pISSN: 1226-7147, eISSN: 2383-9171
https://doi.org/10.17208/jkpa.2025.08.60.4.109

Journal of Korea Planning Association Vol.60, No.4 (2025)
대한국토·도시계획학회지「국토계획」제60권 제4호 pp.109-130

Special Issue

Predicting Housing Vacancies in Depopulating Areas with Spatial GAMLSS*

Lee, So-Yeong** · Bae, Min-Cheul*** · Joo, Hee-Sun****

Abstract

Housing vacancy, exacerbated by population aging and decline, generates negative externalities that extend beyond property owners, impacting real estate markets and community stability. This study examines housing vacancy patterns and determinants in Jinju-si, distinguishing urban and non-urban areas. Using Spatial GAMLSS and hotspot analysis, the research identifies high-risk zones and highlights regional disparities. Urban areas face higher risks in older neighborhoods, declining commercial districts, and poorly connected zones, while non-urban areas with weakened economic bases or limited accessibility also show elevated risks. Conversely, regions with robust infrastructure or economic activity exhibit lower vacancy probabilities. The findings emphasize the importance of tailored vacancy management policies, such as downtown revitalization, transit-oriented development, and economic base strengthening. By providing spatial insights into high-risk zones, this study supports effective resource allocation and targeted interventions for sustainable housing management.

Keywords Housing Vacancy, Urban and Non-Urban Disparities Spatial GAMLSS, Hotspot Analysis, Vacancy Management

Ⅰ. Introduction

The issue of vacant housing has long been recognized as a critical urban challenge worldwide, especially in small and medium-sized cities where vacancy rates in single-family houses have steadily risen due to population decline, aging, and changing housing preferences. This trend disrupts local housing markets and contributes to environmental degradation, heightened crime risks, and stagnation in urban regeneration and regional economic development. Systematically analyzing the causes of housing vacancy is thus essential for effective urban management (Halvorsen and Pollakowski, 1981), and predicting high-risk areas is key to developing proactive policy responses.

Previous studies have employed statistical models based on spatial factors, population density, and accessibility to examine vacancy causes. However, these approaches often overlook spatial interdependencies—the influence of adjacent areas on vacancy patterns—resulting in potentially biased estimations. This highlights the importance of incorporating spatial interactions to enhance both accuracy and policy relevance in vacancy prediction (Basu and Thibodeau, 1998). The Spatial Generalized Additive Models for Location, Scale, and Shape (GAMLSS) model addresses this gap by integrating spatial autocorrelation and capturing nonlinear relationships among variables, offering improved explanatory performance in vacancy analysis.

This study applies the Spatial GAMLSS model to estimate

---

* This study was partially supported by the 2023 National Research Fund of Korea (Grant number: RS-2024-00339015).

** Ph.D. Candidate, Department of Urban Engineering, Gyeongsang National University (First Author: sodol0111@gnu.ac.kr)
*** Postdoctoral, Department of Urban Engineering, Gyeongsang National University (mcbae@chari.re.kr)
**** Associate Professor, Department of Urban Engineering, Gyeongsang National University (Corresponding Author: hsjoo@gnu.ac.kr)

vacancy probabilities and identify high-risk areas in Jinju-si, aiming to support the development of tailored housing management strategies. It further enhances estimation accuracy by leveraging the model's flexibility, as proposed by Rigby and Stasinopoulos (2005), in capturing complex spatial and distributional structures.

The main objectives of this study are as follows:

1. To estimate and compare the probability of housing vacancies in urban and non-urban areas of Jinju-si using the Spatial GAMLSS model;
2. To analyze the spatial effects of housing vacancies by incorporating regional characteristics and spatial autocorrelation.

The findings will provide empirical evidence for formulating more refined and spatially responsive housing policies and serve as a foundation for long-term vacancy management strategies.

Ⅱ. Related Works

1. Characteristics of Housing Vacancies

The occurrence of vacancies in single-family houses is influenced by a complex interplay of various factors, including economic, social, physical, topographical, environmental, and locational characteristics. Previous studies have demonstrated that factors influencing housing vacancies arise from various dimensions. These studies have sought to identify the causes of vacancies by categorizing them into levels such as the housing unit itself (including individual houses and parcels), the community, and the urban or regional scale. This includes individual building factors related to vacant houses, such as size, area, building structure, road width, and building age (Baba and Hino, 2019; Morckel, 2013; Nadalin and Igliori, 2017; Yin and Silverman, 2015). Studies examining socioeconomic characteristics have identified population variables as one of the key factors influencing housing vacancies. Most studies have emphasized that demographics are a primary cause of housing vacancies and that their impact varies depending on the regional context (Bassett et al., 2006; Edward and Joseph, 2005; Immergluck, 2016; Mallach, 2017; Morckel, 2013; Nassauer and Raskin, 2014; Page, 2001; Ribant and Chen, 2020; Wilhelmsson et al., 2011). Additionally, factors such as commercial services, surrounding landscape quality, environmental conditions, topography, and location have been examined as influencing the occurrence of housing vacancies. Some studies have indicated that areas with higher land values tend to exhibit higher vacancy rates (Ma et al., 2020). Discussions on environmental characteristics are closely linked to quality of life, emphasizing factors such as infrastructure accessibility, environmental hygiene, and commuting convenience (Chen et al., 2023; Deng and Ma, 2015; Joo and Lee, 2021; Hackman et al., 2019; Terada et al., 2016).

Generally, older, smaller, deteriorated, inexpensive, and poorly maintained houses have a higher probability of becoming vacant (Baba and Shimizu, 2023; Gu et al., 2019; Morckel, 2014; Sargent et al., 1997). Studies highlighting the importance of land size and price in maintaining housing value (Zhao, 2022) and the role of negative externalities from surrounding environments in promoting housing vacancies (Joo et al., 2022) underscore the necessity of proactive measures to address vacancies.

2. Modeling Approaches to Predicting Housing Vacancy

Predicting housing vacancy has long been a critical research topic across urban studies, real estate economics, and spatial planning. Early studies primarily adopted regression-based frameworks such as logistic regression, Ordinary Least Squares (OLS), and Generalized Linear Models (GLM) to examine the relationships between vacancy occurrence and various socioeconomic or physical attributes. These models were often chosen for their interpretability and statistical rigor. However, they operate under the assumptions of linearity, homoscedasticity, and normally distributed residuals, which may not hold in the context of highly skewed or non-normal vacancy data. A fundamental distinction must be made between explanatory and predictive modeling. Shmueli (2010) emphasizes that models designed for explanation are not necessarily optimal for prediction, and vice versa. In this respect, much of the early literature on housing vacancy focused on hypothesis testing and variable interpretation, rather than on optimizing predictive performance.

Recent advances in computational capacity and the availability of large-scale urban data have shifted scholarly focus toward prediction-oriented modeling frameworks, particularly those emerging from the fields of machine learning and nonparametric statistics. Machine learning algorithms such as decision trees, random forests, support vector machines, and artificial neural networks have been increasingly applied to forecast vacancy occurrences or to classify high-risk housing units (Martin et al., 2017; Newman et al., 2016). These models are data-driven and flexible in functional form, thus well-suited for capturing nonlinear and high-order interactions that are difficult to specify a priori.

Several studies distinguish between binary prediction models (i.e., predicting whether a house is vacant) and risk scoring models that generate continuous vacancy risk indices (Appel et al., 2014; Lee and Newman, 2017; Lee et al., 2021). While machine learning approaches have generally shown higher predictive accuracy, they often lack comparative assessments and are criticized for their “black-box” nature, particularly in policy contexts. Amandolia (2021) contributes a rare comparative evaluation of multiple nonparametric methods, underscoring the importance of validating predictive accuracy across models rather than relying on a single technique.

Nevertheless, most existing models—whether parametric or nonparametric—fail to incorporate spatial autocorrelation, despite the inherently geographic nature of housing vacancy. Spatial proximity often exerts a significant influence on vacancy risk due to neighborhood effects, service access, or physical decline. In response to this gap, this study employs Spatial GAMLSS, which integrates spatial random effects into a flexible distributional modeling framework. The theoretical structure and implementation of the Spatial GAMLSS model are detailed in Section 3.2.

3. Research Gap and Contribution

This study addresses several key gaps identified in the existing literature on housing vacancy prediction and offers new contributions from both methodological and empirical perspectives.

First, previous studies on housing vacancy have predominantly employed traditional statistical models such as multiple regression or hedonic pricing models. While these approaches are effective for identifying linear associations, they often fail to capture nonlinear effects and spatial autocorrelation, both of which are crucial in understanding the geographic distribution of vacancy risk. This study addresses this methodological limitation by employing the Spatial GAMLSS framework. This model offers a flexible structure capable of estimating vacancy probabilities with higher accuracy by accounting for distributional complexity and spatial dependence—factors overlooked in most prior work.

Second, much of the existing research has relied on either aggregated regional data or small-sample household surveys, which limits the spatial granularity and inferential strength of the analysis. In contrast, this study builds a unit-level dataset of single-family housing, incorporating detailed locational, structural, and environmental attributes. The integration of high-resolution data with a spatially-aware statistical model enhances the explanatory power and predictive reliability, particularly in identifying localized high-risk zones.

Third, prior studies have rarely distinguished between urban and non-urban vacancy mechanisms, despite meaningful differences in demographic trends, land use intensity, and infrastructure accessibility. This study explicitly incorporates an urban–non-urban comparative framework, thereby revealing heterogeneous vacancy mechanisms across spatial contexts. This perspective not only enriches theoretical understanding but also supports the development of more context-sensitive housing policies.

In summary, the study makes three primary contributions: (1) It introduces a novel application of Spatial GAMLSS to vacancy risk prediction, advancing methodological approaches in housing vacancy research; (2) It leverages high-resolution, unit-level data for improved spatial inference; and (3) It incorporates an urban–non-urban comparative lens, enabling more nuanced policy insights for managing vacancy in diverse regional settings.

Ⅲ. Methodology

1. Spatial GAMLSS Framework

This section presents the analytical framework employed in this study. A sequential modeling strategy was developed to enhance both distributional flexibility and spatial sensitivity, culminating in a Spatial GAMLSS model with latent structured effects. The methodological procedure consists of five stages (M1–M5), which progressively integrate distributional complexity, nonlinear terms, and spatial random effects. Details of the distributional assumptions, model specification, and evaluation metrics are provided below. To estimate housing vacancy probabilities and identify spatial risk zones, this study adopts the Spatial GAMLSS framework–a flexible semi-parametric regression model capable of addressing the statistical complexities typically found in urban vacancy data. These complexities include distributional skewness, non-linearity, and spatial autocorrelation, which conventional regression approaches often fail to accommodate. Recent empirical applications underscore the strengths of GAMLSS and its spatial extension. Recent methodological overviews highlight that GAMLSS are particularly useful when data exhibit non-normal or bounded characteristics, as they allow modeling the entire distribution—including location, scale, and shape parameters—rather than only the mean. For example, Marmolejo-Ramos et al. (2022) demonstrated through applications in learning analytics how GAMLSS can flexibly fit skewed and bounded data, offering advantages over traditional regression models. Zhang et al. (2015) also utilized a GAMLSS-based model to predict China's annual maximum precipitation, successfully modeling the nonlinear effects of explanatory variables using a Cubic Spline function. These studies have demonstrated that GAMLSS is a powerful tool for analyzing data with non-normal distributions and strong skewness. The GAMLSS model is increasingly being utilized in the housing sector as well. Florencio et al. (2011) introduced GAMLSS for land price prediction, demonstrating superior predictive performance compared to traditional regression models. Similarly, Cajias (2018) reported that GAMLSS produced more accurate results than Generalized Additive Models (GAM) in rent prediction. However, most of these studies did not explicitly incorporate spatial elements, limiting their ability to fully account for the influence of location and surrounding areas. Spatial GAMLSS extends the traditional GAMLSS model by incorporating spatial autocorrelation, making it advantageous for explaining how housing vacancy probabilities can vary depending on location, even under identical housing conditions. For instance, De Bastiani et al. (2018) utilized a Spatial GAMLSS model to analyze real estate price data and effectively captured the spatial patterns of price distributions across regions. Spatial GAMLSS, which accounts for spatial autocorrelation, enables precise analysis of the impact of geographic location on housing vacancies. This makes it a valuable tool for visualizing and predicting the spatial distribution of vacancy probabilities.

The classical GAMLSS model, developed by Rigby and Stasinopoulos (2005), extends the capabilities of GLM and GAM by allowing the dependent variable to follow any parametric distribution from a broad family (e.g., Poisson, gamma, inverse Gaussian, zero-inflated beta). Unlike traditional models that generally focus on modeling only the mean parameter, GAMLSS enables simultaneous modeling of the location (μ), shape (σ), skewness (ν, τ) through separate link functions. This multi-parameter framework enhances the model's flexibility to capture heterogeneity in both central tendency and distributional form. In the context of this study, the μ and σ parameters are modeled as smooth additive functions of explanatory variables using penalized B-splines (P-splines), which allow for a flexible, data-driven representation of non-linear effects without imposing strict functional assumptions. Model selection is conducted based on penalized deviance, Akaike Information Criterion (AIC), and generalized pseudo R^2 to ensure optimal distributional and functional fit.

Crucially, the spatial extension of GAMLSS integrates spatial structure into the estimation process by applying an Intrinsic Conditional Autoregressive (ICAR) prior–equivalent to the Intrinsic Autoregressive (IAR) model–for the spatial random effect linked to the μ parameter. Following the approach by De Bastiani et al. (2018), this spatial formulation assumes that adjacent areal units exhibit correlated latent risks, thereby accounting for unmeasured spatial heterogeneity. This approach is especially relevant when physical, social, or policy factors influencing vacancy risks are spatially clustered but only partially observed in the dataset.

The resulting Spatial GAMLSS model not only improves the accuracy of local vacancy risk estimates but also provides a basis for visualizing the spatial diffusion of vacancy pressures across urban regions. Its ability to flexibly model non-normal distributions, capture non-linear relationships, and incorporate spatial structures makes it a suitable framework for analyzing the spatial patterns and underlying factors of housing vacancy.

In this study, the response variable is modeled as a continuous probability ranging between 0 and 1, representing the likelihood that a given single-family house is vacant. Accordingly, the μ parameter is directly interpreted as the estimated vacancy probability at each observational unit. Although the conceptual outcome is binary (i.e., vacant vs. non-vacant), a continuous probability formulation was adopted to facilitate spatial interpolation and to better capture uncertainty across units with varying characteristics. For the distributional assumption, several candidate distributions–including the Gaussian, gamma, and inverse Gaussian–were empirically evaluated. Based on model performance criteria such as the AIC and generalized R^2, the Exponential Gaussian distribution was selected as optimal. This distribution was particularly well-suited for the data's asymmetric structure, accommodating right-skewed probability values while ensuring stability in parameter estimation. In operationalizing the spatial random effects, the spatial adjacency structure was constructed based on Queen contiguity using administrative units (eup, myeon, and dong). The ICAR prior was applied to the μ parameter to account for latent spatial dependence among neighboring areas, enabling more accurate estimation of spatially correlated vacancy risks.

For binary response variables (0/1), the mean parameter (μ) in the GAMLSS framework corresponds to the predicted probability of success under the binomial distribution. Accordingly, μ represents the expected value of the outcome, conditional on the covariates. This interpretation is consistent with that in GLM, where μ directly reflects the probability of a "success" event, such as a housing unit being vacant. In this study, even though the vacancy status is observed as a binary outcome, a continuous probability formulation was adopted for μ to enable probabilistic interpretation, smooth spatial prediction, and incorporation into subsequent risk mapping.

2. Model Development and Specification

To accurately capture the spatial distribution and statistical characteristics of housing vacancy, this study constructed a five-stage modeling framework (M1–M5). Each stage incrementally expands the analytical scope, beginning with a simple linear regression model and progressively incorporating distributional flexibility, nonlinear covariate effects, multi-parameter estimation, and spatial autocorrelation structures.

In the first stage (M1), the linear regression model assumes a normally distributed response variable Y with mean μ and constant variance σ^2. The mean is expressed as a linear combination of the design matrix X and coefficient vector β, as shown in Equation (1):

$$Y \sim N(\mu, \sigma^2),\ \mu = X\beta \tag{1}$$

The second stage (M2) introduces the Generalized Linear Model (GLM), which extends the response distribution to the exponential family. A link function $g(\cdot)$ is applied to the mean μ, yielding the linear predictor $\eta = g(\mu)$. While this enhances distributional flexibility, the relationships between variables remain linear (Equation (2)):

$$Y \sim D(\mu, \varnothing),\ \eta = g(\mu) = X\beta \tag{2}$$

In the third stage (M3), the GAM is adopted to relax linearity assumptions. Each covariate is modeled with a nonparametric smoothing function $s_j(x_j)$, commonly implemented using P-splines. This enables the model to capture nonlinear covariate effects directly from the data (Equation (3)):

$$Y \sim D(\mu, \varnothing),\ \eta = g(\mu) = X\beta + \sum_{j=1}^{J} s_j(x_j) \tag{3}$$

The fourth stage (M4) adopts the generalized additive models for location, scale, and shape (GAMLSS) framework. Unlike GLM and GAM, which typically model only the mean, GAMLSS enables separate modeling of multiple distribution parameters–mean μ, variance σ, and skewness ν–each with its own link function and smoothing structure. In this study, the exponential Gaussian distribution is used, with identity, log, and log link functions applied to μ, σ, and ν, respectively (Equation (4)):

$$\begin{aligned} &Y \sim D(\mu, \sigma, \upsilon) \\ &\eta_1 = g_1(\mu) = X_1\beta + \sum_j s_{1j}(x_{1j}) \\ &\eta_2 = g_2(\mu) = X_2\beta + \sum_j s_{2j}(x_{2j}) \\ &\eta_3 = g_3(\mu) = X_3\beta + \sum_j s_{3j}(x_{3j}) \end{aligned} \tag{4}$$

In the final stage (M5), spatial effects are integrated into the GAMLSS framework by incorporating a spatial random effect into the mean predictor μ. This accounts for potential spatial autocorrelation among adjacent administrative units. The spatial effect follows an Intrinsic Conditional Autoregressive (ICAR) structure based on a queen contiguity matrix constructed from municipal boundary data. This enables the model to reflect similarities among geographically proximate areas. The spatial random effect model is given by Equation (5):

$$y = Z_{\gamma} + \epsilon \tag{5}$$

The spatial random effect γ is assumed to follow an ICAR prior, and its estimates are computed using the following formulation, where $\lambda = \sigma_{\epsilon}^2/\sigma_{\gamma}^2$ represents the ratio of error variance to spatial signal (Equation (6)):

$$\hat{\gamma} = (ZWZ^T + \lambda G)^{-1} Z^T W_y \tag{6}$$

All models were fitted using the RS algorithm, and model performance was assessed using generalized R^2 and the *AIC*. The generalized R^2 is defined Equation (7):

$$R^2 = 1 - \left(\frac{L(0)}{L(fitted)}\right)^{\frac{2}{n}} \tag{7}$$

Where $L(0)$ denotes the log-likelihood of the null model (intercept only), and L(fitted) represents the log-likelihood of the fitted model, and n is the number of observations. Higher values of generalized R^2 indicate better model performance. The *AIC* is expressed Equation (8):

$$GAIC = -2\hat{l} + k \cdot df \tag{8}$$

Where, $\hat{l}$ is the fitted log-likelihood, df is the effective degrees of freedom, and k is the penalty parameter, set to 2 in the case of *AIC*.

The response variable (probability of being vacant) follows an Exponential Gaussian distribution with parameters μ, σ, and ν. The link function for the parameters was set as the identity function for mean (μ) and the log function for variance (σ) and skewness (ν). This approach allows each parameter of the model to flexibly vary according to the explanatory variables. The final model can be expressed Equation (9).

$$\begin{aligned} \mu &= \beta_0 + pb(F\_a) + \beta_1(R\_s) + pb(B\_a) + pb(O\_s) + \\ &\quad \beta_2(Dist\_from\_admin\_bound) + \beta_3(Dist\_from\_ \\ &\quad w\_bound) + \beta_4(Dist\_from\_env\_pol\_emi) \\ \log(\sigma) &= \gamma_0 + \gamma_1(F\_ar) + \gamma_2(Dist\_from\_w\_bound) \\ \log(\nu) &= \delta_0 + \delta_1(B\_a) + \delta_2(Dist\_from\_w\_bound) \end{aligned} \tag{9}$$

All models were estimated using the RS (residual scoring) algorithm via the GAMLSS package in R. The spatial contiguity matrix was implemented using the spdep package. Model performance was assessed using the Akaike Information Criterion (*AIC*) and generalized R^2, and the detailed model comparison results are presented in Section V.

Ⅳ. Research Design

1. Data and Analytical Procedures

1) Study Area

The empirical focus of this study is Jinju-si, located in Gyeongsangnam-do, Republic of Korea. As of 2021, Jinju had a population of approximately 341,000, classifying it as a medium-sized city. According to the 2021 National Vacancy Survey conducted by the Ministry of Land, Infrastructure and Transport (MOLIT), the total number of vacant housing units in Jinju was 3,370, of which over 60% were single-family dwellings.

Furthermore, data from the Korea Real Estate Board's Vacancy Information System indicate that Jinju's housing vacancy rate in 2020 stood at 10.1%, which exceeds both the national average (8.2%) and the Gyeongsangnam-do provincial average (11.6%), ranking second highest in the province after Changwon-si. These figures underscore Jinju's relatively severe vacancy problem compared to other mixed urban–rural medium-sized cities, supporting the relevance of this case for drawing policy implications applicable to similar urban typologies.

Administratively, Jinju-si is composed of 49 subdivisions, including 15 myeon, 33 dong, and one eup. Urban areas account for approximately 38.68% of Jinju's total land area, while non-urban regions comprise the remaining 61.31%. A spatial visualization of the study area is provided in 〈Figure 1〉. Jinju-si is characterized by pronounced spatial heterogene-

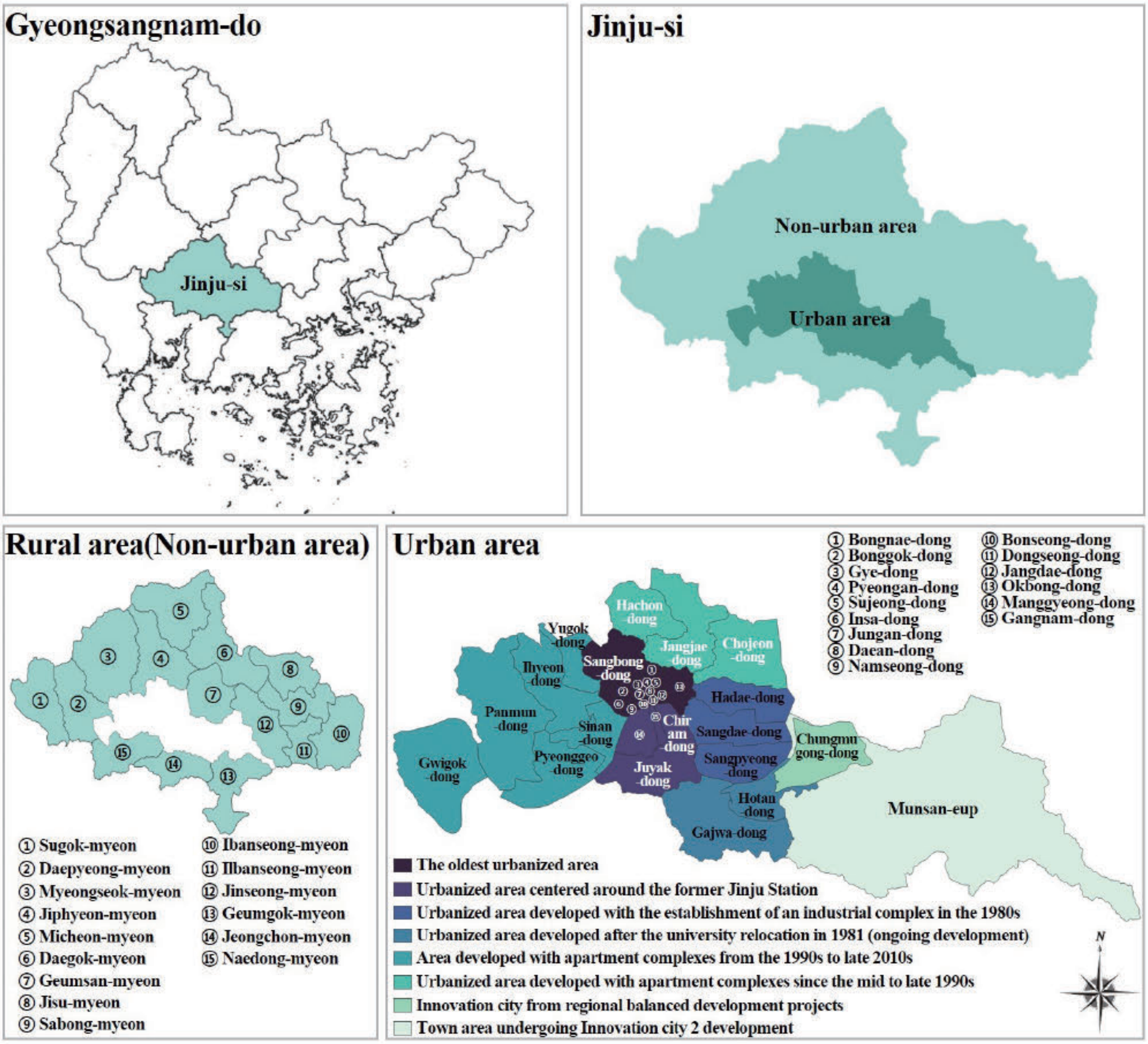


Figure 1. Study area

ity, with the coexistence of older and newer urban districts. Continued peripheral expansion has resulted in the functional decline of the historic city center and the erosion of its commercial centrality—processes that have contributed directly to the rise in vacant housing. This spatial duality makes Jinju particularly well suited for analyzing vacancy dynamics within differentiated urban structures.

In addition to its spatial complexity, Jinju has institutional plans to conduct systematic vacancy monitoring in cooperation with the Korea Real Estate Board, beginning in 2024. Under this initiative, a quinquennial vacancy census will be implemented, including field surveys and classification by structural condition. The resulting dataset will be integrated with a national vacancy information platform, offering high-resolution, address-level spatial data suitable for advanced empirical analysis. These developments enhance Jinju's suitability as a case study for vacancy research requiring fine-grained spatial data.

Moreover, Jinju has been designated as a strategic policy focus area, with various urban regeneration initiatives—including the Urban Regeneration New Deal—actively underway. These ongoing programs further increase the practical relevance of the study's findings for policy application and implementation. Taken together, Jinju's spatial complexity, administrative structure, data accessibility, and policy significance render it a highly appropriate case for this research. The analytical results derived from Jinju are expected to offer generalizable insights for other mixed urban–rural medium-sized cities facing similar vacancy-related challenges.

2) Data

This study estimates the probability of housing vacancies in single-family homes across urban and non-urban areas of Jinju-si. As of 2021, Jinju-si's residential housing stock was recorded at 42,169 units, comprising 25,860 units in urban areas and 16,309 units in non-urban areas. Excluding multi-family housing, multi-household dwellings, official residences, and other types of single-family houses, the number of single-family houses was recorded at 36,605 units, with 20,629 units in urban areas and 15,976 units in non-urban areas, as shown in 〈Figure 2〉 (Statistics Korea, 2024).

The vacancy dataset employed in this study was constructed from address-level administrative records generated through systematic field inspections conducted by authorized municipal personnel in Jinju-si between 2017 and 2020. These on-site surveys identified persistently unoccupied single-family residential units over a multi-year period. The resulting dataset, comprising formally verified and spatially detailed observations, provides a robust empirical foundation for unit-level probabilistic modeling of housing vacancy risk. As of 2020, a total of 2,018 vacant single-family houses had been officially recorded in Jinju-si. The descriptive statistics of this dataset are presented in 〈Table 1〉. To assess whether physical and spatial characteristics differ significantly by vacancy status and location, an independent samples t-test was performed across urban and non-urban contexts. The results indicate that most variables exhibit statistically significant differences between vacant and non-vacant houses, suggesting that vacancy patterns are systematically associated with distinct physical and locational conditions rather than occurring randomly within the residential landscape.

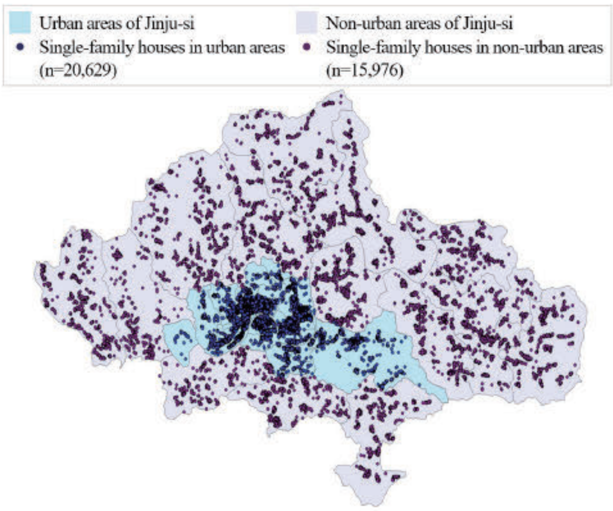


Figure 2. Distribution of single-family houses in Jinju-si

As shown in 〈Table 1〉, the characteristics of single-family houses in Jinju-si exhibit notable variations based on urbanity and vacancy status. In urban areas, non-vacant single-family houses have an average floor area of approximately 91.57 m^2 and an average building age of 46 years. The mean distances to the nearest administrative boundary, water boundary, and environmental pollutant emission facility are 385 m, 542 m, and 3,520 m, respectively.

In contrast, vacant single-family houses in urban areas are generally smaller and older, with an average floor area of 73 m^2 and a mean building age of 53 years. These units are also located slightly closer to environmental and administrative edges, with average distances of 349 m to administrative boundaries, 521 m to water boundaries, and 3,855 m to pollutant-emitting facilities.

In non-urban areas, non-vacant houses display an average floor area of approximately 91 m^2 and an average age of 48 years. Their mean distances to the nearest administrative boundary, water boundary, and pollutant emission facility are 1,229 m, 365 m, and 9,391 m, respectively. By comparison, vacant single-family houses in non-urban settings tend to be smaller (76 m^2), significantly older (63 years), and located closer to potential externalities—on average, 1,050 m from administrative boundaries, 274 m from water boundaries, and 7,924 m from pollutant sources.

These patterns suggest that while the average floor area of single-family houses does not differ substantially between urban and non-urban areas, the urban housing stock tends to be relatively newer. Moreover, across both spatial contexts, vacant houses are consistently characterized by smaller size, older construction, and closer proximity to administrative edges and environmental burdens—highlighting a spatial marginality associated with residential vacancy.

Finally, 〈Figure 3〉 illustrates the spatial distribution of vacant and non-vacant single-family houses across urban and non-urban areas in Jinju-si. Areas with no observed data correspond primarily to commercial zones or mountainous regions where residential housing is sparse or absent.

Table 1. Descriptive statistics

Variables	Classification		Mean	Median	p-value
Floor area (m^2)	Urban	Non-vacant houses	91.57	74.09	0.000***
		Vacant houses	73.15	54.98	
	Non-urban	Non-vacant houses	91.42	84.95	0.000***
		Vacant houses	75.79	69.10	
Building age (year)	Urban	Non-vacant houses	46.27	45.00	0.000***
		Vacant houses	52.59	53.50	
	Non-urban	Non-vacant houses	47.78	43.00	0.000***
		Vacant houses	62.88	69.00	
Distance from administrative boundary (m)	Urban	Non-vacant houses	385.34	296.07	0.000***
		Vacant houses	348.86	277.99	
	Non-urban	Non-vacant houses	1229.07	1124.15	0.000***
		Vacant houses	1049.70	911.50	
Distance from water boundary (m)	Urban	Non-vacant houses	541.83	486.56	0.079*
		Vacant houses	520.85	474.94	
	Non-urban	Non-vacant houses	364.65	249.99	0.000***
		Vacant houses	273.91	200.32	
Distance from environmental pollutant emission facility (m)	Urban	Non-vacant houses	3519.88	3226.06	0.000***
		Vacant houses	3855.04	3944.36	
	Non-urban	Non-vacant houses	9391.27	9452.02	0.000***
		Vacant houses	7923.55	6600.17	

Note) Non-vacant houses N=34,587, Vacant houses N=2,018
*p<0.1, **p<0.05, ***p<0.01

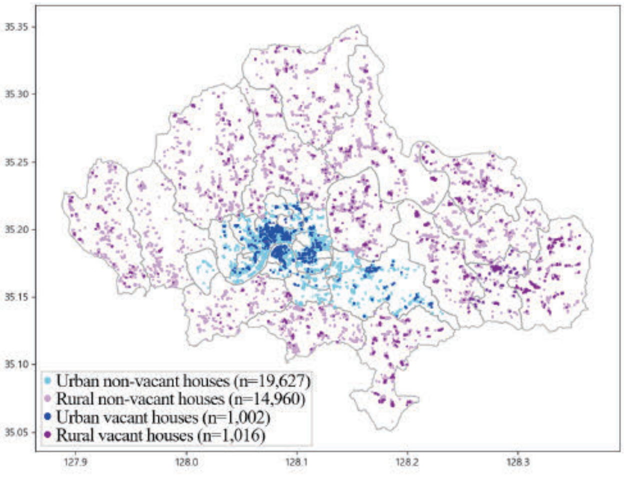


Figure 3. Distribution of vacant and non-vacant houses

2. Explanatory Variables

This study employs seven explanatory variables to estimate the probability of housing vacancy: Floor area, Roof structure, Building age, Owner-occupancy status (dummy variable), Distance from administrative boundaries, Distance from water boundaries, and Distance from environmental pollutant emission facilities. These variables were selected based on prior empirical findings and their theoretical relevance to the structural and spatial dynamics of Jinju-si.

Floor area serves as a proxy for the physical scale and functional capacity of the housing unit. Smaller houses tend to exhibit lower market appeal, are more susceptible to functional obsolescence, and are often bypassed in reinvestment decisions, especially in peripheral or aging neighborhoods. Prior research has consistently shown that reduced floor area is positively associated with higher vacancy rates (Clark and Herrin, 1997; Lester et al., 2013).

Roof structure reflects the structural durability and construction quality of the house. Houses with outdated or vulnerable roofing materials (e.g., wood, panels) are more likely to suffer from leakage, fire hazards, or thermal inefficiency. Such homes are frequently excluded from reinvestment cycles and may face prolonged vacancy, particularly in declining neighborhoods (Liu et al., 2024; Yue et al., 2022).

Building age captures the degree of physical deterioration and structural obsolescence. Older buildings often require greater maintenance, may not comply with contemporary building standards, and are more likely to remain vacant, especially in areas with declining population and weak housing demand (Newman, 2013; Lee et al., 2022).

Owner-occupancy status is a key socio-economic indicator reflecting the likelihood of housing management and residential stability. Owner-occupied homes tend to be better maintained and less prone to vacancy, whereas absentee ownership is often associated with disinvestment and higher vacancy risk (Martin et al., 2017; Molloy, 2016).

Distance from administrative boundary measures spatial marginality. Properties located near jurisdictional borders frequently suffer from reduced access to services, limited infrastructure investment, and lower integration into planning frameworks. These “edge effects” have been associated with housing instability and elevated vacancy rates (Henry et al., 2001; Geoghegan et al., 1997).

Distance from water boundary captures the influence of natural landscape features. While waterfront proximity can enhance property values in some metropolitan areas, in many secondary cities like Jinju-si, water-adjacent lands are often underdeveloped, flood-prone, or designated for non-residential use, thus becoming a locational disadvantage (Newman et al., 2016; Yue et al., 2022).

Distance from environmental pollutant emission facilities accounts for the negative externalities of industrial land uses. Residential properties near pollution-emitting facilities (e.g., factories, waste plants) face health and environmental concerns, often leading to reduced desirability and increased vacancy risks. In Jinju-si, these facilities are typically located in industrial zones with adjacent residential areas showing higher vacancy rates (Cozens and Tarca, 2016; Joo and Lee, 2021).

All spatial variables were operationalized using high-resolution GIS data from national sources such as the National Geographic Information Institute and the Ministry of the Interior and Safety. Euclidean distance metrics and buffer zones were applied to quantify spatial exposures.

3. Analytical Procedure

This study estimates the probability of housing vacancies in single-family houses in Jinju-si and predicts future high-risk vacancy zones through the following four steps.

First, based on existing research and data analysis, potential variables that may influence housing vacancies were identified. The correlations between these variables were examined, and candidate variables for analysis were selected. Through this, the aim was to enhance the accuracy of vacancy probability analysis by considering regional characteristics and the spatial distribution of housing vacancies.

Second, the Spatial GAMLSS model was used to estimate the probability of housing vacancies and analyze the influence of each variable. In this process, spatial autocorrelation and nonlinear relationships were incorporated, allowing for the estimation of housing vacancy probabilities in both urban and non-urban areas. This approach enabled the quantitative assessment of spatial factors influencing housing vacancies in specific areas.

Third, based on the Spatial GAMLSS model, areas with high housing vacancy probabilities were visualized. Specifically, spatial clustering techniques were employed to identify high-risk areas for housing vacancies (hotspots) and stable areas (cold spots), enabling the prediction of future vacancy-prone regions.

Fourth, by comparing high-risk housing vacancy areas classified into urban and non-urban regions, the study identified priority areas for policy intervention. This allowed for the identification of key causes of housing vacancies and the development of region-specific management strategies. As a result, effective policy directions were proposed for the efficient allocation of limited resources and the resolution of housing vacancy issues.

Through the analytical procedures outlined in ⟨Figure 4⟩, this study aims to provide a comprehensive understanding of the probability and location of housing vacancies in single-family houses in Jinju-si, and to establish a practical policy foundation for addressing the housing vacancy issue.

Ⅴ. Analytical Results

1. Vacancy Rate Estimation

1) Model Calibration and Results

To estimate the probability of housing vacancies in Jinju-si, this study employed a sequential modeling strategy designed to incrementally enhance distributional flexibility

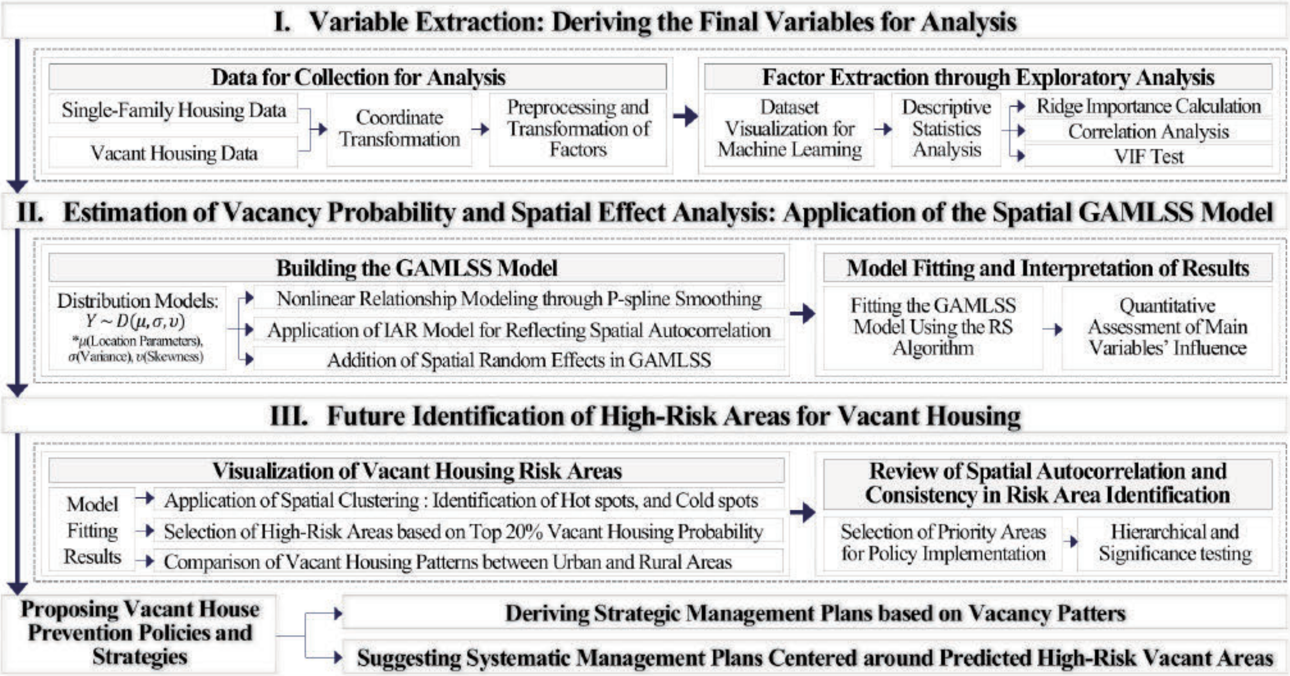


Figure 4. Research flowchart

and spatial sensitivity. The modeling process began with a baseline Gaussian model (M1) and advanced through successive stages incorporating non-Gaussian distributions (M2), non-linear smoothing terms (M3), and multi-parameter distributional structures under the GAMLSS framework (M4). Although each intermediate model showed progressive improvements in fit, detailed outputs are omitted here for analytical clarity and to emphasize the final model of interest.

The final specification, Model 5 (M5), extends the GAMLSS structure by incorporating spatial random effects, which account for latent spatial dependence across housing units. Among all specifications, M5 achieved the best performance, recording the lowest AIC values—99,502 for urban areas and 96,902 for non-urban areas—as well as the highest generalized R^2 values of 0.81 and 0.78, respectively. As shown in 〈Figure 5〉, residual diagnostics revealed no discernible patterns or spatial autocorrelation, affirming the model's robustness and internal consistency.

As presented in 〈Table 2〉, the generalized R^2 increased and *AIC* decreased as the models progressed from M1 to M5 in both urban and non-urban contexts. This pattern indicates

Table 2. Model selection and criteria

Model name			Response variable distribution	Generalized R^2	*AIC*
Urban area	M1	Linear model	Gaussian	0.16	231,143
	M2	Generalized linear model	Exponential Gaussian	0.54	230,544
	M3	Generalized additive model	Exponential Gaussian	0.63	230,450
	M4	Additive GAMLSS model (μ, σ, ν)	Exponential Gaussian	0.74	162,578
	M5	Spatial GAMLSS	Exponential Gaussian	0.81	99,502
Non-urban area	M1	Linear model	Gaussian	0.29	113,909
	M2	Generalized linear model	Exponential Gaussian	0.56	110,808
	M3	Generalized additive model	Exponential Gaussian	0.62	107,556
	M4	Additive GAMLSS model (μ, σ, ν)	Exponential Gaussian	0.68	103,181
	M5	Spatial GAMLSS	Exponential Gaussian	0.78	96,902

steady improvement in model fit, which is further confirmed by the visual diagnostics in ⟨Figure 5⟩. Both M5 models showed minimal deviation from observed values and no systematic residual bias, justifying the model's adoption for the final analysis.

2) Linear Effects

This study simultaneously modeled the spatial structure of the data along with mean (μ), variance (σ), and skewness (v) to analyze housing vacancy probabilities. The parameter estimates derived from the final model are presented in ⟨Table 3⟩ for both urban and non-urban areas.

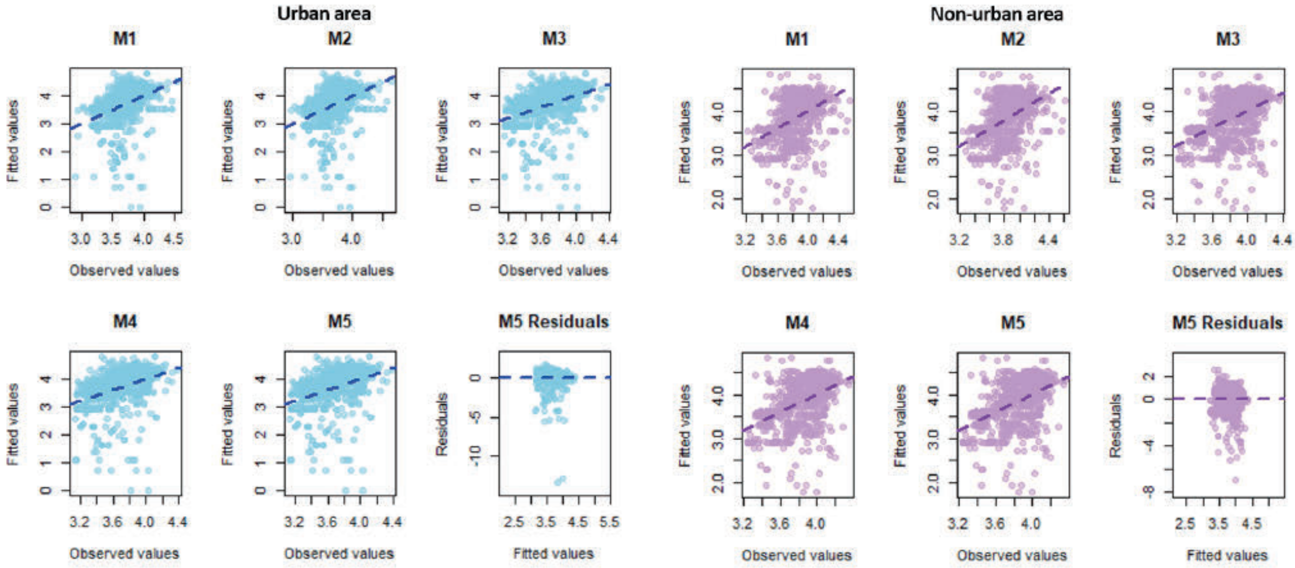


Figure 5. The goodness of the fit from M1 to M5

Note) L: Urban area, R: Non-urban area

Table 3. Results from the final model

Variables	Urban area				Non-urban area (Rural)			
	Estimate	Std. E.	t-value	p-value	Estimate	Std. E.	t-value	p-value
μ coefficients								
Intercept	2.935	0.169	6.600	0.000	4.591	0.266	4.869	0.000
pb (Floor area)	-0.045	0.018	-2.421	0.000	-0.141	0.029	-4.847	0.000
Roof structure	-0.008	0.010	-0.770	0.441	0.016	0.016	0.995	0.020
pb (Building age)	0.043	0.013	3.376	0.000	0.067	0.019	3.505	0.000
Owner-occupancy status	0.018	0.017	0.990	0.322	-0.038	0.027	-1.398	0.062
pb (Distance from administrative boundaries)	-0.061	0.012	-5.290	0.000	0.078	0.018	4.294	0.000
pb (Distance from water boundaries)	0.072	0.009	0.786	0.002	-0.068	0.013	-5.106	0.000
pb (Distance from environmental pollutant emission facility)	-0.244	0.011	-2.062	0.000	-0.265	0.017	-5.920	0.000
σ coefficients								
Intercept	1.363	0.093	1.116	0.000	-2.170	0.139	7.870	0.000
Floor area	-0.045	0.018	-2.520	0.012	-0.076	0.027	-2.770	0.006
Distance from water boundaries	0.038	0.009	4.040	0.000	-0.050	0.013	-3.740	0.000
v coefficients								
Intercept	1.055	0.068	4.882	0.000	2.007	0.097	3.332	0.000
Building age	0.030	0.012	2.480	0.003	0.045	0.018	2.510	0.012
Distance from water boundaries	-0.038	0.009	4.070	0.000	0.081	0.033	-5.731	0.000

The mean parameter (μ) was modeled using seven key explanatory variables influencing housing vacancy probability. These variables include floor area, building age, owner-occupancy status, roof structure, distance from administrative boundaries, distance from water boundaries, and distance from the nearest environmental pollutant emission facility. P-spline smoothing was applied to floor area, building age, and owner-occupancy status to capture the nonlinear relationships between these explanatory variables and the response variable. The variance parameter (σ) was modeled based on two variables—floor area and distance to the nearest environmental pollutant emission facility—to relax the assumption of constant variance in the distribution. This approach accounts for the possibility that the variance of housing vacancy probabilities may vary depending on specific variables. Finally, the skewness parameter (v) was modeled using building age and the distance to the nearest environmental pollutant emission facility to account for the asymmetry of the distribution. This design allows the skewness of the distribution to flexibly adapt to the characteristics of the data. In conclusion, the final model in this study explained housing vacancy probabilities by comprehensively considering mean, variance, and skewness. Notably, the use of P-spline smoothing captured nonlinear relationships, while quantifying spatial characteristics and interactions among variables enabled a more accurate analysis of the complex patterns of housing vacancies.

The analysis results showed the following outcomes for the intercept of the mean parameter (μ) derived from the final model. In urban areas, the intercept was estimated at 2.935, representing the baseline housing vacancy probability in urban areas when controlling for key variables related to housing vacancies. In non-urban areas, the intercept was estimated at 4.591, indicating that the baseline housing vacancy probability in non-urban areas is relatively higher than that in urban areas. The P-spline coefficients for floor area were -0.045 in urban areas and -0.141 in non-urban areas. This indicates that as floor area increases, the probability of housing vacancy tends to decrease, suggesting that larger houses may be more attractive or have higher utility. The larger decrease observed in non-urban areas suggests that floor area has a more significant impact on the probability of vacancy in non-urban areas. Roof structure did not have a significant impact in urban areas; however, in non-urban areas, it showed a positive coefficient of 0.016, indicating a slight positive effect on housing vacancy probability. This suggests that certain roof structures in non-urban areas may increase the likelihood of housing vacancies. The P-spline coefficients for building age were 0.043 in urban areas and 0.067 in non-urban areas, indicating that the probability of housing vacancy increases as building age increases. For owner-occupancy status, no significant results were observed in urban areas. In contrast, in non-urban areas, the coefficient was -0.038, indicating that the likelihood of housing vacancy decreases as the probability of the owner residing in the property increases. The distance from administrative boundaries showed coefficients of -0.061 in urban areas and 0.078 in non-urban areas. This indicates that in urban areas, the probability of housing vacancy decreases as the distance from administrative boundaries increases. Conversely, in non-urban areas, the probability of housing vacancy increases with greater distance from administrative boundaries. The distance from water boundaries had coefficients of 0.072 in urban areas and -0.068 in non-urban areas. This indicates that in urban areas, the probability of housing vacancy decreases for houses closer to water boundaries, whereas in non-urban areas, the opposite trend was observed. The distance from environmental pollutant emission facilities showed a negative relationship in all areas, with coefficients of -0.244 in urban areas and -0.265 in non-urban areas. This suggests that proximity to environmental pollutant emission facilities increases the likelihood of housing vacancy.

The analysis of the variance parameter (σ) revealed that floor area and distance from water boundaries are significant variables influencing the variance of housing vacancy probabilities. For floor area, the coefficients were -0.045 in urban areas and -0.076 in non-urban areas, indicating that larger floor areas tend to reduce the variance in housing vacancy probabilities. This reflects the characteristic that larger houses are more likely to be utilized stably. The distance from water boundaries exhibited contrasting patterns across regions, with a coefficient of 0.038 in urban areas and -0.050 in non-urban areas. In urban areas, proximity to water boundaries increased the variability in housing vacancy probabilities, likely due to differences in residential preferences. In contrast, in non-urban areas, proximity to water boundaries reduced variability, suggesting that houses near water boundaries are often utilized as agricultural or non-residential facilities, leading to more stable usage patterns.

The analysis of the skewness parameter (ν) identified building age and distance from water boundaries as significant variables influencing the asymmetry of housing vacancy probabilities. Building age exhibited a positive relationship across all areas, with coefficients of 0.030 in urban areas and 0.045 in non-urban areas, indicating that older buildings are a major factor increasing the asymmetry of housing vacancy probabilities. This suggests that aged houses are more likely to transition rapidly into vacancies under certain conditions. In urban areas, proximity to water boundaries decreased the asymmetry of housing vacancy probabilities, reflecting the higher utility or relatively stable vacancy likelihood of houses in urban waterfront locations. In contrast, in non-urban areas, proximity to water boundaries increased the asymmetry, indicating that certain houses in non-urban waterfront regions are more likely to transition rapidly into vacancies under specific conditions.

3) Nonlinear Effects

In this study, P-spline smoothers were applied to variables such as floor area, building age, distance from administrative boundaries, distance from water boundaries, and distance from environmental pollutant emission facilities to explain the mean parameter μ. P-spline smoothers were used to capture the nonlinear relationships between the vacancy rate (housing vacancy probability) and the explanatory variables, enabling more precise modeling of these relationships. The degrees of freedom for the smoothing terms were used as a measure to indicate the extent of nonlinearity in the relationship between the explanatory variables and the response variable.

The analysis results showed that the estimated degrees of freedom for key variables such as floor area, building age, distance from administrative boundaries, distance from water boundaries, and distance from environmental pollutant emission facilities were all sufficiently high. A high degree of freedom indicates a nonlinear relationship between the variable and the probability of housing vacancy. Conversely, if the degree of freedom is close to 1.0, it suggests a linear relationship between the variable and the response variable, implying that the variable's effect can be adequately explained using a linear model without the need for smoothing terms. In this study, the degrees of freedom for all five key variables were sufficiently high, confirming the necessity of using P-spline smoothing terms to account for nonlinear relationships in modeling the mean parameter μ. This approach effectively captured the complex relationships between the variables and the response variable, enabling more accurate analysis results for housing vacancy probabilities.

〈Figure 6〉 visually illustrates the relationships between the vacancy rate and the seven key explanatory variables used to describe the mean parameter μ. Notably, the results focus on the five smoothing terms where nonlinear effects are pronounced. First, regarding the relationship between floor area and vacancy rate, urban areas exhibited a trend where the vacancy rate increased as floor area grew from approximately 0 m^2 to 200 m^2. However, beyond 200 m^2, the vacancy rate sharply declined, reflecting a common charac-

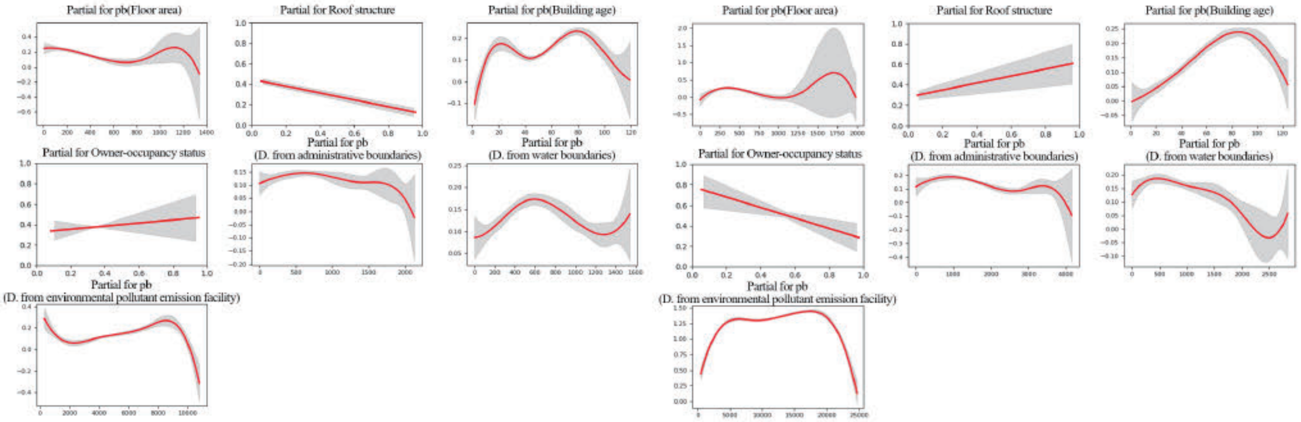


Figure 6. Relationships between the probability of being vacant and explanatory variables

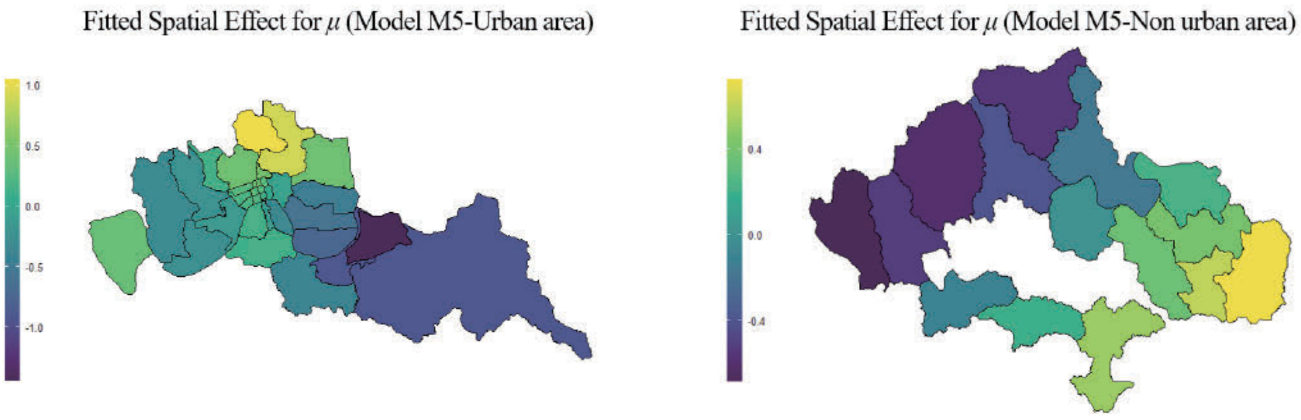
Note 1) L: Urban, R: Non-urban
Note 2) The shaded areas are the 95% confidence bands.

teristic of the real estate market where large single-family houses are considered premium properties with reduced vacancy risks. In contrast, in non-urban areas, the vacancy rate gradually decreased as floor area increased, likely because larger floor areas were perceived as more attractive housing options. The relationship between building age and vacancy rate also revealed notable differences between urban and non-urban areas. In urban areas, the vacancy rate initially increased with older building age but gradually declined after a certain point, reflecting the stabilization of older properties as they secure a fixed position in the housing market. In non-urban areas, however, the vacancy rate continued to rise as building age increased, indicating that aging properties are a significant factor contributing to housing vacancies. The distance from administrative boundaries had opposing effects in urban and non-urban areas. In urban areas, the vacancy rate gradually decreased as the distance from administrative boundaries increased, reflecting a preference for housing located closer to central areas. In contrast, in non-urban areas, the vacancy rate tended to increase with greater distance from administrative boundaries, indicating reduced residential appeal in peripheral locations. The distance from water boundaries also exhibited regional differences. In urban areas, regions closer to water boundaries tended to have lower vacancy rates, likely due to the pleasant environment and scenic advantages. In contrast, in non-urban areas, proximity to water boundaries was associated with higher vacancy rates, which could be attributed to reduced residential appeal caused by flood risks or natural disasters. The distance from environmental pollutant emission facilities showed similar trends in both urban and non-urban areas. Closer proximity to these facilities resulted in higher vacancy rates, likely due to pollution and noise diminishing residential desirability. However, beyond a certain distance from the facilities, these negative effects diminished, leading to lower vacancy rates. Meanwhile, variables such as roof structure and owner-occupancy status showed differing impacts on vacancy rates between urban and non-urban areas. Notably, in non-urban areas, owner-occupancy status exhibited a significant relationship with reduced vacancy rates, whereas this trend was not observed in urban areas.

The remaining panels in ⟨Figure 6⟩ depict the linear relationships between vacancy rates and explanatory variables. These relationships do not exhibit pronounced nonlinear effects, indicating that changes in the explanatory variables influence vacancy rates in a linear manner.

4) Spatial Effects on μ

In this study, spatial effects were incorporated into the final model (M5) to estimate the mean (μ) of the response variable representing the probability of housing vacancy. The analysis results estimated the pure error effect (σe)at 32.38, which reflects the overall variability not explained by the model. Additionally, the spatial random effect (σb) was estimated at 1.07, reflecting variability attributable to inter-regional differences. To provide a more intuitive interpretation of the spatial effects, the analysis visualized the spatial effects rather than examining the variance values of the random effects. The spatial effects of μ derived from the final model are visually represented in ⟨Figure 7⟩. The analysis revealed that areas within urban regions of Jinju-si with


Fitted Spatial Effect for μ (Model M5-Urban area)

Fitted Spatial Effect for μ (Model M5-Non urban area)

Figure 7. Fitted spatial effect for μ for model M5

Note) L: Urban, R: Non-urban

insufficient commercial facilities and transportation infrastructure exhibited relatively low residential desirability, resulting in a higher risk of housing vacancies. In contrast, new urban developments and areas near transit hubs were found to have a lower risk of housing vacancies, with relatively low μ values observed in these regions. This indicates high residential desirability, stable housing demand, and a reduced likelihood of vacancies in these areas. These spatial patterns highlight the variations in housing demand and residential environments across different areas within Jinju-si, underscoring the need for region-specific vacancy management policies.

The GAMLSS model offers the advantage of flexibly modeling not only the mean (μ) but also all distribution parameters (e.g., variance, skewness) based on explanatory variables. In this study, the visualization of spatial effects on μ revealed that there remain unexplained spatial distributions of housing vacancy probabilities. This is likely due to the model's inability to include certain significant spatial factors as explanatory variables. These residual spatial effects were effectively accounted for through the Spatial GAMLSS model, highlighting the critical role of spatially-informed modeling in analyzing housing vacancy patterns.

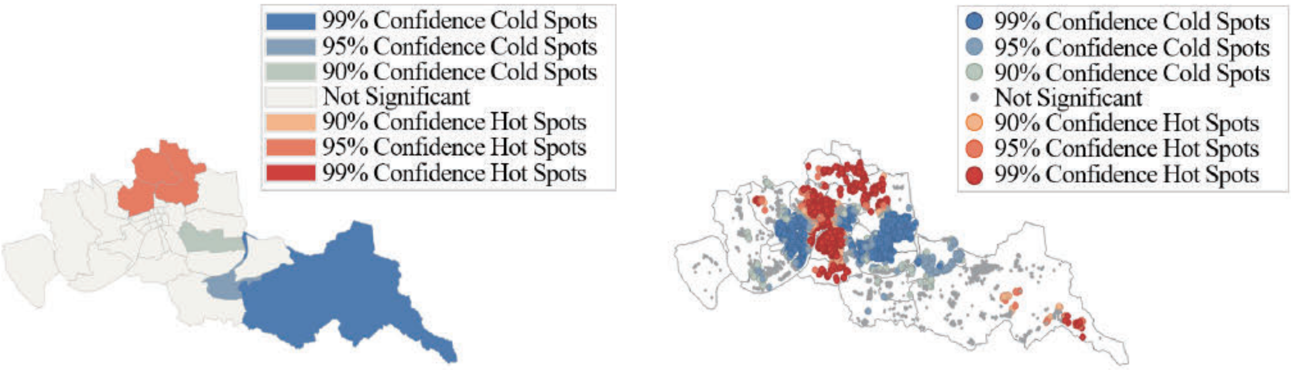
2. Prediction of Future Vacant Houses

This study conducted a hotspot analysis in Jinju-si to identify areas and buildings at high risk of housing vacancy in urban and non-urban regions. The hotspot analysis was based on housing vacancy probabilities estimated using the Spatial GAMLSS model, and the results were derived for urban areas, non-urban areas, and the entire region of Jinju-si.

Two types of hotspot analyses were conducted to reflect both meso- and micro-level spatial patterns. First, the spatially estimated vacancy probabilities (μ) were aggregated to the eup–myeon–dong level, and the average values were used to identify area-based risk clusters using the Getis-Ord Gi* statistic (Figures 8(a) and 9(a)). Second, the same μ values at the individual housing unit level were used as input for point-based hotspot analysis (Figures 8(b) and 9(b)), allowing the detection of statistically significant clusters of high-risk properties. This dual-scale analysis facilitates a more comprehensive interpretation of vacancy risk across both regional and building levels.

〈Figure 8(a)〉 illustrates the areas at risk of housing vacancy in the urban regions of Jinju-si. The risk of housing vacancy was low in new urban developments and areas near transit hubs, whereas it was high in areas with aging residential and infrastructure facilities, as well as gateway regions on the urban outskirts. This suggests that housing vacancy risk increases in areas with lower residential desirability and declining housing demand. In particular, traditional old city centers and early-developed residential areas exhibited a higher risk of housing vacancies. In contrast, newly developed urban districts, areas urbanized in connection with industrial complexes, and regions with active educational and economic activities showed a lower likelihood of housing vacancies. These findings indicate that residential preferences and the distribution of economic activities within the city significantly influence the risk of housing vacancies.

〈Figure 8(b)〉 visualizes buildings in Jinju-si's urban areas with a high risk of housing vacancy. High-risk buildings were predominantly located in areas with aging residential environments or on the urban outskirts. These areas often



(a) High-risk areas for housing vacancy

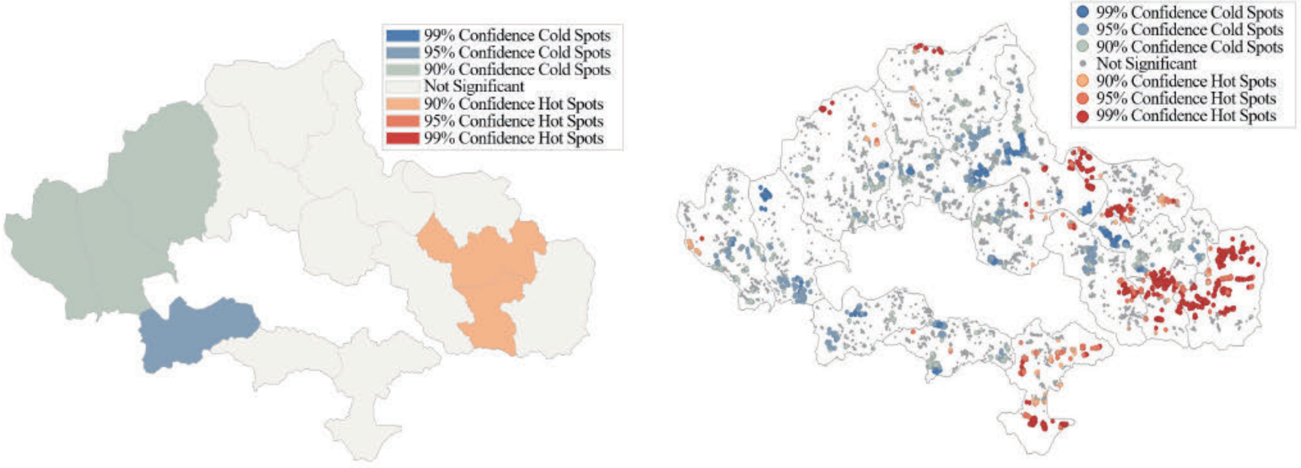
(b) Houses with high vacancy rates

Figure 8. Vacancy risk in urban areas based on hotspot analysis

exhibited low residential desirability due to inadequate transportation infrastructure or limited accessibility to the city center. Additionally, regions with declining old commercial districts or mixed urban-rural residential environments showed relatively higher vacancy risks. In contrast, buildings with a low risk of vacancy were identified in areas where recent development has been active or where urban infrastructure is well-established. These areas exhibited high residential desirability and low vacancy rates due to the stable maintenance of residential and commercial functions, as well as well-developed transportation and industrial infrastructure. Notably, regions undergoing new urban development or industrial complex construction were found to have stable housing demand and a low risk of vacancy. Summarizing the results of vacancy risk predictions in the urban areas of Jinju-si, it was found that areas with aging residential environments and poor transportation accessibility had a high risk of housing vacancies. In contrast, regions characterized by active urbanization and economic activities exhibited a lower likelihood of vacancy occurrence. These findings suggest that factors such as the level of urbanization, economic activity, and residential desirability play a significant role in influencing housing vacancy risks.

〈Figure 9(a)〉 illustrates the high-risk areas for housing vacancy in the non-urban regions of Jinju-si. The analysis revealed that vacancy risk was elevated in industrial areas with a manufacturing base and rural areas with low accessibility to urban centers. These regions shared several common economic and social characteristics. First, areas with limited transportation accessibility exhibited a pronounced risk of housing vacancy. Regions lacking connectivity to major transportation networks or located far from urban centers exhibited reduced residential desirability due to limited mobility for residents and lower accessibility to local commercial and living infrastructure. In such areas, the loss of functionality in former transportation hubs, caused by rail relocations or route closures, contributed to higher vacancy rates. Second, areas with weakened economic foundations due to industrial restructuring showed a higher risk of housing vacancies. While regions with active manufacturing industries in the past experienced the development of industrial and agro-industrial complexes, recent declines in manufacturing competitiveness and operational rates have led to economic stagnation. As a result, reduced worker inflows and declining local housing demand have increased the likelihood of housing vacancies. Third, natural constraints were also found to influence housing vacancy. In agriculture-focused regions, frequent flooding and natural disasters often resulted in poor living conditions. These factors contributed to long-term resident outmigration, ultimately leading to an increase in housing vacancies. In contrast, areas with well-connected transportation networks or relatively stable tourism and agricultural bases showed lower vacancy risks. Overall, the risk of housing vacancies in non-urban areas exhibited significant variation depending on industrial and economic structures, accessibility to urban centers, and the level of regional infrastructure.

〈Figure 9(b)〉 presents the spatial distribution of buildings with a high risk of vacancy in the non-urban areas of Jinju-si.



(a) High-risk areas for housing vacancy

(b) Houses with high vacancy rates

Figure 9. Vacancy risk in non-urban areas based on hotspot analysis

The analysis revealed that vacancy risks were predominantly high in regions engaged in agriculture- and manufacturing-based economic activities, as well as in peripheral areas with limited transportation accessibility. Regions with weakened economic foundations exhibited higher risks of housing vacancies. Areas previously characterized by thriving agriculture and manufacturing activities have recently experienced increasing vacancy rates due to industrial restructuring and a decline in economic activity. These regions, often concentrated with agro-industrial or manufacturing complexes, have seen decreased building utilization as industrial competitiveness has eroded and population outflows have intensified. Moreover, areas with limited accessibility to urban centers also demonstrated elevated vacancy risks. These regions often suffer from poor connectivity to major transportation infrastructure or are located at considerable distances from city centers, restricting access to essential residential and commercial amenities. This lack of accessibility diminishes the residential appeal of these areas, acting as a significant factor that exacerbates vacancy risks. In contrast, areas with lower vacancy risks were rural regions characterized by relatively high accessibility to urban centers and well-developed residential and commercial infrastructure. These areas, benefiting from a balanced influence of urbanization, maintained stable economic activities and exhibited reduced vacancy risks. These findings underscore the need for vacancy management policies tailored to the specific characteristics of non-urban areas. In regions experiencing economic stagnation, strategies aimed at industrial revitalization and enhancing residential appeal are essential. Meanwhile, in areas with limited transportation accessibility, policies focused on improving physical connectivity could serve as effective measures to mitigate housing vacancy issues.

Ⅵ. Conclusion

1. Research Summary and Implications

This study analyzed the characteristics and key factors of housing vacancy occurrences in urban and non-urban areas of Jinju-si, highlighting significant differences in the spatial patterns and causes of vacancies between these two regions. The findings underscore the necessity for tailored vacancy management strategies that account for regional characteristics. Using a Spatial GAMLSS model, the probabilities of housing vacancy occurrences in urban and non-urban areas were estimated, enabling the effective prediction of high-risk zones and vulnerable properties. This approach provided a detailed understanding of the spatial features and influential factors driving housing vacancy issues. The results offer valuable insights for devising targeted policy responses to address these challenges effectively.

In urban areas, the risk of housing vacancy was higher in old city centers, early-developed residential neighborhoods, and areas with declining commercial activity. This trend can be attributed to the gradual marginalization of older urban districts during the urbanization process, alongside the weakening of commercial and infrastructure functions. Notably, urban fringe areas with limited transportation infrastructure or reduced accessibility also exhibited a higher likelihood of vacancies. In contrast, regions with well-established urban infrastructure, such as new towns and areas near transportation hubs, showed relatively low vacancy risks. These findings highlight the necessity of policy interventions to address spatial imbalances in urban economic activity and housing demand. Strategies such as urban center regeneration, commercial district revitalization, and transit-oriented development are essential to mitigate these disparities. In non-urban areas, the risk of housing vacancy was higher in regions where agricultural and manufacturing-based economic activities have weakened and in peripheral areas with limited accessibility to city centers. This trend appears to be driven by a decline in residential attractiveness due to economic base deterioration caused by industrial structural changes and insufficient connectivity to transportation networks. Conversely, areas with stable agricultural and tourism bases or convenient access to urban centers exhibited relatively low vacancy risks. These findings underscore the need to address housing vacancy issues in non-urban areas through economic base strengthening and infrastructure enhancement. Furthermore, a comparative analysis of vacancy factors between urban and non-urban areas revealed distinct differences. For instance, owner-occupancy status was identified as a significant factor in suppressing housing vacancy in non-urban areas, while it did not exert a notable influence in urban areas. This suggests that residential stability and a sense of community play a critical role in mitigating vacancy risks in non-urban contexts. The distance to water boundaries had contrasting effects depending on regional characteristics. In urban areas, homes closer to water boundaries were associated with higher residential desirability. In contrast, in non-urban areas, proximity to water boundaries increased the likelihood of housing vacancy due to greater exposure to natural disaster risks. These findings highlight the necessity of tailored vacancy management policies that consider regional characteristics and adopt differentiated approaches for urban and non-urban areas.

The housing vacancy probabilities estimated using the Spatial GAMLSS model enabled the visualization of high-risk vacancy areas and individual at-risk buildings. The analysis revealed that in urban areas, vacancy risks were concentrated in aging residential zones and areas with poor transportation accessibility. In non-urban areas, regions with weakened agricultural and manufacturing bases, as well as peripheral locations, exhibited higher vacancy risks. Additionally, risk predictions at the individual building level provided practical evidence for identifying buildings in urgent need of management, facilitating the efficient allocation of resources and prioritization of problem-solving efforts. In conclusion, this study underscores the necessity of adopting differentiated approaches that incorporate regional characteristics and problem-specific factors to address housing vacancy issues in urban and non-urban areas. In urban areas, priorities should include the revitalization of old downtowns, activation of commercial districts, and development of transit-oriented hubs. In non-urban areas, strategies such as revitalizing industrial complexes, improving transportation networks, and strengthening economic foundations are required. Proactive interventions targeting high-risk areas and buildings can prevent the exacerbation of housing vacancy issues. Furthermore, region-specific policies can promote efficient resource utilization and effective problem-solving tailored to the unique needs of each area.

2. Limitations and Future Research Directions

This study conducted a comparative analysis of housing vacancy characteristics in urban and non-urban areas of Jinju-si using a Spatial GAMLSS model and hotspot analysis. It identified the key factors and spatial distributions associated with housing vacancies and effectively predicted high-risk areas and individual properties prone to vacancy. By analyzing the differentiated factors driving housing vacancies in urban and non-urban areas, the study highlighted the necessity of formulating tailored policies based on regional characteristics. Furthermore, it provided policymakers with actionable data that can serve as practical decision-making tools. In urban areas, the risk of housing vacancies was found to be higher in declining commercial districts and aging residential neighborhoods, whereas in non-urban areas, limited transportation accessibility and weakened economic foundations emerged as key factors. These findings underscore the necessity for differentiated vacancy management strategies tailored to the specific contexts of urban and non-urban areas. They also provide critical baseline data for developing detailed intervention plans targeting high-risk areas and individual properties prone to vacancy.

Nevertheless, this study has four key limitations. First, it is based on cross-sectional data and does not account for temporal trends in housing vacancies. Incorporating longitudinal data through time-series analysis could enable a deeper understanding of the dynamic changes in housing vacancies and provide a more comprehensive evaluation of the effects of policy interventions over time. Second, while this study focuses on the regional characteristics of Jinju-si, the findings may not be directly applicable to other areas due to varying local contexts. Comparative studies across diverse regions are necessary to generalize the results. Third, this research primarily emphasizes spatial factors and does not fully account for the interactions with socio-economic variables such as income levels, employment rates, and educational environments. Integrating these factors could provide a clearer understanding of the complex causes behind housing vacancies. Fourth, although the study offers policy implications, it does not evaluate the feasibility or effectiveness of the proposed policies in real-world applications.

To address these limitations, future research should proceed in the following directions. First, it is essential to classify urban areas into more specific types, such as old downtowns, new developments, and commercial centers, as well as non-urban areas into agricultural or tourism-focused zones, to analyze the characteristics and factors of housing vacancies in greater detail. This approach would enable the development of tailored policy recommendations for each subregion. Additionally, utilizing temporal data to examine the dynamic changes in housing vacancy rates and predict long-term vacancy risks is imperative. Such analyses could provide a foundation for designing sustainable housing vacancy management and response strategies. Third, while this study focused primarily on spatial factors, housing vacancy issues are intricately linked to socioeconomic variables such as income levels, educational environments, and demographic structures. For instance, regions with lower income levels may face difficulties in maintaining and repairing homes, increasing the likelihood of vacancies. Similarly, areas with inadequate educational facilities may see reduced residential appeal, leading to population outflows. From a demographic perspective, aging populations and youth migration can be significant drivers of housing vacancies. By integrating socioeconomic factors with spatial considerations, future research could more effectively identify the complex causes of housing vacancies. To achieve this, it would be beneficial to include a broader range of variables—such as household income, educational attainment, population density, age distribution, and employment rates—into the modeling process. Employing dimensionality reduction techniques, such as principal component analysis, or hierarchical modeling approaches could help manage computational complexity while maintaining analytical rigor. This comprehensive approach would provide a clearer understanding of the multifaceted nature of housing vacancy dynamics.

In conclusion, this study compared and analyzed the factors and spatial characteristics of housing vacancies in urban and non-urban areas of Jinju-si, emphasizing the necessity of tailored vacancy management policies based on regional characteristics. Future research should adopt a long-term and integrated approach to identify the root causes of housing vacancies and propose practical and effective policies for urban and regional regeneration. This study serves as a starting point for understanding housing vacancy trends and improving more reliable predictive models.

References

1. Amandolia, K.A., 2021. "Exploration of Algorithmic Bias in Machine Learning Methods Used to Predict Vacancy Rates", Master's Thesis, The University of North Carolina at Chapel Hill.
2. Appel, S.U., Botti, D., Jamison, J., Plant, L., Shyr, J.Y., and Varshney, L.R., 2014. "Predictive Analytics Can Facilitate Proactive Property Vacancy Policies for Cities", *Technological Forecasting and Social Change*, 89: 161-173.
3. Baba, H. and Hino, K., 2019. "Factors and Tendencies of Housing Abandonment: An Analysis of A Survey of Vacant Houses in Kawaguchi City, Saitama", *Japan Architectural Review*, 2(3): 367-375.
4. Baba, H. and Shimizu, C., 2023. "The Impact of Apartment Vacancies on Nearby Housing Rents over Multiple Time Periods: Application of Smart Meter Data", *International Journal of Housing Markets and Analysis*, 16(7): 27-41.
5. Bassett, E.M., Schweitzer, J., and Panken, S., 2006. *Understanding Housing Abandonment And Owner Decision-Making in Flint, Michigan: An Exploratory Analysis*, Lincoln Institute of Land Policy Working Paper.
6. Basu, S. and Thibodeau, T.G., 1998. "Analysis of Spatial Autocorrelation in House Prices", *The Journal of Real Estate Finance and Economics*, 17: 61-85.
7. Cajias, M., 2018. "Is There Room for Another Hedonic Model? The Advantages of the GAMLSS Approach in Real Estate Research", *Journal of European Real Estate Research*, 11(2): 204-245.
8. Chen, M., Bolt, G., Yu, L., and Hooimeijer, P., 2023. "The Impact of the Residential Environment on Chinese Older People's Aging-in-Place Intentions: A Mediation and Moderation Analysis", *Habitat International*, 140: 102908.
9. Clark, D.E. and Herrin, W.E., 1997. "Historical Preservation Districts and Home Sale Prices: Evidence from the Sacramento Housing Market", *Review of Regional Studies*, 27(1): 29-48.
10. Cozens, P. and Tarca, M., 2016. "Exploring Housing Maintenance and Vacancy in Western Australia: Perceptions of Crime and Crime Prevention through Environmental Design (CPTED)", *Property Management*, 34(3): 199-220.
11. De Bastiani, F., Rigby, R.A., Stasinopoulous, D.M., Cysneiros, A.H., and Uribe-Opazo, M.A., 2018. "Gaussian Markov Random Field Spatial Models in GAMLSS", *Journal of Applied Statistics*, 45(1): 168-186.
12. Deng, C. and Ma, J., 2015. "Viewing Urban Decay from the Sky: A Multi-Scale Analysis of Residential Vacancy in a Shrinking Us City", *Landscape and Urban Planning*, 141: 88-99.
13. Edward, L.G. and Joseph, G.K., 2005. "Urban Decline and Durable Housing", *The Journal of Political Economy*, 113: 345-375.
14. Florencio, L., Cribari-Neto, F., and Ospina, R., 2011. "Real Estate Appraisal of Land Lots Using GAMLSS Models", arXiv preprint arXiv:1102.2015.

15. Geoghegan, J., Wainger, L.A., and Bockstael, N.E., 1997. "Spatial Landscape Indices in a Hedonic Framework: An Ecological Economics Analysis Using GIS", *Ecological Economics*, 23(3): 251-264.

16. Gu, D., Newman, G., Kim, J.H., Park, Y., and Lee, J., 2019. "Neighborhood Decline and Mixed Land Uses: Mitigating Housing Abandonment in Shrinking Cities", *Land Use Policy*, 83: 505-511.

17. Hackman, D.A., Robert, S.A., Grübel, J., Weibel, R.P., Anagnostou, E., Hölscher, C., and Schinazi, V.R., 2019. "Neighborhood Environments Influence Emotion and Physiological Reactivity", *Scientific Reports*, 9(1): 9498.

18. Halvorsen, R. and Pollakowski, H.O., 1981. "Choice of Functional Form for Hedonic Price Equations", *Journal of Urban Economics*, 10(1): 37-49.

19. Henry, M.S., Schmitt, B., and Piguet, V., 2001. "Spatial Econometric Models for Simultaneous Systems: Application to Rural Community Growth in France", *International Regional Science Review*, 24(2): 171-193.

20. Immergluck, D., 2016. "Examining Changes in Long-term Neighborhood Housing Vacancy during the 2011 to 2014 U.S. National Recovery", *Journal of Urban Affairs*, 38: 607-622.

21. Joo, H.S. and Lee, S.Y., 2021. "Spatial Analysis of Abandoned Houses and Their Influencing Factors in South Korea", *Applied Sciences*, 11(18): 8576.

22. Joo, H.S., Lee, S.Y., Kang, S.J., and Kim, S.Y., 2022. "Vacant House Characteristics by Use Area and Their Application to Sustainable Community", *Applied Sciences*, 12(21): 10696.

23. Lester, T.W., Kaza, N., and Kirk, S., 2013. "Making Room for Manufacturing: Understanding Industrial Land Conversion in Cities", *Journal of the American Planning Association*, 79(4): 295-313.

24. Lee, J. and Newman, G., 2017. "Forecasting Urban Vacancy Dynamics in A Shrinking City: A Land Transformation Model", *ISPRS International Journal of Geo-Information*, 6(4): 124.

25. Lee, J., Kim, H., and Kim, H., 2021. "Commercial Vacancy Prediction Using LSTM Neural Networks", *Sustainability*, 13(10): 5400.

26. Lee, J., Newman, G., and Lee, C., 2022. "Predicting Detached Housing Vacancy: A Multilevel Analysis", *Sustainability*, 14(2): 922.

27. Liu, Q., Sho, K., Corcoran, J., and Ando, N., 2024. "Effects of Neighbourhood Characteristics on the Distribution of Vacant Houses in Toda City, a Satellite City of Tokyo Metropolis of Japan", *Scientific Reports*, 14(1): 2045-2322.

28. Ma, J., Cheng, J.C., Jiang, F., Chen, W., and Zhang, J., 2020. "Analyzing Driving Factors of Land Values in Urban Scale Based on Big Data and Non-Linear Machine Learning Techniques", *Land Use Policy*, 94: 104537.

29. Mallach, A., 2017. "What We Talk about When We Talk about Shrinking Cities: The Ambiguity of Discourse and Policy Response in the United States", *Cities*, 69:109-115.

30. Marmolejo-Ramos, F., Tejo, M., Brabec, M., Kuzilek, J., Joksimovic, S., Kovanovic, V., González, J., Kneib, T., Bühlmann, P., Kook, L., Briseño-Sánchez, G., and Ospina, R., 2022. Distributional Regression Modeling via Generalized Additive Models for Location, Scale, and Shape: An Overview through a Data Set from Learning Analytics", *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 13(1): e1479.

31. Martin, H., Whitaker, S.D., Oduro, I., Johnson, É., Richter, F.G.C., and Urban, A.H., 2017. "Predictive Modeling of Surveyed Property Conditions And Vacancy", the paper presented at he 18th Annual International Conference on Digital Government Research, 358-367, New York.

32. Molloy, R., 2016. "Long-term Vacant Housing in the United States", *Regional Science and Urban Economics*, 59: 118-129.

33. Morckel, V.C., 2013. "Empty Neighborhoods: Using Constructs to Predict the Probability of Housing Abandonment", *Housing Policy Debate*, 23(3): 469-496.

34. Morckel, V.C., 2014. "Spatial Characteristics of Housing Abandonment", *Applied Geography*, 48: 8-16.

35. Nadalin, V. and Igliori, D., 2017. "Empty Spaces in the Crowd. Residential Vacancy in São Paulo's City Centre", *Urban Studies*, 54(13): 3085-3100.

36. Nassauer, J.I. and Raskin, J., 2014. "Urban Vacancy and Land Use Legacies: A Frontier for Urban Ecological Research, Design, and Planning", *Landscape and Urban Planning*, 125: 245-253.

37. Newman, G., 2013. "A Conceptual Model for Measuring Neglect in Historic Districts", *Journal of Preservation, Education, and Research*, 6: 45-57.

38. Newman, G., Lee, J., and Berke, P., 2016. "Using the Land Transformation Model to Forecast Vacant Land", *Journal of Land Use Science*, 11(4): 450-475.

39. Page, M., 2001. "Comment on James R. Cohen's 'Abandoned Housing: Exploring Lessons from Baltimore'", *Housing Policy Debate*, 12(3): 457-463.

40. Ribant, M. and Chen, X., 2020. "A Typology of US Shrinking Cities", *The Professional Geographer*, 72(1): 152-164.

41. Rigby, R.A. and Stasinopoulos, D.M., 2005. "Generalized Additive Models for Location, Scale and Shape (GAMLSS)", *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 54(3): 507-554.

42. Sargent, J.D., Bailey, A., Simon, P., Blake, M., and Dalton, M.A., 1997. "Census Tract Analysis of Lead Exposure in Rhode Island Children", *Environmental Research*, 74(2): 159-168.

43. Shmueli, G., 2010. "To Explain or to Predict?", *Statistical Science*, 25(3): 289-310.

44. Terada, T., Tanaka, T., and Yokohari, M., 2016. "Directions for Vacant Lot Management in the Outer Suburbs of the Tokyo Metropolitan Region", *Urban and Regional Planning Review*, 3: 66-84.

45. Wilhelmsson, M., Andersson, R., and Klingborg, K., 2011. "Rent control and Vacancies in Sweden", *International Journal of Housing Markets and Analysis*, 4(2):105-129.

46. Yin, L. and Silverman, R., 2015. "Housing Abandonment and Demolition: Exploring the Use of Micro-level and Multi-year

Models", *International Journal of Geo-Information*, 4(3): 1184-1200.

47. Yue, X., Wang, Y., and Zhang, H.O., 2022. "Influences of the Plot Area and Floor Area Ratio of Residential Quarters on the Housing Vacancy Rate: A Case Study of the Guangzhou Metropolitan Area in China", *Buildings*, 12(8): 1197.
48. Zhang, D.D., Yan, D.H., Wang, Y.C., Lu, F., and Liu, S.H., 2015. "GAMLSS-based Nonstationary Modeling of Extreme Precipitation in Beijing–Tianjin–Hebei Region of China", *Natural Hazards*, 77: 1037-1053.
49. Zhao, W., 2022. "The Long-Run Effects of Minimum Lot Size Zoning on Housing Redevelopment", *Journal of Housing Economics*, 55: 101806.
50. Statistics Korea. "Number of Houses by Type of Housing, Eup/Myeon/Dong (Years ending in 0 and 5), Si/ Gun/Gu (Other years) in 2021", Accessed December 22, 2024. https://kosis.kr/index/index.do

Date Received	2025-03-28
Reviewed(1st)	2025-04-27
Date Revised	2025-05-02
Reviewed(2nd)	2025-06-12
Date Accepted	2025-06-12
Final Received	2025-06-30