BUILIDING A MODEL FOR FUTURE ESTIMATION OF VACANT HOUSES DISTRIBUTION USING GOVERNMENT STATISTICS

Tokyo City University Member ○Kotaro MIZUTANI
Tokyo City University Member Yuki AKIYAMA
Kyoto University Non-Member Hiroki BABA

1. OBJECTIVE

In Japan, the number of vacant houses is increasing due to social factors such as declining birth rates, aging populations, and migration from small cities to large ones. According to the 2018 Housing and Land Survey, approximately 3.49 million are potentially regarded as mismanaged vacant houses, accounting for 5.6% of the total[1]. The presence of vacant houses causes external diseconomies, including pest infestations and risks such as roof tiles and other materials being dislodged by typhoons and other severe weather conditions. Local governments have initiated measures aimed at repurposing or decommissioning and formulating future plans for vacant house countermeasures in response to the "Act on Special Measures Concerning Promotion of Measures against Vacant Houses and Vacant Buildings" enacted in May 2015. However, while the quantity or proportion of vacant houses is a key factor in formulating effective plans to deal with the problem in the future, there are a few examples demonstrating the accurate estimation of the number of vacant houses. In addition, it is important to note that the existing estimates face limitations such as being restricted to macro-scale estimates at the prefectural level and their estimation accuracy is often inadequate[2), 3), 4)]. Therefore, this study aims to construct a model to estimate the present and future vacant house rate at the municipal level in Japan by utilizing the National Census and the Housing and Land Survey, and to achieve results that can contribute to the planning of future vacant house countermeasures.

2. METHOD

(1) Database construction and Prediction methods

This study utilized data from the National Census and the Housing and Land Survey. The explanatory variables used in this study were the National Census data (in 2000, 2005, 2010, 2015, and 2020), and the dependent variable was the vacancy rate, which was defined as the number of the "other dwellings" vacant houses excluding those with confirmed usage, divided by the total number of houses, based on the Housing and Land Survey data (in 2003, 2008, 2013, and 2018). The study predicted the vacancy rate 8 years later for all municipalities in Japan. Prediction models for the vacancy rate were created at each time epoch, and these models were then extrapolated using census data from future years. The models relied solely on National Census information that was consistent between the years 2000 and 2020, as the disclosure of National Census data varies by year.

(2) Building machine learning model

In this study, we used the LightGBM regression model as a machine learning model. We conducted accuracy verification by comparing the estimated values obtained from the model with the actual vacancy rates obtained from the Housing and Land Survey. The evaluation metrics were the determination coefficient (R2) and the root mean squared error (RMSE).

3. RESULT

(1) Model evaluation

Table 1 shows the estimated vacancy rates after 8 years at each time epoch, all of which are generally highly accurate. Additionally, the extrapolation results indicate that the model developed is highly versatile and that its accuracy is assured.

(2) Analysis of estimated factors

Next, we utilized the SHAP (Shapley Additive exPlanations) method to investigate the key factors that influenced the estimated vacancy rate. **Fig.1** shows a beeswarm plot based on the SHAP values for the highly influential features. Our findings indicate that the areas with

Table 1 Estimation of vacancy rate after 8 years at each time epoch.

	2000→2008		2005→2013		2010→2018	
	R2	RMSE	R2	RMSE	R2	RMSE
2000→2008	**0.6851**	**0.0182**	*0.7441*	*0.0207*	*0.6590*	*0.0265*
2005→2013	-	-	**0.7924**	**0.0186**	*0.7649*	*0.0220*
2010→2018	-	-	-	-	**0.8150**	**0.0195**

Note.1 : XXXX→YYYY : where XXXX is the year of the National Cencus Survey and YYYY is the year of the prediction.
Note.2 : **Bolded values** indicate the prediction accuracy of the validation data, and *italicized values* indicate the accuracy of the extrapolation.

Keywords Vacant house, Future prediction, Government statistics, Machine learning, Urban planning
Contact 1-28-1, Tamazutsumi, Setagaya-ku, Tokyo 158-8557, Japan.
Tokyo City University TEL 03-5707-0104

a high percentage of people aged 65 or above and a low percentage of people within the 15-64 age bracket, as well as areas with a high percentage of owner-occupied homes, show a rising trend in vacancy rate.

(3) Analysis of estimated vacancy rates throughout Japan

The validation results for the model constructed using the National Census (2010) as the explanatory variable and the Housing and Land Survey (2018) as the objective variable are presented in **Fig.2**, and **Fig.3** shows the projected vacancy rate in 2028, extrapolated from the National Census (2020) using the same model. Based on these figures, the vacancy rate in Japan is increasing in numerous regions, with the average vacancy rate per municipality increasing from 9.40% in 2018 to 12.32% in 2028, an increase of approximately 3 percentage points. According to our projections, the vacancy rate is expected to increase by 2028 not only in mountainous areas at such as Chugoku, Shikoku, and Kyushu regions, where the rate was already high in 2018, but also nationwide, except for metropolitan areas like Tokyo and Osaka. This highlights the need for nationwide measures to address the issue. Furthermore, this projection can be applied to towns and villages with a population of less than 15,000, which are not covered by the Housing and Land Survey, allowing for an assessment of the relative and future vacancy rate in these previously unsurveyed areas.

4. FUTURE PROSPECTS

In this study, we constructed a model using open data from the National Census and the Housing and Land Statistics Survey at the municipal level to estimate the distribution of vacant houses. In the future, we plan to use data at the basic survey area level, which is more micro in spatial scale than the municipal level, for more precise estimation. To promote measures against vacant houses and national land planning based on the estimation results, stakeholders from industry, government, and academia are developing a dashboard system that allows them to view the results on a map. Furthermore, we plan to continue conducting interviews with local governments to clarify their needs in order to accelerate actions toward promoting measures against vacant houses using the results of this study.

ACKNOWLEDGMENTS

This study is supported by the Digital Urban Spatial Information Research and Development Unit, Advanced Research Institute, Tokyo City University. We are gratefully deep gratitude them.

REFERENCES

1) Ministry of Land, Infrastructure and Transport : Curr entstatus and issues of vacant house policy and dire ction of study, <https://www.mlit.go.jp/policy/shingikai /content/001518774.pdf>, (Last viewed on March 8, 2023).
2) Yuko, K., Toshinori, A. and Keisuke, M. : Factor an alysis and estimation of rate of vacant dwellings, *Jo urnal of the City Planning Institute of Japan*, Vol. 5 0, No. 3, pp. 1017-1024, 2015.
3) Yuki, A. and Ryosuke, S. : Future Estimation of Va cant House Distribution Using Micro Estimation Dat a of Future Population, *Research Abstracts on CSIS DAYS 2015*, C09, 2015.
4) Masahiro, I., Keisuke, M., Yuko, K. and Toshinori, A: Method of grasping detailed regional distribution of v acant dwelling based on the number of dwellings and households, *Journal of the City Planning Institute of Japan*, Vol. 51, No. 3, pp. 833-838, 2016.

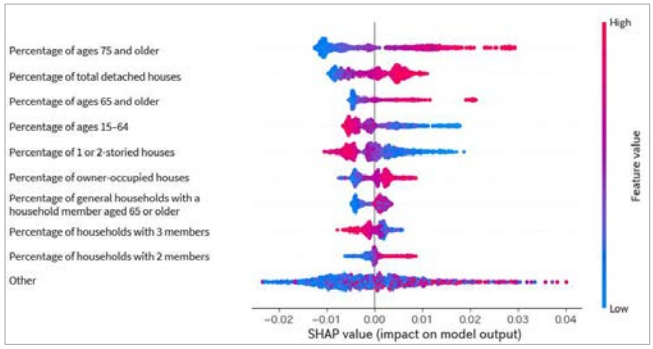

Fig.1 Beeswarm plot based on SHAP values.

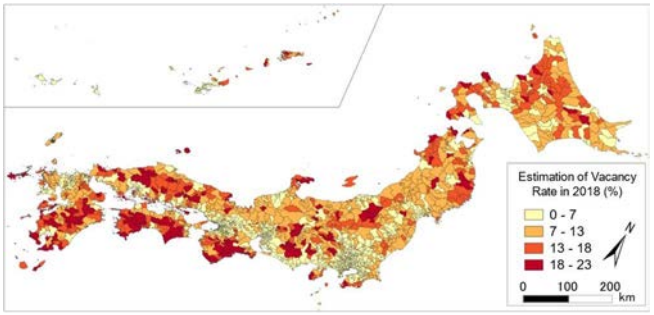

Fig.2 Estimation of vacancy rate in 2018.

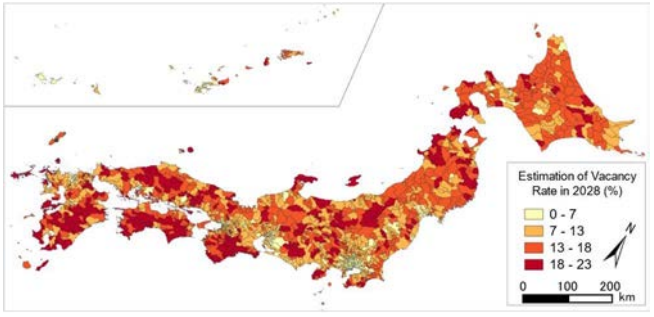

Fig.3 Estimation of vacancy rate in 2028.