contrary to existing visual inspection methods that collect micro-level vacancy data. The new inspection method will be a fast and easily replicable process and unlike existing methods will include vacancy data from apartment buildings. Additionally, an estimation model will be developed to extrapolate the volume of vacant housing at non-surveyed locations. The aforementioned surveying and estimation models will employ public and open-source data that are available on the national level in Japan, allowing the process to be easily replicable by researchers and municipalities.

3. Data Collection Methodology

3.1. House Vacancy Surveying

The house vacancy surveying method used in this study will not depend on the visual inspection of individual houses. Instead, the method introduced by Ishikawa, et al. (2016) will be employed to obtain the volume of vacant dwelling units in the targeted areas. This method provides house vacancy data of a targeted area on the macro-level through [Eq.1]. Unlike Ishikawa, et al. (2016) implementation, the volume of the dwelling units for target areas will be collected through a field survey instead of using building point data—which provides unreliable figures in some prefectures. The adoption of this method into the surveying process will be done by integrating the household data—which is provided with a resolution of 250 m—from the census. As “Households” are defined as people who share their livelihood and address (excluding dormitories) or single persons who live alone (Statistics Bureau, 2015), they are a representation of the occupied dwelling units in a given area. Consequently, collecting the number of dwelling units in the same area through a complementary field survey will lead to obtaining that area’s volume of vacant dwelling units. In this context, a vacant dwelling unit is a dwelling unit (detached housing, row housing, and apartment housing regardless of the ownership status) that was not occupied by any individual at the time of conducting the 2015 national Japanese census.

\[ \text{vacant dwelling units} = \text{dwelling units} - \text{occupied dwelling units} \quad \text{[Eq.1]} \]

Another consequence of targeting the vacancy surveying issue from the macro-level and eliminating the dependence on subjective visual inspection is the simplicity of the alternative field survey. Collecting the number of dwelling units can be done by counting the number of residential mailboxes, nameplates, or utility meters in the target area, consequently decreasing the volume of collected figures to just one.

3.2. City Selection

The 20 Japanese “Designated Cities” are the prime applicants for the findings of this study, as their large sizes make the conventionally used surveying methods inefficient to apply throughout their city areas, causing them to settle for sample surveys instead. Moreover, conducting this study on one of the “Designated Cities” would increase the applicability of the findings of this study, as these cities house various characteristics compared to the monotone
residential nature of rural cities. The candidate cities were ordered based on the proximity of their vacancy ratio to the median vacancy ratio of the designated cities \(v_{\text{MEd}}\) \[Eq.2\].

\[
\text{ordering index} = |v_{\text{ratio}} - v_{\text{MEd}}| \quad \text{[Eq.2]}
\]

Out of the candidate pool for this study [table 1], Kyoto city was chosen as the target city for the study. As of 2015, Kyoto city had a vacancy ratio of 14.0% which is in close proximity to the median of the designated cities [13.2%]. Moreover, the aging rate of Kyoto [25.8%] is similar to that of Japan [26.7%] making it a good representation of the conditions in Japan (Statistics Bureau, Ministry of Internal Affairs and Communications, 2015). As surveying all areas in Kyoto city is not possible, the decision was made to survey selected areas in the city called pixels—which are square areas whose boundaries are defined in the census and have a resolution of 250 m.

Table 1. Results of the Ordering of the Japanese Designated Cities Based on the Ordering Index.

| Order | City       | Vacancy Ratio [%](2015) | Ordering Index [%] | Order | City       | Vacancy Ratio [%](2015) | Ordering Index [%] |
|-------|------------|-------------------------|--------------------|-------|------------|-------------------------|--------------------|
| 1     | Nagoya     | 13.16                   | 0.04               | 11    | Kitakyushu | 14.33                   | 1.13               |
| 2     | Sakai      | 13.25                   | 0.05               | 12    | Niigata    | 12.01                   | 1.19               |
| 3     | Kobe       | 13.05                   | 0.15               | 13    | Chiba      | 11.52                   | 1.68               |
| 4     | Shizuoka   | 13.62                   | 0.42               | 14    | Okayama    | 15.73                   | 2.53               |
| 5     | Hamamatsu  | 13.91                   | 0.71               | 15    | Sagamihara | 10.64                   | 2.56               |
| 6     | Kyoto      | 14.03                   | 0.83               | 16    | Kawasaki   | 10.41                   | 2.79               |
| 7     | Kumamoto   | 14.07                   | 0.87               | 17    | Yokohama   | 10.09                   | 3.11               |
| 8     | Sapporo    | 14.08                   | 0.88               | 18    | Sendai     | 9.97                    | 3.23               |
| 9     | Hiroshima  | 14.11                   | 0.91               | 19    | Saitama    | 9.88                    | 3.32               |
| 10    | Fukuoka    | 12.23                   | 0.97               | 20    | Osaka      | 17.18                   | 3.98               |

3.3. Pixel Selection

To test the effect of the zonal variation of the residential characteristics of the pixels on the accuracy of the estimation models, the decision was made to exclusively include pixels that only house one residential subzone of the Japanese zoning system. Hence, a large sample size from each residential subzone was required. Subsequently, the decision was made to cluster the seven residential subzones into three residential classes and survey a quota of 10 pixels per class. Moreover, as the most recent demographic data are of the 2015 census, selected pixels need not have gone through significant changes in their built environment since 2015 to allow a seamless and accurate restoration process of the built environment in the selected pixels.

As Kyoto city houses 3,706 pixels—which are situated in residential, commercial, and industrial zones—a selection process was conducted to randomly select suitable pixels for the study. To ensure that the selected pixels fit the restrictions set in this study in terms of characteristics and data availability, the pixels went through two selection phases (the detailed flow of the selection process is as illustrated in figure 2).
Phase 1
i. Exclude all pixels that host non-residential zones and/or host more than one residential subzone (i.e., a combination of two or more residential subzones).
ii. Exclude all pixels with concealed data.

Phase 2
i. Exclude all pixels with no building footprint data.
ii. Exclude all pixels with parks and/or vacant lots that occupy more than 20 % of the pixel’s area.

As illustrated in figure 2, phase 1 was done for all pixels at once while phase 2 was done separately for each residential subzone. During phase 2, the input was a randomly selected pixel from the appropriate residential subzone. Phase 2 was repeated until all candidate pixels went through the process or when the quota of the appropriate zone was reached.

The summary of the pixel selection process is provided in table 2. The difference between the input and output of the first phase shows that pixels that exclusively house residential zones are sparse. However, that does not limit the applicability of this study as such residential pixels inherently contain large volumes of housing vacancies. Moreover, pixels from “Category II Exclusively Low-rise Zone” were not found in Kyoto, while those from “Category II Mid/high-rise Oriented Residential Zone” were excluded from the study as they did not provide the required number of pixels.

A supplementary phase was conducted to select an additional 5 pixels from each of “Category I Exclusively Low-rise Zone” and “Category I Mid/high-rise Oriented Residential Zone” to reach the 10 pixels quota for the “Exclusively Low-rise Class” and “Mid/high-rise Class”. Finally, “Residential Class” failed to reach its quota after all pixels went through the filtration process; however, the decision was made to include this class in the study for the purpose of exhaustively studying the zonal variability of the collected data. It is to be noted that the single pixel from
“Category II Residential” is in the vicinity of Kiyomizu-dera which is a lively touristic hub of Kyoto city. Therefore, the residential choices of its residents are altered due to the commercial appeal of touristic activity as well as the negative externalities associated with a significantly large touristic activity. Since these conditions represent a significant deviation from the average pixel in Kyoto city, this pixel was excluded from this study.

Table 2. Summary of the Pixel Selection Process.

| Class                                | Residential Sub-zone                | Phase 1 | Phase 2 | Supplementary Phase | Total |
|--------------------------------------|------------------------------------|---------|---------|---------------------|-------|
|                                      |                                    | Input   | Output  | Input               | Output |
| Exclusively Low-rise Residential Class | Category I Exclusively Low-rise Residential | 78      | 78      | 5                   | 10    |
|                                      | Category II Exclusively Low-rise Residential | 0       | 0       | 0                   |       |
| Mid/high-rise Oriented Residential Class | Category I Mid/high-rise Oriented Residential | 82      | 82      | 5                   | 10    |
|                                      | Category II Mid/high-rise Oriented Residential |            | 7       | 7                   | 0     |
| Residential Class                    | Category I Residential             | 26      | 26      | 3                   | 3     |
|                                      | Category II Residential            | 8       | 8       | 1                   |       |

After the pixels were selected from Kyoto city, each one was assigned a unique ID which will be used to identify the pixels in this study. The spatial distribution of the selected pixels is illustrated in figure 3.

3.4. Data Collection

Demographic data about the study pixels was obtained from the 2015 Japanese census (National Statistics Center, 2015). In the context of this study, the building footprint data will function as the foundation of data collection. Hence, precise building footprint data that accurately represent the conditions in 2015 is required.

The building footprint data of the study pixels was obtained from OpenStreetMap (OpenStreetMap Contributors, 2019). However, a calibration process was required to ensure that the data accurately reflected the conditions in 2015. This calibration process consisted of three phases. The first phase consisted of adjusting the edges of the building footprint data, creating missing footprints, and merging buildings that were separated into several footprints. This phase was done on the OpenStreetMap platform using a combination of aerial imagery provided by Bing, and base maps from the Japan GSI Standard Map and Japan GSI KIBAN 2500. The second phase
consisted of adjusting the building footprint data to match the conditions of 2015 while the third phase consisted of recording the changes in the data compared to 2019. The second and third phases were done using historical aerial imagery from Google Earth Pro.

Figure 3. Map of Targeted Pixels Within Kyoto City.

To ensure that the pixels’ built environment in the case of this study was the same as the 2015 census. The assignment of buildings to their appropriate pixels was done using the centroids of the buildings based on the methodology employed in the census (Statistics Bureau, 2015).

Finally, the residential characteristics of study pixels were collected through an extensive field survey of the study pixels. It was during this field survey that the newly developed vacancy surveying method was applied. The buildings in the study pixels were used as the collection points during the survey. In addition to collecting the number of dwelling units from each building, the following was collected.

- Survey accessibility (1 if the building is accessible for surveying, otherwise 0).
- Residential index\(^1\) (1 if the building is residential, otherwise 0).
- Apartment index (1 if the residential building is an apartment building, otherwise 0).
- Number of dwelling units (number of residential mailboxes).

Due to the demolition and inaccessibility of some buildings, the survey coverage\(^2\) of the pixels’ was seldom 100%. Hence, a restoration process consisting of two phases was conducted to increase the accuracy of the data relative to the actual conditions present in 2015. The first restoration phase solely targeted demolished buildings. In this phase, historical Google Street View was used when possible to survey demolished buildings. The second phase targeted

---

1 Buildings that contained a mixture of residential and non-residential uses were still considered residential.

2 Survey coverage is the ratio of surveyed buildings area to total built area in a given pixel.
demolished buildings that could not be restored in the first phase and inaccessible buildings. The latter buildings were all assumed to be residential buildings and to follow the average conditions in their respective pixels. [Eq. 3] was used to extrapolate the dwelling units in these buildings.

\[
dwelling units_{k,j} = \min \left( \frac{x_{k,j} \times N_j}{\sum_i x_{i,j} \times y_{i,j}}, 1 \right)
\]  

[Eq.3]

Where,

\begin{align*}
dwelling units_{k,j}: & \quad \text{Restored dwelling units in building } k \text{ in pixel } j. \\
x_{k,j}: & \quad \text{Top surface area of building } k \text{ in pixel } j. \\
N_j: & \quad \text{Number of dwelling units in pixel } j. \\
y_{i,j}: & \quad \text{Survey accessibility of building } i \text{ in pixel } j.
\end{align*}

The summary of the collected and derived data with their respective resolutions, sources, descriptions, and formulae is provided in table 3.

| Data                  | Resolution | Source                  | Description                                                                 |
|-----------------------|------------|-------------------------|-----------------------------------------------------------------------------|
| Occupied Dwelling Units | 250 m      | 2015 National Census    | The number of occupied dwelling units in any given pixel. Obtained as the number of households in that pixel |
| Dwelling Units        | Building level | Field survey            | The number of residential mailboxes in any given residential building or alternatively the number of gas meters or nameplates |
| Vacancy Ratio         | 250 m      | 2015 National Census and field survey | The vacancy ratio for any given pixel \( vacancy ratio = \frac{100 \times (dwelling units - occupied dwelling)}{dwelling units} \) |
| Apartments            | Building Level | Field survey            | The number of apartments in any given residential building. If any apartments = \( \sum apartment index \times dwelling units \) |
| Residential Buildings | 250 m      | Field survey            | The number of residential buildings in any given pixel \( \sum residential index \times dwelling units \) |
| Apartment Ratio      | 250 m      | Field survey            | The ratio of apartments to dwelling units in any given pixel \( apartment ratio = \frac{\sum apartment index \times dwelling units}{\sum dwelling units} \) |

4. Zonal Variation of the Residential Characteristics of Study Pixels

The Japanese zoning system imposes regulatory restrictions in terms of land use, density, height, and building characteristics to maintain an environment within a specified quality of urban areas (Akashi, 2007). As the collected and derived data in this study mainly describe the residential environment of the study pixels, they will be impacted by these zoning regulations. Consequently, it is essential to evaluate the zonal variation of the study’s data to consider the effect of the imposed regulations. Moreover, as the vacancy estimation models will be dependent
on the zoning regulations of the target areas, analyzing the zonal variations of the data will serve as a tool to determine the appropriate indices used for the estimation models.

The range of the apartment ratio is larger for the mid/high-rise class than in the exclusively low-rise class. This is apparent in figure 4, where the histogram of the data shows that the exclusively low-rise pixels’ apartment ratio is concentrated around lower values while the mid/high-rise class’ pixels show a wide distribution along the apartment ratio spectrum. Overall, the residential class had the largest values. This implies that the impact of apartments on household composition and vacancy values variation is the largest for the mid-high class’ pixels.

The effect of the apartment ratio on dwelling characteristics is apparent when comparing the variations of the volume of residential buildings and dwelling units. As shown in figure 5, the mid/high-rise class had the most residential buildings followed by the residential class and then the exclusively low-rise class. Nevertheless, when the dwelling units within the residential buildings are considered, the residential class has the largest values of dwelling units followed by the mid/high rise and exclusively low-rise classes, respectively as shown in figure 6. Moreover, the difference between the classes’ dwelling units values is larger compared to that of residential buildings. This implies that a combination of residential buildings and apartment ratio values as indices for the dwelling units in a given area is more suitable than exclusively using the volume of residential buildings. Furthermore, the volume of households follows a pattern similar to that of the dwelling units as shown in figure 7, making it a potential estimator for dwelling units.
Figure 5. Zonal Variation of the Volume of the Residential Buildings of Study Pixels.

Figure 6. Zonal Variation of the Volume of Dwelling Units of Study Pixels.

Figure 7. Zonal Variation of the Volume of Households in Study Pixels.
The vacancy data of the surveyed pixels are summarized in table 4. All the study pixels had positive vacancy ratios, except pixel 21 which implies that the number of dwelling units in that pixel is underreported. As the survey coverage in that pixel is 98%, it is presumed that the underreporting is a result of an error in the surveying process rather than the inability of [Eq.3] to restore inaccessible buildings. Moreover, pixel 13 is an outlier in the mid/high-rise class as it has an extremely low vacancy ratio when compared with the rest of the class as shown in figure 8. For these reasons, pixels 13 and 21 will be disregarded from further analysis.

| Class            | Pixel | Occupied Dwelling Units [dwelling units] | Vacant Dwelling Units [dwelling units] | Vacancy Ratio [%] |
|------------------|-------|------------------------------------------|----------------------------------------|-------------------|
| Exclusively Low-rise | 1     | 265                                      | 27                                     | 9.4               |
|                  | 2     | 229                                      | 64                                     | 21.8              |
|                  | 3     | 327                                      | 67                                     | 17.0              |
|                  | 4     | 223                                      | 21                                     | 8.6               |
|                  | 5     | 295                                      | 9                                      | 3.0               |
|                  | 6     | 344                                      | 75                                     | 18.0              |
|                  | 7     | 250                                      | 21                                     | 7.6               |
|                  | 8     | 303                                      | 57                                     | 15.8              |
|                  | 9     | 253                                      | 72                                     | 22.2              |
|                  | 10    | 220                                      | 22                                     | 9.2               |
| Mid/high-rise    | 11    | 516                                      | 84                                     | 14.0              |
|                  | 12    | 538                                      | 71                                     | 11.7              |
|                  | 13    | 502                                      | 10                                     | 1.9               |
|                  | 14    | 214                                      | 24                                     | 10.1              |
|                  | 15    | 338                                      | 27                                     | 7.3               |
|                  | 16    | 447                                      | 51                                     | 10.2              |
|                  | 17    | 354                                      | 50                                     | 12.4              |
|                  | 18    | 607                                      | 75                                     | 11.0              |
|                  | 19    | 453                                      | 77                                     | 14.5              |
|                  | 20    | 500                                      | 130                                    | 20.6              |
| Residential      | 21    | 415                                      | -11                                    | -2.8              |
|                  | 22    | 496                                      | 31                                     | 6.0               |
|                  | 23    | 675                                      | 113                                    | 14.4              |
5. House Vacancy Estimation

Given that the number of occupied dwelling units within the study pixels is already known through the household numbers provided by the census, the volume of vacant dwelling units can be estimated through accurate estimation of the dwelling units of that pixel using [Eq.1]. In this model, the dwelling units of the study pixels will be estimated through a linear regression using independent variables that describe the residential environment within these pixels. The selection of the independent variables will be zonal (i.e., each zone will have different independent variables). It is to be noted that no model is developed for the residential class due to the small sample size collected in this study.

5.1. Exclusively Low-rise Class

For the exclusively low-rise class pixels, the volume of residential buildings and occupied dwelling units are used as independent variables. This is due to the strong influence of both variables on the volume of dwelling units.

The summary of the regression results for the exclusively low-rise class is provided in table 5. As apparent from the coefficient of determination, the model explained most of the variations in the real dwelling units’ values. Moreover, the independent variables were significant in estimating the volume of dwelling units in the exclusively low-rise class. The accuracy of the estimations is evaluated using the root mean square error RMSE [Eq.4] and mean absolute error MAE [Eq.5]. Subsequently, the estimated volume of vacant dwelling units was obtained using [Eq.1]. As shown in table 6, the overall accuracy of the model is average except for pixel 5 which has a vacancy ratio significantly smaller than the rest of the class’ pixels.

\[
RMSE = \sqrt{\frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}{n}} \tag{Eq.4}
\]

\[
MAE = \frac{\sum_{i=1}^{n} |\hat{y}_i - y_i|}{n} \tag{Eq.5}
\]
Where,

\(i\): Pixel \(i\) in the appropriate class.

\(n\): Sample size of the appropriate class.

\(\hat{y}_i\): Actual value of the dependent variable of pixel \(i\).

\(y_i\): Estimated value of the dependent variable of pixel \(i\).

### Table 5. Summary of the Results of the Dwelling Units Estimation (Exclusively Low-rise Class).

| Sample Size | Coefficient of Determination "R-Squared" | 0.93 |
|-------------|------------------------------------------|------|
|             | RMSE [dwelling units] | 15.55 |
|             | MAE [dwelling units] | 13.23 |

| Coefficient | Significance [P-value] |
|-------------|------------------------|
| Intercept   | -58.68                 | -    |
| Residential Buildings | 0.48 | ** |
| Occupied Dwelling Units | 0.95 | *** |

*: \(p < 0.1\); **: \(p < 0.05\); ***: \(p < 0.01\); *: \(p > 0.1\)

### Table 6. Summary of the Results of the Vacant Dwelling Units Estimation (Exclusively Low-rise Class).

| Pearson Correlation Coefficient | 0.77 |
|----------------------------------|------|

| Pixel | Vacant Dwelling Units [ dwelling unit ] | Estimated Vacant Dwelling Units [ dwelling unit ] | Error [%] |
|-------|----------------------------------------|-----------------------------------------------|----------|
| 1     | 27                                     | 19                                            | 28       |
| 2     | 64                                     | 40                                            | 37       |
| 3     | 67                                     | 77                                            | 15       |
| 4     | 21                                     | 31                                            | 45       |
| 5     | 9                                      | 39                                            | 301      |
| 6     | 75                                     | 67                                            | 11       |
| 7     | 21                                     | 24                                            | 21       |
| 8     | 57                                     | 36                                            | 37       |
| 9     | 72                                     | 66                                            | 8        |
| 10    | 22                                     | 35                                            | 58       |

### 5.2. Mid/high-rise Class

Similar to the exclusively low-rise class, the occupied dwelling units will be used as an independent variable. Nonetheless, due to the high variation inflicted on the volume of dwelling units by the apartment ratio variation, an apartment ratio index will be used as the additional independent variable. In the context of this study, the ratio of occupied dwelling units to residential buildings provides an accurate apartment ratio index as shown in figure 9.

The results of the estimation model for the mid/high-rise class pixels are summarized in table 7. It can be deduced that the estimated volumes of dwelling units explain most of the variation of the actual dwelling units in the mid/high-rise class. Subsequently, the vacant dwelling units were estimated using [Eq.1]. The summary of the results shows that the model’s vacant dwelling units estimation accuracy is satisfactory [table 8].
Figure 9. Linear Regression Between the Apartment Ratio and the Occupied Dwelling Units to Residential Buildings Ratio (Mid/high-rise Class).

Table 7. Summary of the Results of the Dwelling Units Estimation Model (Mid/high-rise Class).

| Sample Size |                  | 9 |
|-------------|------------------|---|
| Coefficient of Determination "R-Squared" |                  | 0.99 |
| RMSE [dwelling units] |                  | 11.62 |
| MAE [dwelling units] |                  | 9.80 |

| Coefficient | Significance [P-value] |
|-------------|------------------------|
| Intercept   | **                     |
| Occupied Dwelling Units to Residential Buildings Ratio | *** |
| Occupied Dwelling Units | *** |

*: p < 0.1; **: p < 0.05; ***: p < 0.01; -: p > 0.1

Table 8. Summary of the Results of the Vacant Dwelling Units Estimation (Mid/high-rise Class).

| Pearson Correlation Coefficient |                  | 0.93 |
|-------------------------------|------------------|------|
| Pixel                         | Vacant Dwelling Units [dwelling unit] | Estimated Vacant Dwelling Units [dwelling unit] | Error [%] |
| 11                            | 84               | 95   | 13   |
| 12                            | 71               | 76   | 7    |
| 14                            | 24               | 22   | 10   |
| 15                            | 27               | 45   | 68   |
| 16                            | 51               | 61   | 20   |
| 17                            | 50               | 39   | 22   |
| 18                            | 75               | 74   | 1    |
| 19                            | 77               | 56   | 27   |
| 20                            | 130              | 121  | 7    |
5.3. Model Cross-validation

The models’ predictive performance against new data was evaluated using the “Leave-one-out” cross-validation method. The root mean square percentage error RMSPE [Eq.6] and mean absolute percentage error MAPE [Eq.7] were used to validate the accuracy of the model. The results are as shown in table 9. For the exclusively low-rise class case, the overall predictive performance of the model was average except for the fifth iteration. In that regard, it can be deduced that the performance of the model against pixels with low vacancy ratios (i.e., less than 6%) is subpar due to the limited low-vacancy ratio training samples. On the other hand, the mid/high-rise class model’s predictive performance was satisfactory as evident in its RMSPE and MAPE values.

\[
RMSPE = \sqrt{\frac{1}{n} \sum_{i} \left(\frac{\hat{y}_i - y_i}{\hat{y}_i}\right)^2}
\]  

[Eq.6]

\[
MAPE = \frac{1}{n} \sum_{i} \left|\frac{\hat{y}_i - y_i}{\hat{y}_i}\right|
\]  

[Eq.7]

5.4. Validation Against an Existing Vacancy Estimation Method

The precision of the developed vacancy estimation model was compared to the model developed by Akiyama, et al. (2018). This method estimates the volume of vacant detached houses in Kagoshima City with a resolution of 250 m. The information used to estimate the volume of vacant dwelling units within a given pixel consists of the building resident register, building registration information, and closed hydrant information. Moreover, this method relies on supplementary information such as the presence of nameplates. This is in contrast with the simpler model introduced in this study, which significantly facilitates the replicability of its method.

As shown in table 10, the several precision indices are similar in both methods. However, both methods were applied to different urban settings (i.e., Kyoto and Kagoshima). Furthermore, Akiyama, et al. (2018) estimated housing vacancies on the detached house level compared to this study’s dwelling units level. Moreover, unlike the current study, Akiyama, et al. (2018) have no restriction on the zoning areas considered in their model. Therefore, direct comparison between the precision of the two methods is inconclusive. Nonetheless, the similarity of the precision indices in both studies is a promising result given the simplicity of this study’s models.
### Table 9. Summary of the Results of the Cross-validation Process of the Dwelling Units Estimation Model.

| Dwelling Units | Iteration | Actual [dwelling units] | Estimation [dwelling units] | RMSPE | MAPE |
|----------------|-----------|-------------------------|----------------------------|-------|------|
|                | 1         | 271                     | 275                        | 0.07  |      |
|                | 2         | 394                     | 412                        |       |      |
|                | 3         | 242                     | 260                        |       |      |
|                | 4         | 293                     | 262                        |       |      |
|                | 5         | 304                     | 341                        |       |      |
|                | 6         | 360                     | 332                        |       |      |
|                | 7         | 244                     | 256                        |       |      |
|                | 8         | 419                     | 406                        |       |      |
|                | 9         | 325                     | 314                        |       |      |
|                | 10        | 292                     | 281                        |       |      |

| Vacant Dwelling Units | Actual [dwelling units] | Estimation [dwelling units] | RMSPE | MAPE |
|-----------------------|-------------------------|----------------------------|-------|------|
|                       | 1                       | 21                        | 25    | 1.34|
|                       | 2                       | 67                        | 85    |      |
|                       | 3                       | 64                        | 40    |      |
|                       | 4                       | 9                         | 33    |      |
|                       | 5                       | 57                        | 46    |      |
|                       | 6                       | 21                        | 29    |      |
|                       | 7                       | 75                        | 33    |      |
|                       | 8                       | 72                        | 62    |      |
|                       | 9                       | 27                        | 16    |      |

### Table 10. Summary of the Comparison Between Akiyama, et al. (2018) Vacant Dwellings Estimation Method and the Current Study Estimation Method.

|                      | Akiyama, et al. (2018) | Current Study |
|----------------------|------------------------|---------------|
| Observations         | 24                     | 19            |
| Pearson Correlation Coefficient | 0.89                  | 0.88          |
| Average Ratio (Estimation/Actual) | 1.21                  | 1.20          |
| Total Ratio (Estimation/Actual)   | 1.01                  | 1.00          |
| RMSPE                 | 1.06                   | 0.81          |
| MAPE                  | 0.44                   | 0.40          |

### 6. Conclusion

This study focused on developing tools for house vacancy surveying and estimation to accelerate research related to the spatial distribution of vacant houses in Japan. The need for efficient and reliable house vacancy estimation and surveying methods is evident in the literature due to the uneconomical nature of existing surveying methods. In this study, a new house vacancy
surveying method that provides house vacancy data with a resolution of 250 m is introduced. This method is the first step in developing economic studies that evaluate the factors shaping the spatial distribution of vacant houses on the neighborhood level in Japanese cities.

The new surveying method was employed in selected parts of Kyoto city in an extensive field survey which additionally identified the zonal variation of the residential environment of the study pixels. Using the latter findings, a dwelling units estimation model was developed for the exclusively low-rise and mid/high-rise residential zones. The results from the aforementioned estimation model succeeded in extrapolating the volume of vacant houses in the study pixels with varying accuracies. The evaluation of the predictive performance of the models against new data during the cross-validation process highlighted the need for more testing using low-vacancy ratio pixels. Consequently, at its current state, this model represents a powerful tool for researchers and municipalities to determine the volume of vacant houses in pixels from the exclusively low-rise and the mid/high-rise classes and identify areas with a large concentration of vacant houses within these residential zones. Furthermore, the precision of the model is similar to the model developed by Akiyama, et al. (2018) which identifies vacant detached houses on the same spatial resolution. Although the different objectives and target area of Akiyama, et al. (2018)’s model results in the comparison being inconclusive, the simplicity of this paper’s model represents a potential advantage as a complete housing vacancy surveying tool.

As the developed estimation models rely on the zonal variation of the residential characteristics of the pixels, they are unique to the exclusively low-rise and mid/high-rise classes. No models were developed for the residential class due to the small sample size of exclusive residential zones present in Kyoto city, moreover, pixels in non-residential zones were not considered in this study. Consequently, estimation models for non-residential zones and pixels that host a mixture of land-use zones need to be developed. Ultimately, expanding the model from its zonal dependence is the next step for its consideration as a complete tool for vacancy estimation in Japanese cities. This will free the application of the model from any zonal pre-requisites such as the presence of a lower threshold of all zones in a given city.

References

Abe, S., Nakagawa, D., Matsunaka, R. & Oba, T., 2014. Study on the Factors to Transform Underused Land Focusing on the Influence of Railway Stations in Central Areas of Japanese Local Cities. Land Use Policy, Volume 41, pp. 344-356.

Akashi, T., 2007. Urban Land Use Planning System in Japan, World: JICA-Net.

Akiyama, Y. et al., 2018. Monitoring of Spatial Distribution of Vacant Houses Using Municipal Public Data In Kagoshima City, Kagoshima Prefecture. A study on the Estimation Method of Spatial Distribution of Vacant Houses Using Municipal Public Data (Part 1). Journal of Architecture and Planning (Transactions of AIJ), 83(744), pp. 275-283. In Japanese

Baba, H. & Asami, Y., 2017. Regional Differences in the Socio-economic and Built-environment Factors of Vacant House Ratio as a Key Indicator for Spatial Urban Shrinkage. Urban and Regional Planning Review, Volume 3, pp. 251-267.

Chen, Z. et al., 2015. Estimating House Vacancy Rate in Metropolitan Areas Using NPP-VIIRS Nighttime Light Composite Data. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 8(5), pp. 2188-2197.

GU, J. & Asami, Y., 2016. Vacant Housing Market in Tokyo 23 Wards: Based on Landlord's Optimal Search Model. Urban and Regional Planning Review, Volume 3, pp. 31-49.
Ishikawa, M. et al., 2016. Estimation of Current and Future Distribution of Vacant Dwellings Within Municipalities - With a Focus on The Difference Between the Number of Dwellings and Households -. *Journal of the City Planning Institute of Japan*, 51(3), pp. 833-838. In Japanese

Kumagai, K., Matsuda, U. & Ono, Y., 2016. Estimation of Housing Vacancy Distributions: Basic Bayesian Approach Using Utility Data. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Volume XLI-B2, pp. 709-713.

Martinez-Fernandez, C. et al., 2016. Shrinking Cities in Australia, Japan, Europe and the USA: From a Global Process to Local Policy Responses. *Progress in Planning*, Volume 105, pp. 1-48.

Ministry of Land, Infrastructure, Transport and Tourism, 2017. *Trends Concerning Land in FY2016. Basic Measures in Relation to Land in FY2017*, s.l.: Ministry of Land, Infrastructure, Transport and Tourism.

National Institute of Population and Social Security Research in Japan, 2017. *Population Projections of Japan (2017): 2016 to 2065*. [Online] Available at: http://www.ipss.go.jp/pp-zenkoku/e/zenkoku_e2017/pp_zenkoku2017e.asp [Accessed 5 2019].

National Statistics Center, 2015. *E-Stat. Statistics on a Map (Statistical GIS) (E-Stat. Chizu de miru tōkei (tōkei GIS)).* [Online] Available at: https://www.e-stat.go.jp/ [Accessed 05 07 2019]. In Japanese

Nomura Research Institute, 2016. *The 2030 Housing Market. ~An Increase in the Mobile Population is the Key to Invigorating the Housing Market in an Era With a Declining Population~*. [Online] Available at: https://www.nri.com/en/knowledge/report/lst/2016/cc/papers/0607 [Accessed 6 2019].

OpenStreetMap Contributors, 2019. *Plant Dump retrieved from https://planet.osm.org*. [Online] [Accessed 10 2019].

Rybczynski, W. & Linneman, P. D., 1999. How to save our shrinking cities. *Public Interest*, Volume 135, pp. 30-44.

Sadayuki, T., Kanayama, Y. & Arimura, T. H., 2019. Evaluating the Externality of Vacant Houses in Japan: The Case of Toshima Municipality, Tokyo. *RIEEM Discussion Paper Series*, Issue 1901.

Sakamoto, K., Iida, A. & Yokohari, M., 2017. Spatial Emerging Patterns of Vacant Land in a Japanese City Experiencing Urban Shrinkage A Case Study of Tottori City. *Urban and Regional Planning Review*, Volume 4, pp. 111-128.

So, T., 2017. Verifying Vacancy Rates Derived From the Housing and Land Survey. *Journal of Architecture and Planning (Transactions of AIJ)*, 82(737), pp. 1775-1781. In Japanese

Statistics Bureau, Ministry of Internal Affairs and Communications, 2015. *Statistics Dashboard*. [Online] Available at: https://dashboard.e-stat.go.jp/en/dataSearch [Accessed 06 2019].

Statistics Bureau, D.-G. f. P. P. (. S. a. S. R. a. T. I., 2015. *IV Explanation of The Terms Used in The Results of The National Census. (IV Kokuseichōsa no kekka de mochiru yōgo no kaisetsu).*
[Online]
Available at: https://www.stat.go.jp/data/kokusei/2015/users-g/pdf/04.pdf
[Accessed 17 12 2019]. In Japanese

Wiechmann, T. & Pallagst, K. M., 2012. Urban Shrinkage in Germany and the USA: A Comparison of Transformation Patterns and Local Strategies. *International Journal of Urban and Regional Research*, 36(2), pp. 216-280.