

## Research Article

# Temporal Variation and Spatial Heterogeneity in Boarding/Alighting Patterns at Urban Railway Stations: Implications for Estimating Optimal Construction Scale—A Case Study in Seoul, South Korea

WooKeol Cho , Kyoungtae Kim , Inmook Lee , and Soyoung (Iris) You 

*Innovative Transportation and Logistics Research Center, Korea Railroad Research Institute, 176 Railroad Museum Road, Uiwang, Republic of Korea*

Correspondence should be addressed to Soyoung (Iris) You; [syyou@krii.re.kr](mailto:syyou@krii.re.kr)

Received 25 October 2023; Revised 8 March 2024; Accepted 4 April 2024; Published 23 April 2024

Academic Editor: Yu Qian

Copyright © 2024 WooKeol Cho et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Urban railways have become a prominent mode of public transportation within cities owing to their connectivity with other modes of transport and environmental friendliness. Various policies, such as the expansion of metropolitan areas and the development of megacities, have further emphasized the pivotal role of urban railways. Consequently, more railway stations are expected to be constructed in developed cities. However, the temporal variation in boarding and alighting patterns at each railway station is often overlooked. Failing to account for this variation, specifically the differences in peak-hour concentration rates, in railway station design may cause increased conflicts among users owing to concentrated demands during specific time periods, exacerbating congestion and diminishing the appeal of the urban railway systems. Therefore, this study investigated the correlation between the temporal variation in boarding and alighting patterns and the attributes (location) of railway stations in Seoul, South Korea, and analyzed the spatial heterogeneity of this correlation. Initially, the factors influencing the peak-hour concentration rates in railway stations were identified using a linear regression model. Peak hours were defined as morning and afternoon peaks and boarding and alighting were differentiated to account for the directional aspects of temporal variations in boarding and alighting patterns. The correlation between boarding and alighting patterns and the attributes of railway station influence zones was determined, and a geographically weighted regression model was estimated to analyze the spatial heterogeneity of this correlation based on railway station location. The analysis results revealed that railway stations in the southeastern and downtown areas of Seoul exhibited varying impacts of station attributes on boarding and alighting patterns even when the station attribute influence zones were identical. The contribution of this study is to evaluate the priorities of railway projects and its corresponding transportation policies. Regarding the policy goal recently announced by the Korean government, “Achieving Commute Times in 30-min range,” our finding will provide a good measure of accessibility whether it succeeds or not.

## 1. Introduction

Urban railways are a primary mode of public transportation for local and regional trips within cities and metropolitan areas and play a central role in the public transportation network. In addition, their environmental friendliness has led to the introduction and expansion of urban rail systems in many countries in accordance with the global focus on reducing carbon emissions. In Seoul, South Korea, approximately four million people use the urban railway system daily. With integration with

other modes of transportation, such as buses, personal mobility (PM), and bike-sharing, the demand for urban railways is projected to continuously increase over time.

Moreover, the emergence of concepts, such as metropolitan regions and megacities, may impact the growing demand for urban railways. These concepts were proposed as alternatives for initial regionally balanced development, aimed at enhancing the competitiveness of surrounding cities and improving the efficiency of various urban projects through collective resource mobilization. These concepts are exemplified by initiatives in

Germany, France, and the United Kingdom. However, a growing concern has emerged in that developing these metropolitan regions and megacities may lead to unintended consequences, such as the “straw effect,” where population and investment concentrate in major cities, causing a decline in surrounding cities. Paradoxically, this trend may further strain transportation infrastructure in major cities owing to increased transportation demands.

Adding new urban railway stations to already developed cities is necessary to address this problem. In this process, the boarding and alighting demands at each railway station are predicted using origin–destination (O/D) and road/rail network data. Typically, station size is estimated based on daily boarding and alighting demands, which may lead to over- or undersizing of railway stations owing to the lack of consideration of the temporal variation in boarding and alighting demands, particularly the peak-hour concentration rates. For example, determining the appropriate station size based solely on daily boarding and alighting demands and constructing the station without accounting for temporal variations in these demands may increase congestion within the railway station, resulting in user discomfort and potential safety hazards. In the Seoul metropolitan area, the Seoul Subway Line 9, currently in operation, has failed to adequately consider the traffic demand concentrated during the morning and afternoon peak hours in its design process. As a result, during these time periods, almost all sections of the line experienced severe congestion, leading to significant discomfort for citizens, including safety concerns. Furthermore, the Gimpo Gold Line, which connects Seoul and Gimpo (city adjacent to western Seoul), has also been designed and constructed without sufficient consideration for peak-hour traffic demand. As a result, frequent safety incidents occur among passengers waiting inside the stations. In addition, considering the directional nature of the temporal variation in demand is crucial. For example, in areas with concentrated business facilities, alighting demand may peak in the morning, whereas boarding demand may peak in the afternoon. Failing to account for this may exacerbate conflicts within railway stations. In other words, failure to consider the concentration of peak-hour traffic demand not only deteriorates the level of service of public transportation services but also increases the risk of safety accidents.

When establishing new railway stations in relatively less-developed cities, the demand characteristics can be observed, and strategies can be continually formulated according to land-use changes around the railway station. However, in fully developed areas, the characteristics of boarding and alighting patterns must be considered when determining the appropriate station size.

Therefore, this study aims to identify the factors influencing the temporal variation in boarding and alighting patterns at urban railway stations in South Korea, focusing on Seoul, and analyze the spatial heterogeneity of these influences. Factors, including station location and the attributes of station influence zones, were selected. Statistical models were estimated using the data collected to explain these attributes. Initially, a linear regression model was applied to

Seoul’s urban railway stations to ascertain the factors influencing boarding and alighting patterns. In addition, a geographically weighted regression (GWR) model, a widely used spatial regression tool, was employed to analyze the size and directionality of the influencing factors based on station location and influence zone attributes. One-week public transportation card data were used to ensure the reliability of the research results and account for daily variations in boarding and alighting patterns. To the best of our knowledge, research on the analysis of boarding and alighting patterns at urban railway stations considering station locations and influence zone attributes for the purpose of informing safety planning has been limited.

The remainder of this paper is organized as follows: Section 2 reviews previous related studies. Section 3 introduces the study area, details of the data collection methods, and basic statistics of the collected data. Section 4 introduces the methodology used in the study. Section 5 presents the results of model estimation. Finally, Section 6 concludes the paper and offers recommendations.

## 2. Literature Review

Previous studies on congestion within urban railway stations analyzed congestion levels using various data and methodologies. Analyses of the factors influencing pedestrian flow within railway stations and studies on pedestrian simulations have been prominent. Ahn et al. [1] performed a simulation based on a gravity model using survey data on pedestrian routes to analyze pedestrian flows for railway station planning and management. Yang and Tang [2] analyzed the effects of adjusting departure times for urban railway passengers based on fare discounts to alleviate peak-hour passenger concentrations. Teng et al. [3] conducted SP/RP surveys to analyze the psychological factors influencing passengers in crowded and conflicting pedestrian environments within railway stations. They quantitatively analyzed the impact using regression models. Li et al. [4] focused on the correlation between bike-sharing and urban railways and analyzed the correlation between land use and railway station usage patterns using bike-sharing usage data. Li et al. [5] developed a method based on rail network data for analyzing congestion patterns in an urban railway network resulting from train delays. Wang et al. [6] analyzed the characteristics of unexpected passenger concentrations in an intelligent urban rail transit network and proposed methods to manage this congestion. Jiao et al. [7] analyzed the correlation between land use and boarding and alighting patterns based on the spatiotemporal similarity of the patterns. The related prior studies mentioned above are summarized in Table 1.

Liang et al. [8] calculated an appropriate fare through a bilevel model using train operation data to establish a fair adjustment strategy for alleviating peak-hour congestion. Yu et al. [9] utilized smart card data (SCD) and clustering algorithms to categorize the heterogeneity in peak-hour travel patterns at railway stations, assuming that the occurrence times of peak-hour travel patterns vary per station. Huang et al. [10] estimated a weighted linear regression model using

TABLE 1: Literature review of congestion in railway stations.

| Research          | Research objective  | Data                            | Methodology  |
|-------------------|---|---------------------------------|--|
| Ahn et al. [1]    | Analysis of passenger pedestrian flow for railway station planning and management   | Survey data (walking path)      | Gravity model, VISSIM simulation                           |
| Yang and Tang [2] | Analysis of peak-hour passenger departure time adjustment effects due to fare discounts   | OD, network                     | Traffic assignment (UE), queuing theory                    |
| Teng et al. [3]   | Analysis of psychological factors influencing passengers' perception of pedestrian environment within railway stations                              | Survey data (SP, RP)            | Perceived congestion quantization model (regression model) |
| Li et al. [4]     | Analysis of the correlation between land use and railway station utilization patterns   | Bike-sharing usage history data | Random forest model  |
| Li et al. [5]     | Development of methodology for analyzing congestion patterns in urban rail networks due to train delays   | Rail network data               | Passenger flow congestion model                            |
| Wang et al. [6]   | Analysis of the characteristics of sudden passenger flow and proposal of congestion management methods in an intelligent urban rail transit network | Simulation data                 | Statistical analysis                                       |
| Jiao et al. [7]   | Analysis of the correlation between land use and boarding/alighting patterns based on spatiotemporal similarity in boarding/alighting patterns      | Smart card data                 | Clustering algorithm                                       |

socioeconomic indicators and land-use data to develop an accessibility-based method for estimating peak-hour light rail transit ridership. Li et al. [11] used SCD to estimate hybrid and ARIMA models for short- to medium-term prediction of urban railway boarding and alighting demands. Wang et al. [12] derived optimal train operation plans using an optimization method to develop a coordinated operation strategy for individually operated railway stations to minimize passenger delays. Gulhan et al. [13] employed a timetable-based assignment methodology to develop accessibility indicators for urban public transit (UPT) during the planning stage. Gutiérrez et al. [14] estimated a distance-decay-weighted regression model to develop a station-level boarding estimation model integrated with a geographic information system (GIS) system. The related prior studies mentioned above are summarized in Table 2.

The above literature review indicates that studies simultaneously considering the spatial heterogeneity resulting from the locational characteristics of railway stations and the different urban railway boarding and alighting patterns on a daily and hourly basis are limited. In addition, studies simultaneously estimating models while considering socioeconomic indicators, boarding and alighting patterns, and land-use characteristics and providing causal relationships between the estimated results and characteristics of the study area are also scarce. Therefore, this study makes a distinctive contribution compared with existing studies.

### 3. Methodology

**3.1. Methodological Framework.** This study identified the factors influencing the temporal variation in boarding and alighting patterns at urban railway stations. Public transportation card data and statistical models were used to analyze the spatial heterogeneity of these influences. First, SCD were processed to calculate the morning and afternoon peak-hour concentration rates of boarding and alighting demands at each urban railway station. Using the station-specific peak-hour concentration rates as the dependent variable and the locational attributes of the station as independent variables, a linear regression model was estimated to determine the factors influencing the peak-hour concentration rates of boarding and alighting demands at each station. In addition, a GWR model was estimated to elucidate spatial heterogeneity based on the locational attributes of each station. The results of the model estimation (estimated regression coefficients) were visualized using GIS tools. The methodological framework of this study is illustrated in Figure 1.

**3.2. GWR Model.** This study employed a GWR model to analyze the factors influencing the temporal variability of urban railway station usage demand and the spatial heterogeneity of these influences. GWR is a regression analysis model that can be used to analyze data collected in a spatial unit. It has been applied in various fields, such as transportation and geography [15–24].

Considering spatial autocorrelation (the property in which spatially adjacent samples have a high correlation), the GWR model enables the analysis of ripple effects that occur due to

data aggregation and spatial proximity in the study area. Owing to these spatial characteristics, estimating data with spatial dependency and heterogeneity using the commonly used ordinary least squares regression analysis model violates the assumption of linearity and results in errors, reducing the efficiency of parameter estimation. To address the heteroscedasticity arising from spatial heterogeneity, the GWR model utilizes a weighting function and considers regression coefficients as functions of location, allowing for an analysis in which coefficients vary depending on spatial position. This can be expressed mathematically as follows:

$$y_i(\mu) = \beta_{0i}(\mu) + \beta_{1i}(\mu)x_1 + \beta_{2i}(\mu)x_2 + \cdots + \beta_{mi}(\mu)x_m + \varepsilon_i(\mu), \quad (1)$$

where  $y_i$  = dependent variable ( $i = 1, 2, \dots, n$ ),  $n$  is the number of observations,  $x_{mi}$  =  $m$ th independent variable of observation  $i$ ,  $\beta_{mi}$  =  $m$ th regression coefficient of observation  $i$ ,  $\varepsilon_i$  = error term,  $\mu$  = spatial coordinates.

What distinguishes GWR from conventional regression models is the inclusion of  $\mu$  in each term, indicating that parameter estimation is performed for a given  $\mu$  and is relevant only for that specific location. The regression coefficients according to location were estimated using the weighted least squares estimation method, expressed as follows:

$$\hat{\beta} = [X^T W(u) X]^{-1} X^T W(u) Y. \quad (2)$$

In this equation, each element of the geographical weight matrix  $W(u)$  is calculated based on the weighting function determined by the kernel function. Although the kernel function can take various forms, this study employs the most common Gaussian function, expressed as follows:

$$W_i(u) = \exp\{-0.5(d_i(u)/h)^2\}, \quad (3)$$

where  $W_i(u)$  = weight of observation  $i$  in spatial coordinates,  $d_i(u)$  = distance between observation  $i$  and spatial coordinates  $\mu$ ,  $h$  = bandwidth.

### 4. Data Collection

**4.1. Study Area.** This study focused on subway stations in Seoul, South Korea, home to ~10 million residents. As the largest city in South Korea, Seoul exhibits heterogeneous land use and demographic composition within its urban area. Not only the Seoul metropolitan area but also most major metropolitan cities have several distinct regions with varying attributes due to the different connectivity and relevance to an entire metropolitan area. To address this issue, this study divided Seoul into five regions and interpreted the results by considering the characteristics of each region. First, the southeastern region, the most developed area in Seoul, encompasses a high proportion of commercial facilities and a relatively large residential population. Moreover, due to the relatively high proportion of business facilities in this region, there is a significant population commuting to and from the cities in

TABLE 2: Literature review of spatiotemporal travel patterns at railway stations.

| Research              | Research objective   | Data  | Methodology                              |
|-----------------------|--|---|--|
| Liang et al. [8]      | Developing fare adjustment strategies for peak-hour congestion alleviation                                   | Train operation schedule data   | Bilevel model                            |
| Yu et al. [9]         | Analysis of heterogeneity in peak travel patterns at each railway station and its relationship with land use | Smart card data   | Clustering method                        |
| Huang et al. [10]     | Development of peak-hour LRT ridership estimation method based on accessibility                              | Socioeconomic attributes, land-use attributes                             | Weighted linear regression model         |
| Li et al. [11]        | developing short- to medium-term passenger flow forecasting models for metro stations                        | Smart card data   | Hybrid model, ARIMA                      |
| Wang et al. [12]      | Development of coordinated operations for peak-hour railway stations to minimize passenger delays            | Simulation data   | Mathematical optimization method         |
| Gulhan et al. [13]    | Development of UPT accessibility metrics at the planning stage   | OD, network   | Timetable-based assignment               |
| Gutiérrez et al. [14] | Development of station-level boarding and alighting prediction models integrated with GIS data               | Socioeconomic attributes, network, station-specific boarding volume, etc. | Distance-decay weighted regression model |

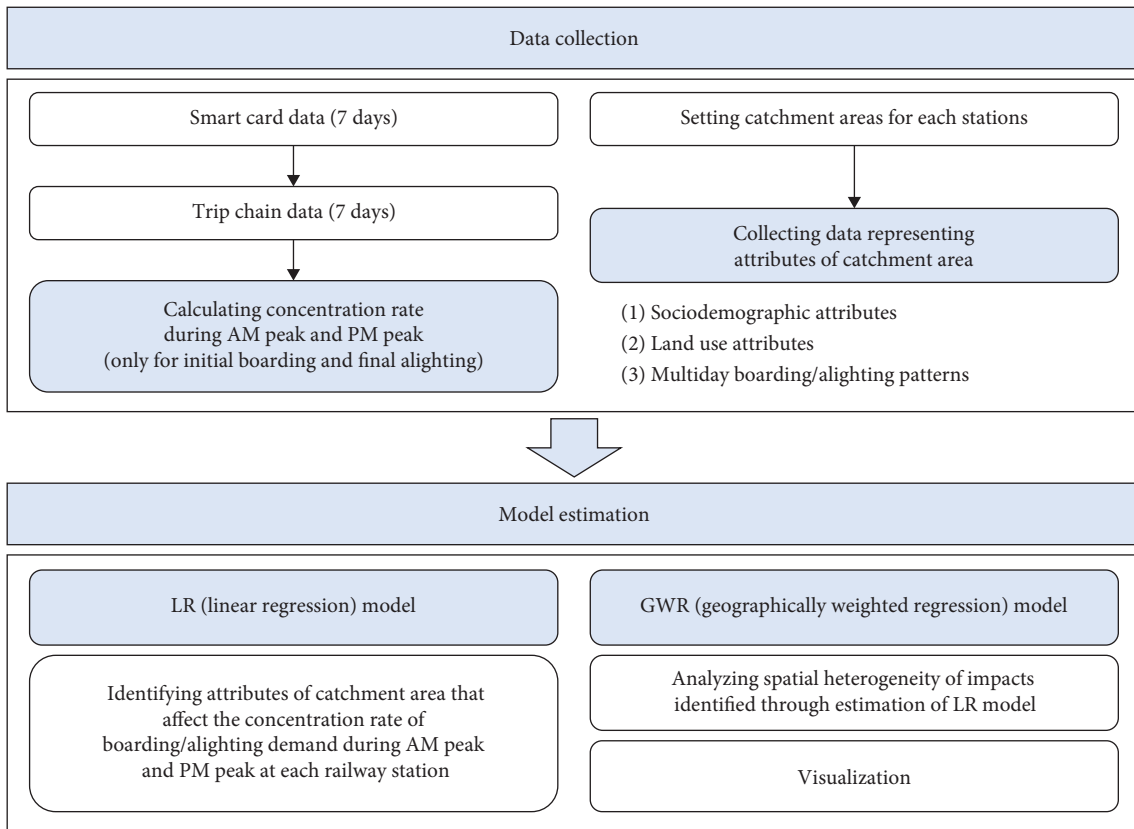


FIGURE 1: Methodological framework.



FIGURE 2: Classification of regions in Seoul, South Korea.

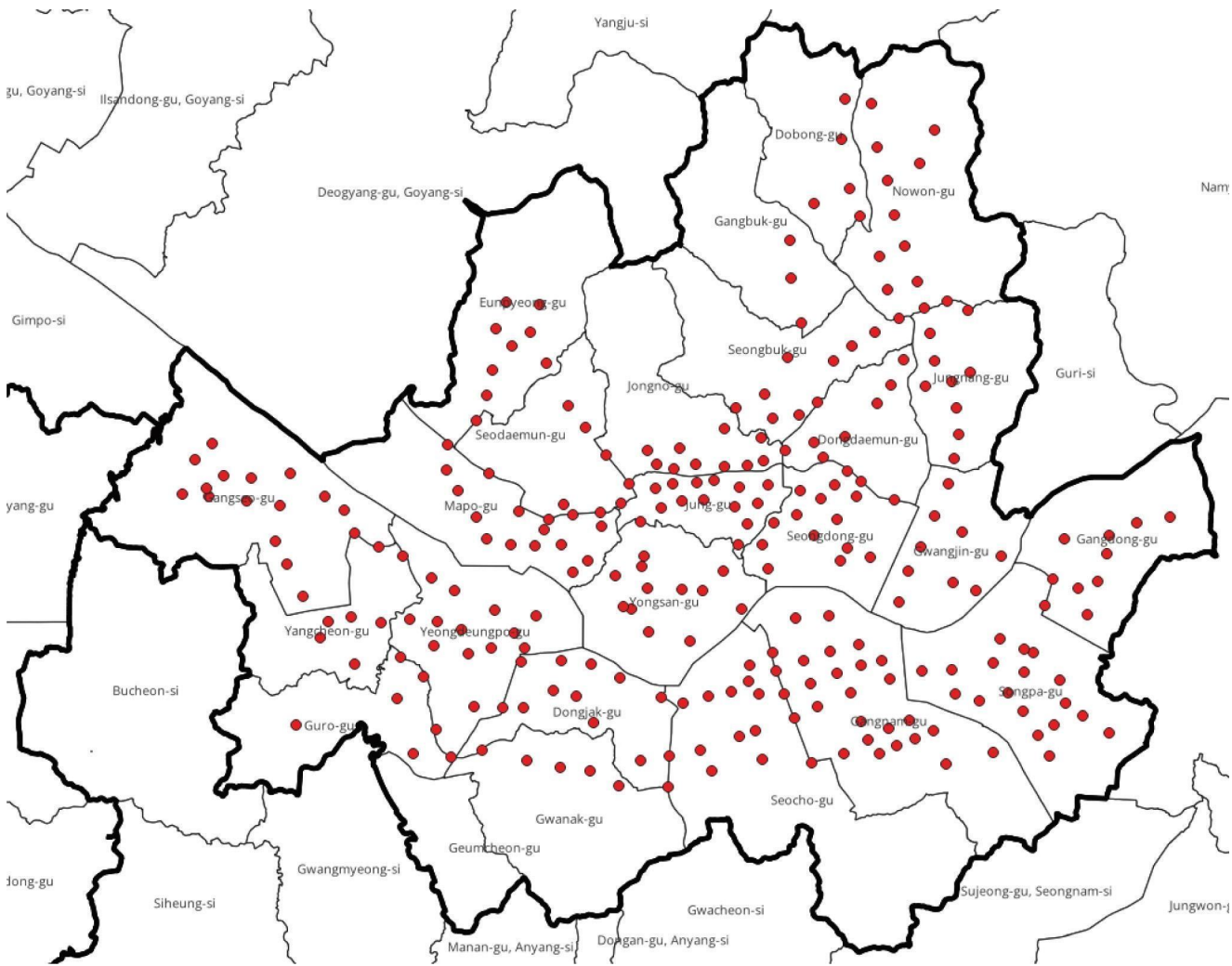
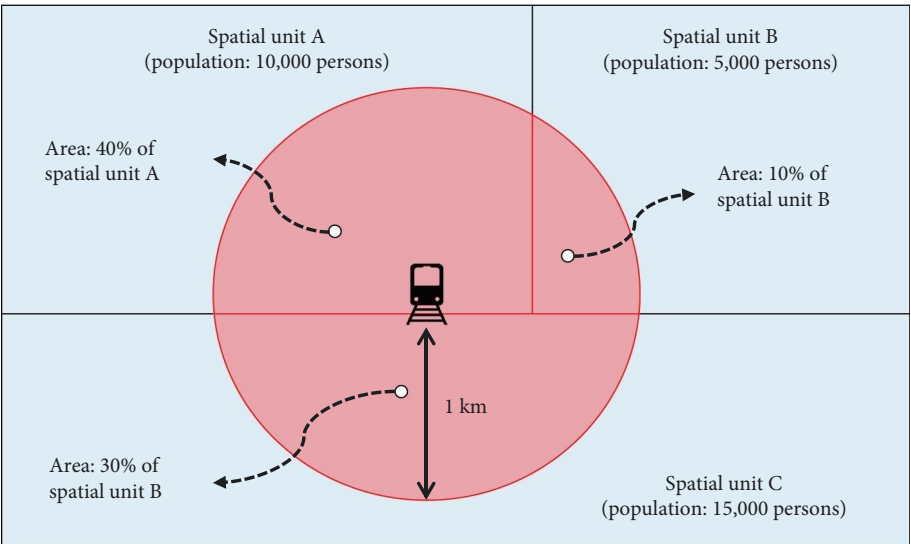


FIGURE 3: Railway stations used in this study.



※ Population of catchment area= (10,000 × 0.4) + (5,000 × 0.1) + (15,000 × 0.3) = 9,000

FIGURE 4: Estimation of sociodemographic attributes in the catchment areas of railway stations.

TABLE 3: Dependent/independent variables.

| Group  | Variables   |
|--|---|
| Dependent variables  | (i) BOAMR: concentration rate of boarding during AM peak (%)      |
|  | (ii) BOPMR: concentration rate of boarding during PM peak (%)     |
|  | (iii) ALAMR: concentration rate of alighting during AM peak (%)   |
|  | (iv) ALPMR: concentration rate of alighting during PM peak (%)    |
| Independent variables  | Multiday boarding/alighting attributes                            |
|  | (i) BOWDR: weekday concentration rate of boarding (%)             |
|  | (ii) ALWDR: weekday concentration rate of alighting (%)           |
|  | (iii) BOWER: weekend concentration rate of boarding (%)           |
|  | (iv) ALWER: weekend concentration rate of alighting (%)           |
|  | Infra attributes  |
|  | (v) NOLINE: number of transfer lines (lines)                      |
|  | (vi) SALEAREA: area of sale facilities (1,000 m <sup>2</sup> )    |
|  | (vii) CBAREA: area of business facilities (1,000 m <sup>2</sup> ) |
|  | Land-use attributes   |
|  | (viii) NOSHOP: number of hospitals (places)                       |
|  | (ix) NOGHOSP: number of general hospitals (places)                |
|  | (x) NOBED: number of beds in hospitals (places)                   |
|  | (xi) POP10–POP70: population by age (1,000 person)                |
| (a) 10, 20, 30, 40, 50, 60, over 70s                         |   |
| Sociodemographic attributes                                  |   |
| (xii) TPOP, MPOP, FPOP: population by gender (1,000 persons) |   |
| (b) Total, male, female                                      |   |
| (xiii) NOEMP: number of employees (1,000 persons)            |   |
| (xiv) NOHH: number of households (1,000 households)          |   |

TABLE 4: Statistical analysis results for the collected data.

| Group                 |          | Mean   | Std. dev | Max      | Min   |
|-----------------------|----------|--------|----------|----------|-------|
| Dependent variables   | BOAMR    | 1.61   | 0.27     | 2.25     | 0.70  |
|                       | BOPMR    | 1.11   | 0.16     | 1.94     | 0.81  |
|                       | ALAMR    | 1.72   | 0.16     | 2.17     | 1.22  |
|                       | ALPMR    | 1.11   | 0.16     | 1.94     | 0.81  |
| Independent variables | NOLINE   | 1.23   | 0.49     | 4.00     | 1.00  |
|                       | NOHH     | 29.07  | 10.95    | 63.94    | 4.45  |
|                       | TPOP     | 66.31  | 24.46    | 133.35   | 9.74  |
|                       | MPOP     | 32.18  | 11.91    | 65.01    | 4.92  |
|                       | FPOP     | 34.12  | 12.59    | 68.35    | 4.82  |
|                       | NOEMP    | 45.15  | 38.83    | 220.76   | 4.21  |
|                       | SALEAREA | 45.26  | 63.75    | 384.90   | 0.37  |
|                       | CBAREA   | 213.53 | 301.59   | 2,228.35 | 0.86  |
|                       | NOHOSP   | 139.51 | 84.40    | 541.94   | 8.44  |
|                       | NOGHOSP  | 0.44   | 0.55     | 2.95     | 0.00  |
|                       | NOHBED   | 625.44 | 430.44   | 2,162.96 | 14.10 |
|                       | POP10    | 5.44   | 2.45     | 15.03    | 0.57  |
|                       | POP20    | 10.47  | 4.38     | 28.72    | 1.39  |
|                       | POP30    | 10.69  | 4.24     | 24.44    | 1.73  |
|                       | POP40    | 10.63  | 4.03     | 20.72    | 1.48  |
|                       | POP50    | 10.45  | 4.17     | 22.85    | 1.50  |
|                       | POP60    | 7.78   | 3.12     | 17.55    | 1.21  |
|                       | POP70    | 6.42   | 2.54     | 12.94    | 0.97  |
|                       | BOWDR    | 79.19  | 3.92     | 89.16    | 64.28 |
|                       | ALWDR    | 79.24  | 3.94     | 89.07    | 65.38 |
| BOWER                 | 20.81    | 3.92   | 35.72    | 10.84    |       |
| ALWER                 | 20.76    | 3.94   | 34.62    | 10.93    |       |



TABLE 5: Estimation results of the linear regression model (BOAMR).

| Variable                 | Est          | SE       | T-stat | p-Value        |
|--------------------------|--------------|----------|--------|----------------|
| Intercept                | 20.059813*** | 1.957122 | 10.250 | $\leq 0.00001$ |
| Area of sales facilities | -0.036284*** | 0.008352 | -4.344 | 0.00002        |
| Number of households     | 0.581989***  | 0.045974 | 12.659 | $\leq 0.00001$ |
| Number of hospitals      | -0.053876*** | 0.005917 | -9.106 | $\leq 0.00001$ |
| Number of transfer lines | -3.637749*** | 1.049986 | -3.465 | 0.00062        |

-Number of observations: 267. -Adjusted  $R^2$ : 0.5462. \* Significance code: \*\*\* ( $p \leq 0.001$ ).

TABLE 6: Estimation results of linear regression model (BOPMR).

| Variable                    | Est         | SE       | T-stat  | p-Value        |
|-----------------------------|-------------|----------|---------|----------------|
| Intercept                   | 27.28373*** | 1.618062 | 16.862  | $\leq 0.00001$ |
| Area of business facilities | 0.004742*** | 0.001404 | 3.377   | 0.00084        |
| Number of households        | -0.47414*** | 0.037326 | -12.703 | $\leq 0.00001$ |
| Number of hospitals         | 0.046304*** | 0.00488  | 9.489   | $\leq 0.00001$ |
| Number of transfer lines    | 2.023987*   | 0.851589 | 2.377   | 0.01818        |

-Number of observations: 267. -Adjusted  $R^2$ : 0.5074. \* Significance code: \*\*\* ( $p \leq 0.001$ ), \* ( $p \leq 0.05$ ).

TABLE 7: Estimation results of linear regression model (ALAMR).

| Variable                    | Est         | SE       | T-stat | p-Value        |
|-----------------------------|-------------|----------|--------|----------------|
| Intercept                   | 32.18867*** | 1.712945 | 18.791 | $\leq 0.00001$ |
| Area of business facilities | 0.006701*** | 0.001863 | 3.597  | 0.00038        |
| Number of households        | -0.97772*** | 0.238769 | -4.095 | 0.00005        |
| Number of general hospitals | 2.769963**  | 1.022812 | 2.708  | 0.00721        |
| Population of 20 and 30s    | 0.707859*   | 0.313516 | 2.258  | 0.02478        |

-Number of observations: 267. -Adjusted  $R^2$ : 0.3006. \* Significance code: \*\*\* ( $p \leq 0.001$ ), \*\* ( $p \leq 0.01$ ), \* ( $p \leq 0.05$ ).

TABLE 8: Estimation results of the linear regression model (ALPMR).

| Variable  | Est         | SE       | T-stat | p-Value        |
|---|-------------|----------|--------|----------------|
| Intercept                                       | 13.3878***  | 1.788965 | 7.484  | $\leq 0.00001$ |
| Area of sales facilities                        | -0.03093*** | 0.004588 | -6.741 | $\leq 0.00001$ |
| Number of households                            | 0.329395*** | 0.02637  | 12.491 | $\leq 0.00001$ |
| Number of hospitals                             | -0.02204*** | 0.003447 | -6.393 | $\leq 0.00001$ |
| Concentration rate of alighting during weekends | 0.311175*** | 0.071765 | 4.336  | $\leq 0.00001$ |

-Number of observations: 267. -Adjusted  $R^2$ : 0.5330. \* Significance code: \*\*\* ( $p \leq 0.001$ ).

TABLE 9: Comparison of goodness of fit.

| Models | Linear regression model | GWR model               |
|--------|-------------------------|-------------------------|
| BOAMR  | Adjusted $R^2 = 0.5462$ | Adjusted $R^2 = 0.7668$ |
| BOPMR  | Adjusted $R^2 = 0.5074$ | Adjusted $R^2 = 0.7797$ |
| ALAMR  | Adjusted $R^2 = 0.3006$ | Adjusted $R^2 = 0.7469$ |
| ALPMR  | Adjusted $R^2 = 0.5330$ | Adjusted $R^2 = 0.8716$ |

TABLE 10: Estimation results of the GWR model (BOAMR).

| Variable                 | Min      | 1st Qu   | Median   | 3rd Qu   | Max      | Global  |
|--------------------------|----------|----------|----------|----------|----------|---------|
| Intercept                | 0.404451 | 11.24496 | 16.77398 | 24.668   | 50.66738 | 20.0598 |
| Area of sales facilities | -0.17327 | -0.07272 | -0.03163 | -0.01523 | 0.095999 | -0.0363 |
| Number of households     | -0.12268 | 0.546268 | 0.643507 | 0.772032 | 1.326634 | 0.582   |
| Number of hospitals      | -0.29164 | -0.08486 | -0.05048 | -0.03045 | 0.022822 | -0.0539 |
| Number of transfer lines | -16.2668 | -6.15996 | -3.94211 | -1.60716 | 0.642216 | -3.6377 |

-Number of observations: 267. -Adjusted  $R^2$ : 0.7668.

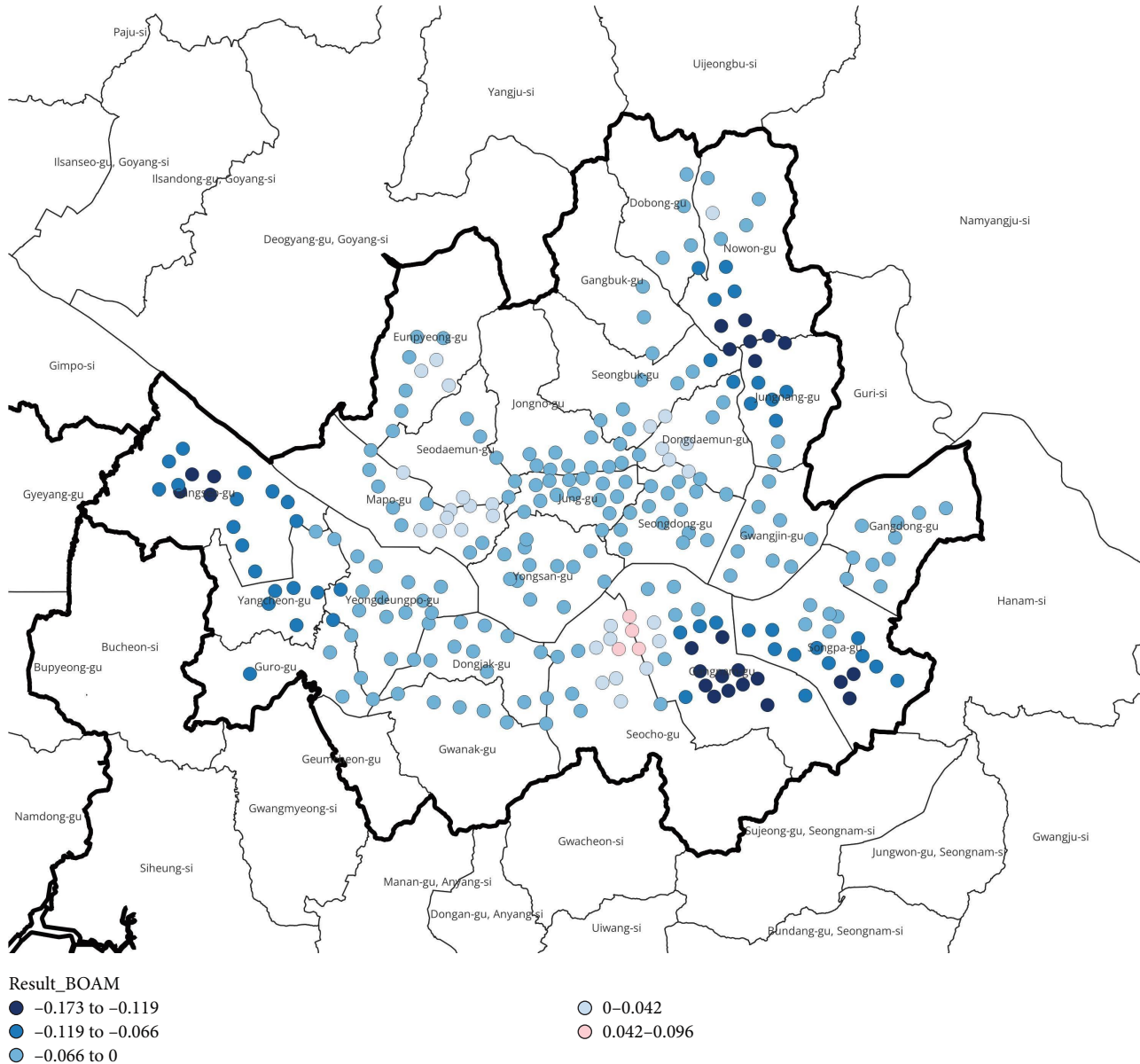


FIGURE 5: BOAMR—SALEAREA.

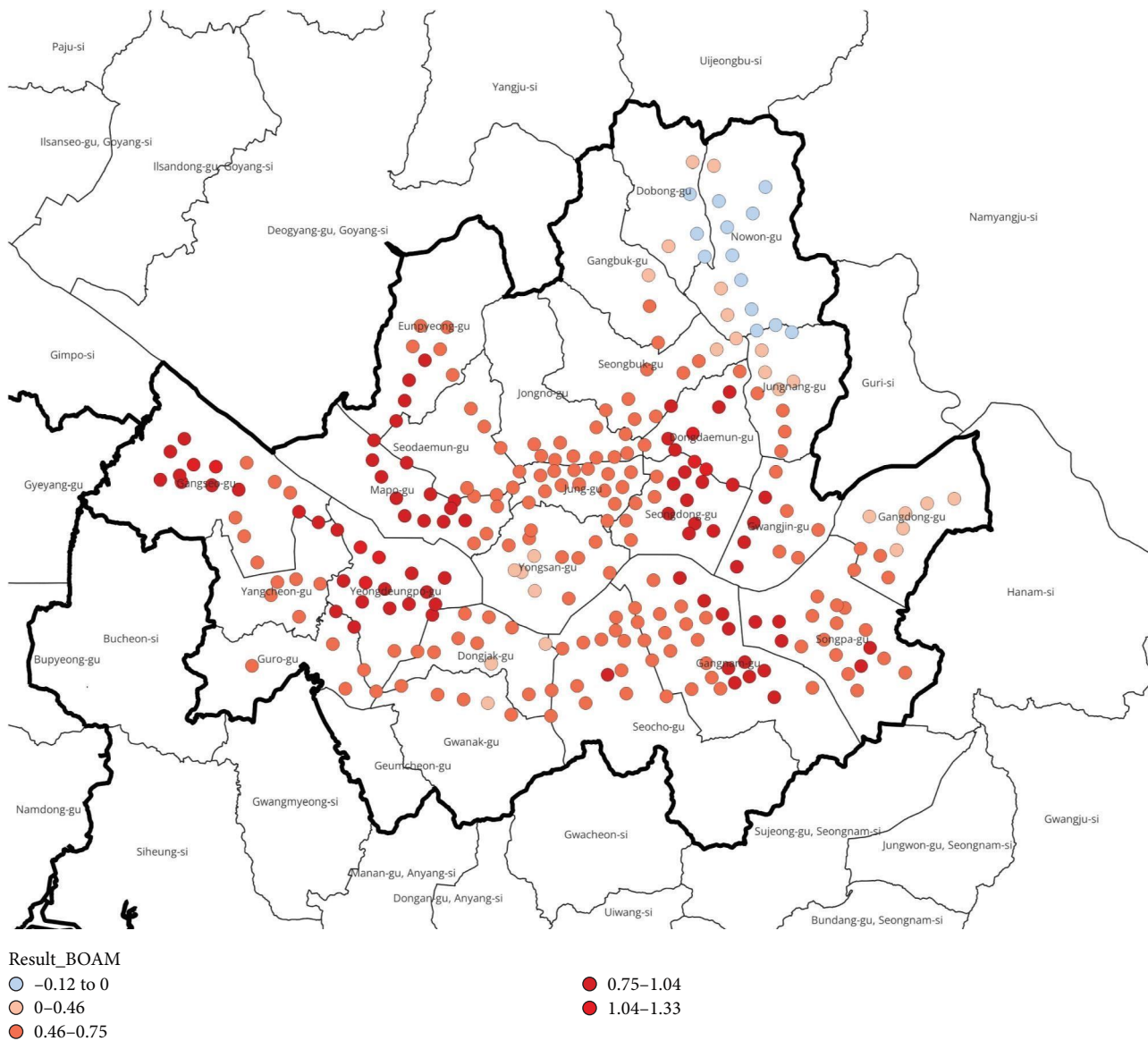


FIGURE 6: BOAMR—NOHH.

the Seoul metropolitan area, and it attracts a high level of floating population within the area at all time periods. Furthermore, from an urban planning perspective, this area has the highest connectivity in terms of railway and public transportation with cities within the Seoul metropolitan area. Therefore, changes in land use and public transit infrastructure within this area are likely to have significant spatial ripple effects. Seoul City Hall is located in the downtown region and serves as Seoul’s cultural center. Although it has a smaller residential population, it concentrates on a significant number of business facilities, similar to Gangnam. Since this area served as the central hub of Seoul before large-scale urban development occurred in the southeastern region, similar to the southeastern region, it experiences high floating population traffic compared to other regions. However, it is the area

where policies to restrict car traffic are most heavily enforced to preserve cultural heritage sites and alleviate congestion. As a result, travel demand from/to this area heavily relies on public transportation infrastructure. Moreover, geographically situated at the center of Seoul, it shares high connectivity with other cities within the Seoul metropolitan area, like the southeastern region. The northeastern and southwestern regions are characterized by a lower proportion of commercial and business facilities than other areas, with a primary focus on residential areas. While these areas also have several downtowns within them, they have a relatively higher proportion of the population commuting to and from other areas. Consequently, public transportation services in this area tend to prioritize connectivity to the major downtowns in Seoul rather than to other cities within the Seoul metropolitan area.

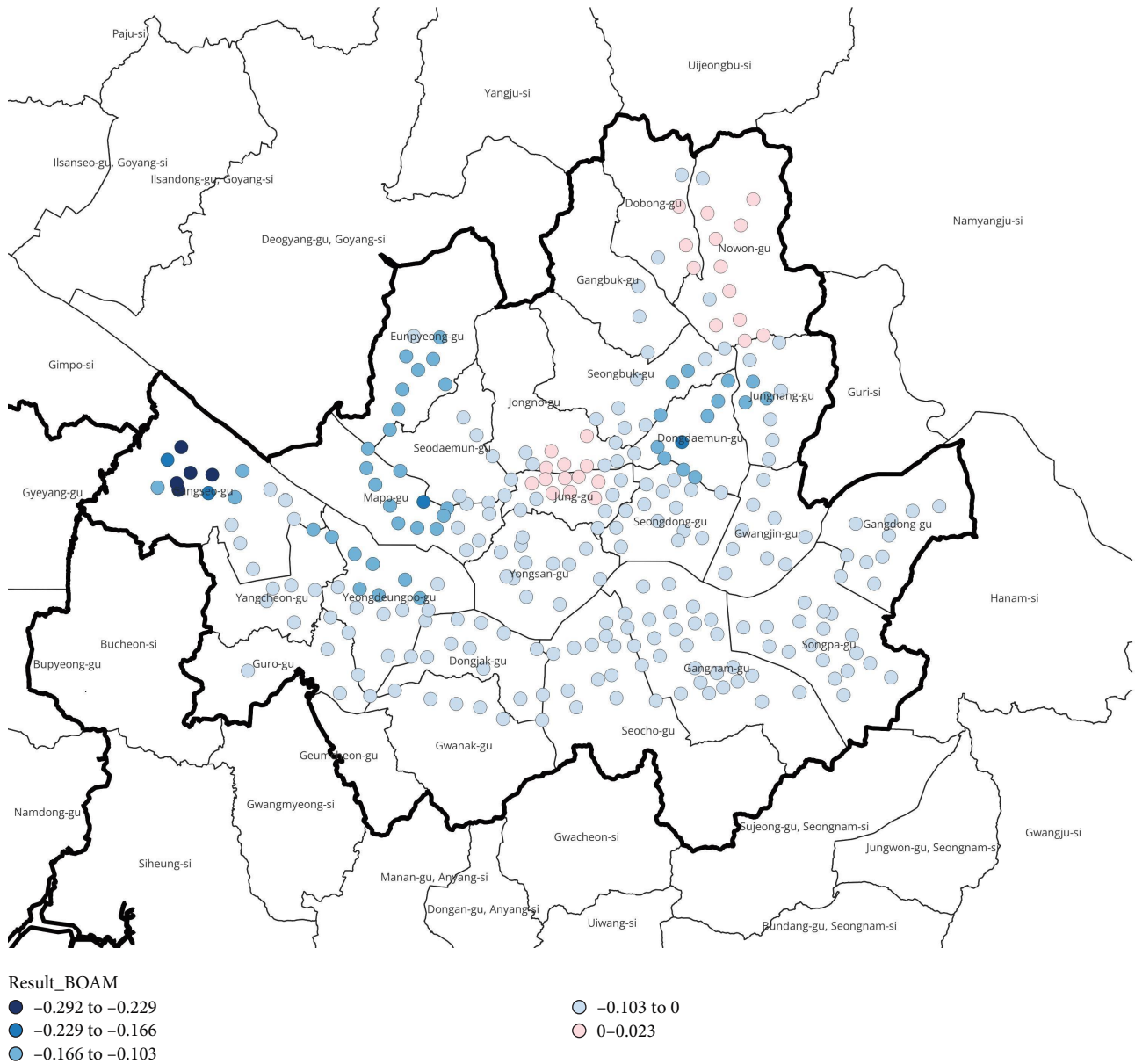


FIGURE 7: BOAMR—NOHOSP.

Finally, the northwestern regions, such as the downtown region, are not dominated by specific facilities. Rather, commercial and residential facilities are evenly distributed in this area. Figure 2 shows the categorization of each region.

4.2. SCD. This study used SCD to analyze the temporal variability of subway station boarding and alighting patterns. Using these data, trip-chain information was constructed, focusing on the initial boarding and final alighting stations for analyzing the boarding and alighting patterns by subway station. SCD covering a 1-week period, from November 11, 2019 to November 17, 2019, was used to account for the daily variability of boarding and alighting patterns. The analysis

was conducted at 267 subway stations in Seoul, South Korea, which were within a 1 km radius from the central point of each station, encompassing the entire influence zone within Seoul. Figure 3 shows the locations of the subway stations.

This study used the concentration ratio of boarding and alighting demands by time period as dependent variables to analyze the temporal variability of boarding and alighting patterns by subway station. The time periods were categorized into morning peak hours, nonpeak afternoon hours, and evening peak hours. The morning and evening peak hours covered 7:00 AM to 10:00 AM and 5:00 PM to 8:00 PM, respectively. The boarding and alighting demands were calculated based on the number of users

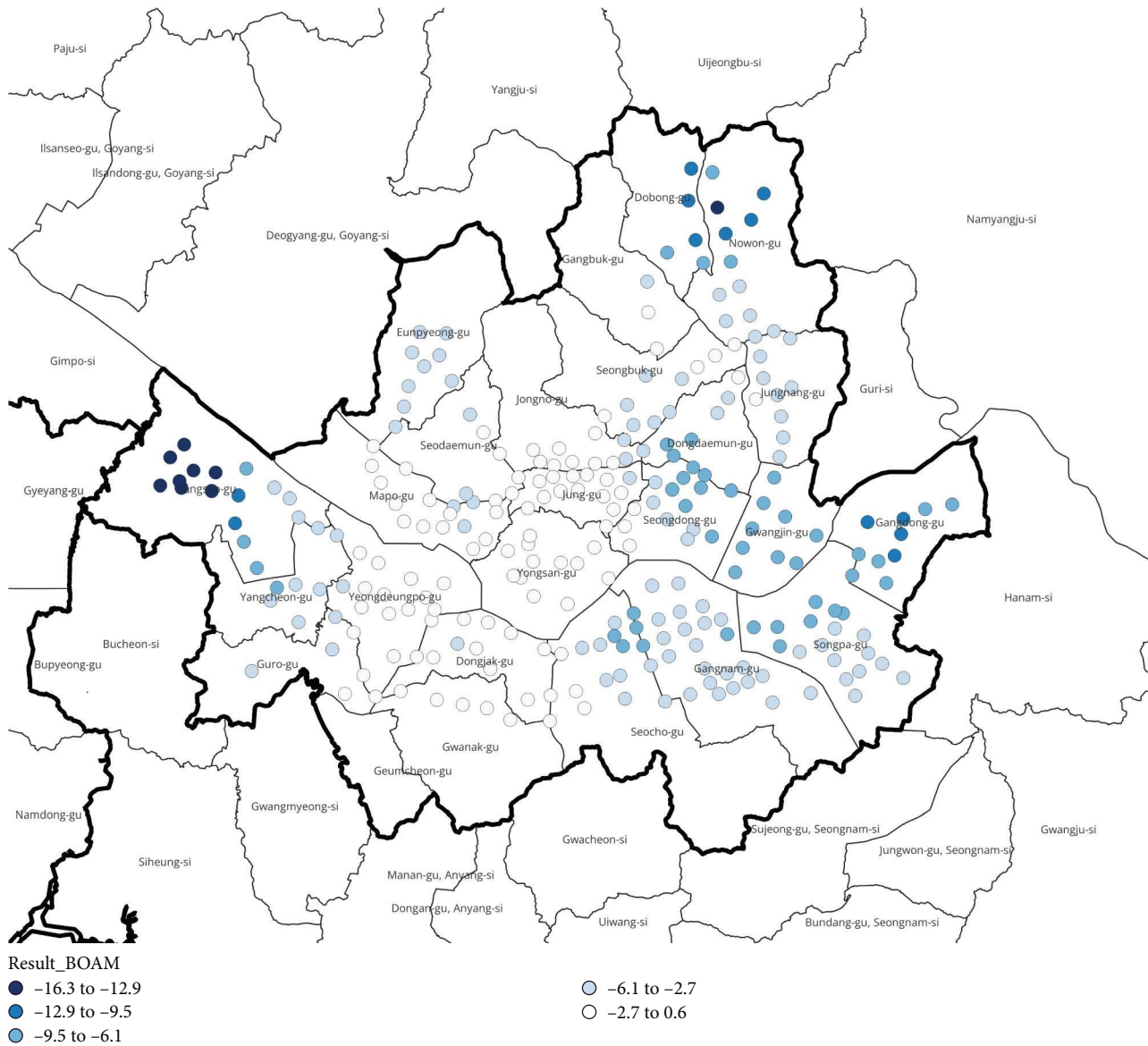


FIGURE 8: BOAMR—NOLINE.

TABLE 11: Estimation results of the GWR model (BOPMR).

| Variable                    | Min      | 1st Qu   | Median   | 3rd Qu   | Max      | Global  |
|-----------------------------|----------|----------|----------|----------|----------|---------|
| Intercept                   | 6.598517 | 22.5504  | 29.03194 | 35.4985  | 46.96715 | 27.2837 |
| Area of business facilities | -0.02232 | 0.003302 | 0.006481 | 0.009922 | 0.032487 | 0.0047  |
| Number of households        | -1.79778 | -0.63481 | -0.49031 | -0.35087 | 0.075294 | -0.4741 |
| Number of hospitals         | -0.02748 | 0.015701 | 0.038332 | 0.061451 | 0.460137 | 0.0463  |
| Number of transfer lines    | -3.44601 | 0.437954 | 2.018026 | 3.439513 | 6.922237 | 2.024   |

—Number of observations: 267. —Adjusted  $R^2$ : 0.7797.

who boarded and alighted at each subway station during their respective time periods. The concentration ratio of boarding and alighting demands is defined as the proportion of boarding and alighting demands during a specific time period compared with the total demand. In addition,

variables, such as the total boarding and alighting demands by the subway station, weekday concentration ratio, weekend concentration ratio, and operational characteristics of the subway system, including transfer routes, were calculated and used as independent variables for correlation analysis with

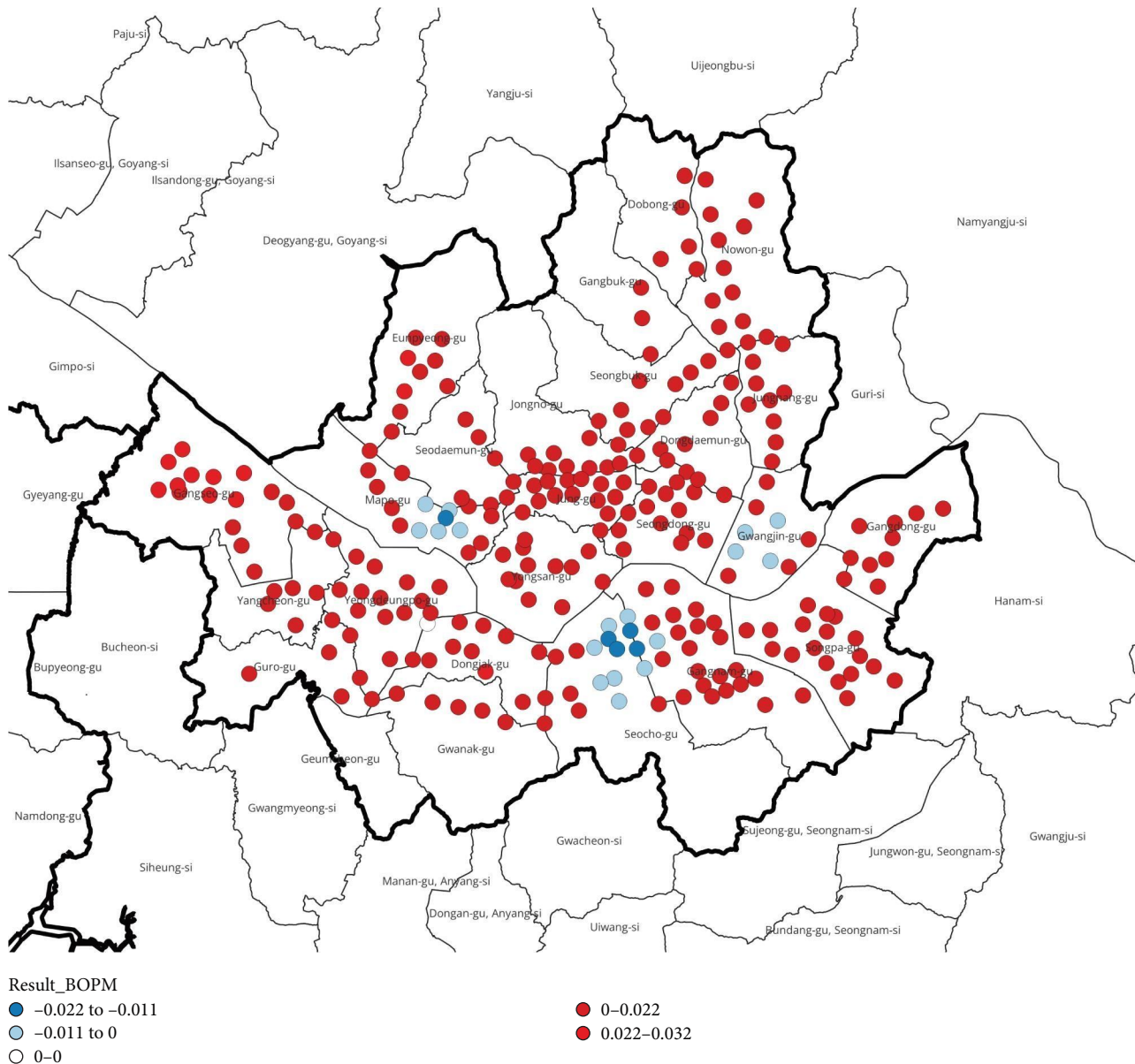


FIGURE 9: BOPMR—CBAREA.

the concentration ratio of boarding and alighting demand by time period.

**4.3. Land Use and Catchment Area Attributes.** The attributes of the influence zone of each subway station were broadly categorized into land use and sociodemographic attributes. For land-use attributes, the analysis dataset was constructed by aggregating the number and area of facilities by facility type within the influence zone of the subway station, based on point of interest (POI) data. Regarding sociodemographic attributes, estimates of the sociodemographic attributes of the influence zone were made based on data for the smallest spatial unit (administrative district) where sociodemographic attribute data were available, using the overlap ratio between the influence zone and the area of each spatial unit as the

basis. When the entire influence zone of a subway station was included within one spatial unit, the sociodemographic attributes of the influence zone were estimated based on the ratio of the area of the influence zone to that of the spatial unit. When the influence zone of a subway station overlapped with two or more spatial units, the sociodemographic attributes of the influence zone were estimated based on the overlap ratio of each spatial unit. Figure 4 illustrates the method for estimating the sociodemographic attributes of the influence zone, and Table 3 lists the final constructed independent variables.

**4.4. Review of Collected Data.** Table 3 lists the independent and dependent variables based on the data collected in this study. The boarding and alighting patterns at the stations

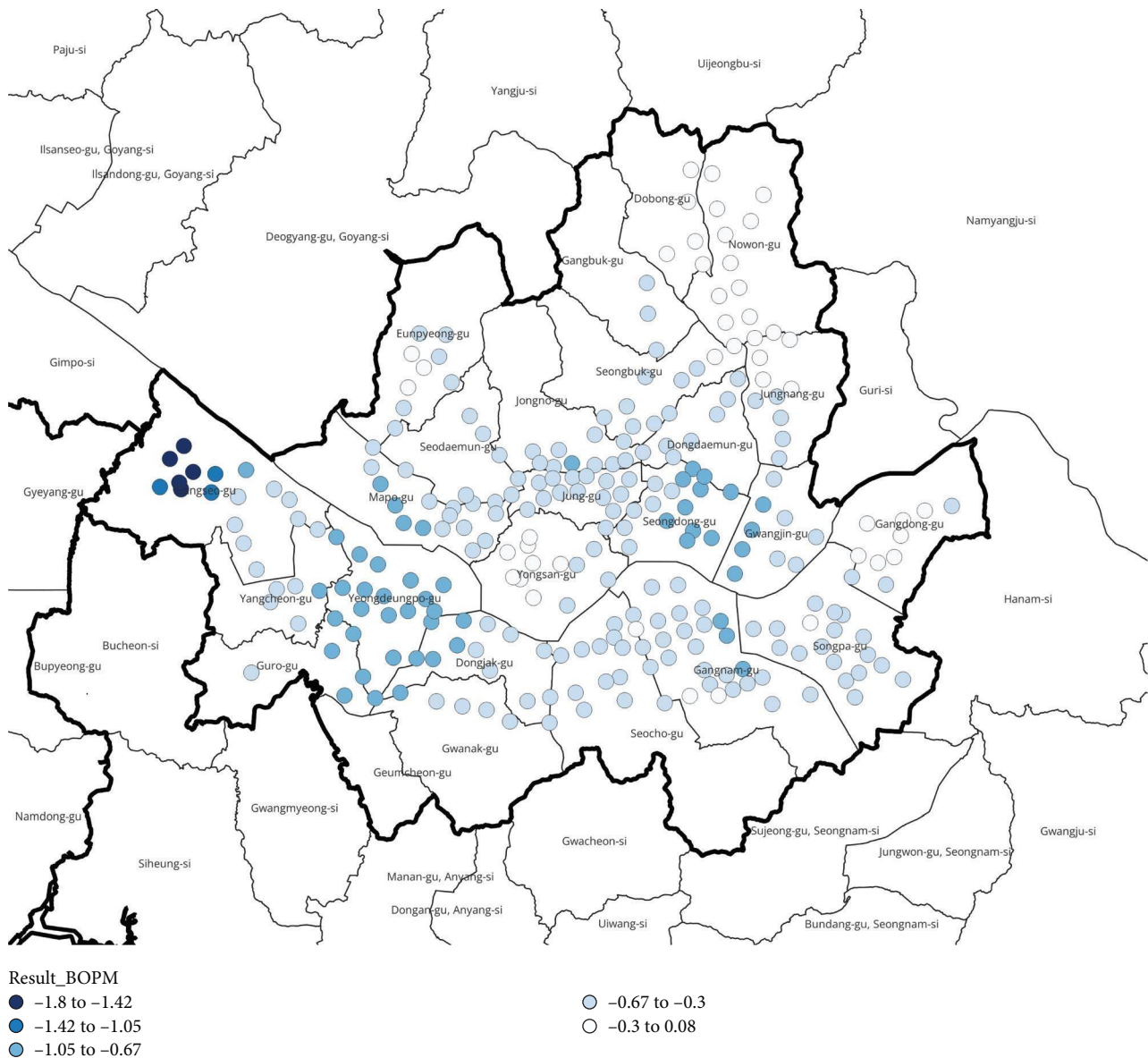


FIGURE 10: BOPMR—NOHH.

were collected based on the sociodemographic characteristics and land-use attributes of the surrounding areas. Table 4 lists the basic statistical analysis results for the collected data.

### 5. Model Estimation Results

This study aimed to elucidate the heterogeneity of the factors influencing boarding and alighting patterns based on the locational characteristics of railway stations. The results of a linear regression model that did not consider spatial heterogeneity were compared with those of a GWR model. Initially, the statistically significant variables were analyzed using a linear regression model. Subsequently, the identified variables were used to estimate the GWR model.

**5.1. Linear Regression Model.** The results of the linear regression analysis indicated that land use and socioeconomic indicators within the vicinity of the station most significantly influenced the peak boarding and alighting rates. Specifically, a wider area of retail facilities and a higher number of hospitals within the influence zone led to lower boarding rates during the morning peak hours. In addition, higher boarding rates during the morning peak hours were associated with more households within the influence zone, whereas stations with more transfer routes had lower boarding rates during this time. Regarding the afternoon peak boarding rate, stations with larger office facility areas and more hospitals in the influence zone had higher boarding rates. Moreover, higher boarding rates during the afternoon peak were associated with more households within the influence zone, and

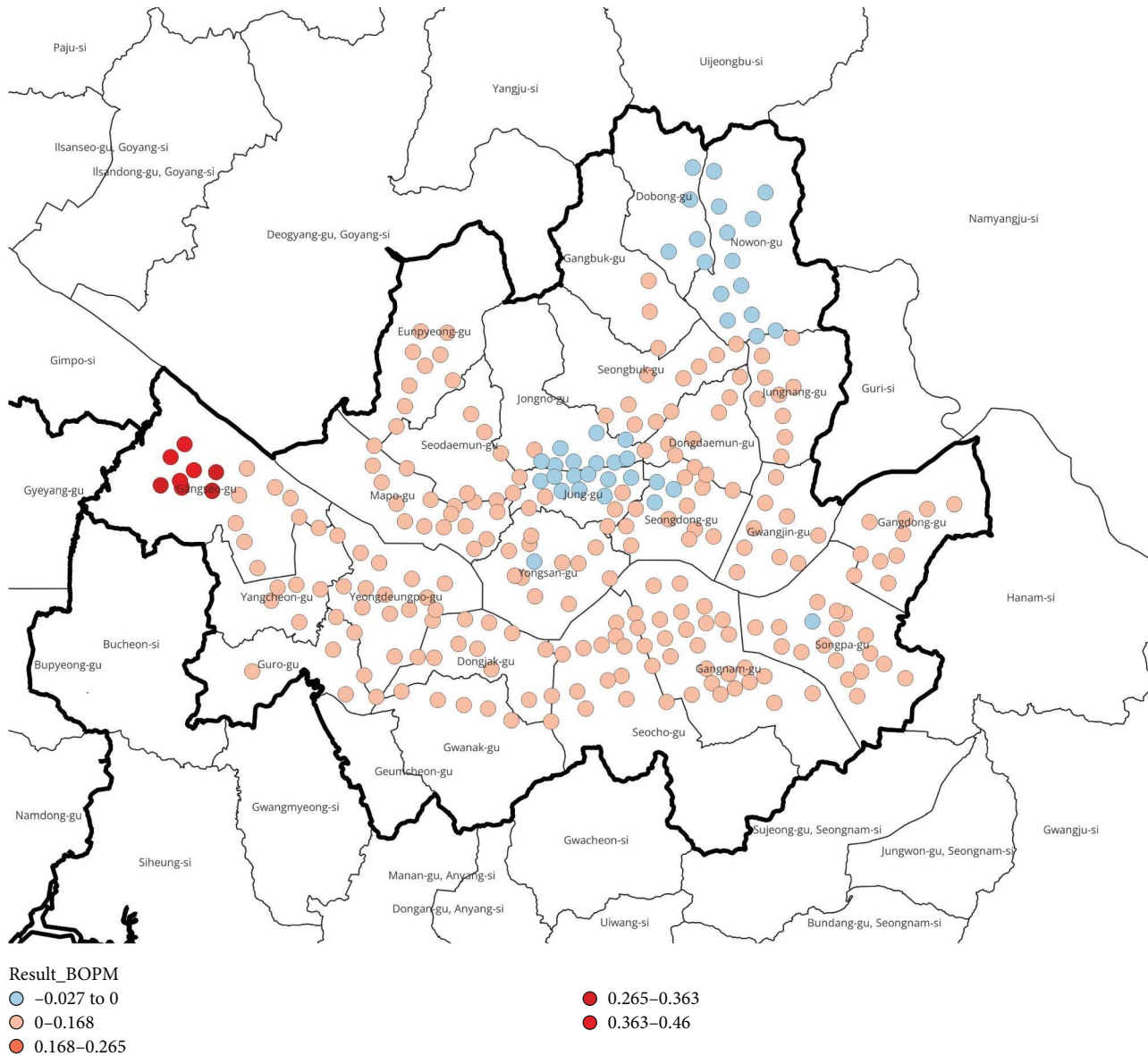


FIGURE 11: BOPMR—NOHOSP.

stations with more transfer routes had higher morning boarding rates. Regarding the morning peak alighting rate, stations with larger office facility areas and more comprehensive hospitals within the influence zone exhibited higher alighting rates. In addition, lower morning alighting rates were associated with more households within the influence zone. Stations with a higher population of individuals in their 20 and 30s within the influence zone exhibited higher morning alighting rates. Finally, for the afternoon peak alighting rate, stations with larger retail facility areas and more hospitals in the influence zone had lower alighting rates. Moreover, higher alighting rates during afternoon peaks were associated with more households within the influence zone. Stations with more passengers alighting on weekends than on weekdays showed higher alighting rates during the afternoon peak. The estimation results for BOAMR are presented

in Table 5, for BOPMR in Table 6, for ALAMR in Table 7, and for ALPMR in Table 8.

The analysis of factors influencing the boarding and alighting concentration rates during AM peak and PM peak hours revealed that the attributes affecting the peak-hour travel demand concentration rates vary depending on the time of day and boarding/alighting status. Furthermore, it was observed that even when the variables affecting the AM peak and PM peak boarding/alighting patterns are the same, their directional impact is opposite. The estimation results of the linear model presented in this study figure out factors that are not considered in existing transportation demand forecasting and optimal station scale estimation procedures. Therefore, it is deemed that including these factors in the analysis process could enhance the accuracy and reliability of the estimating and forecasting results.



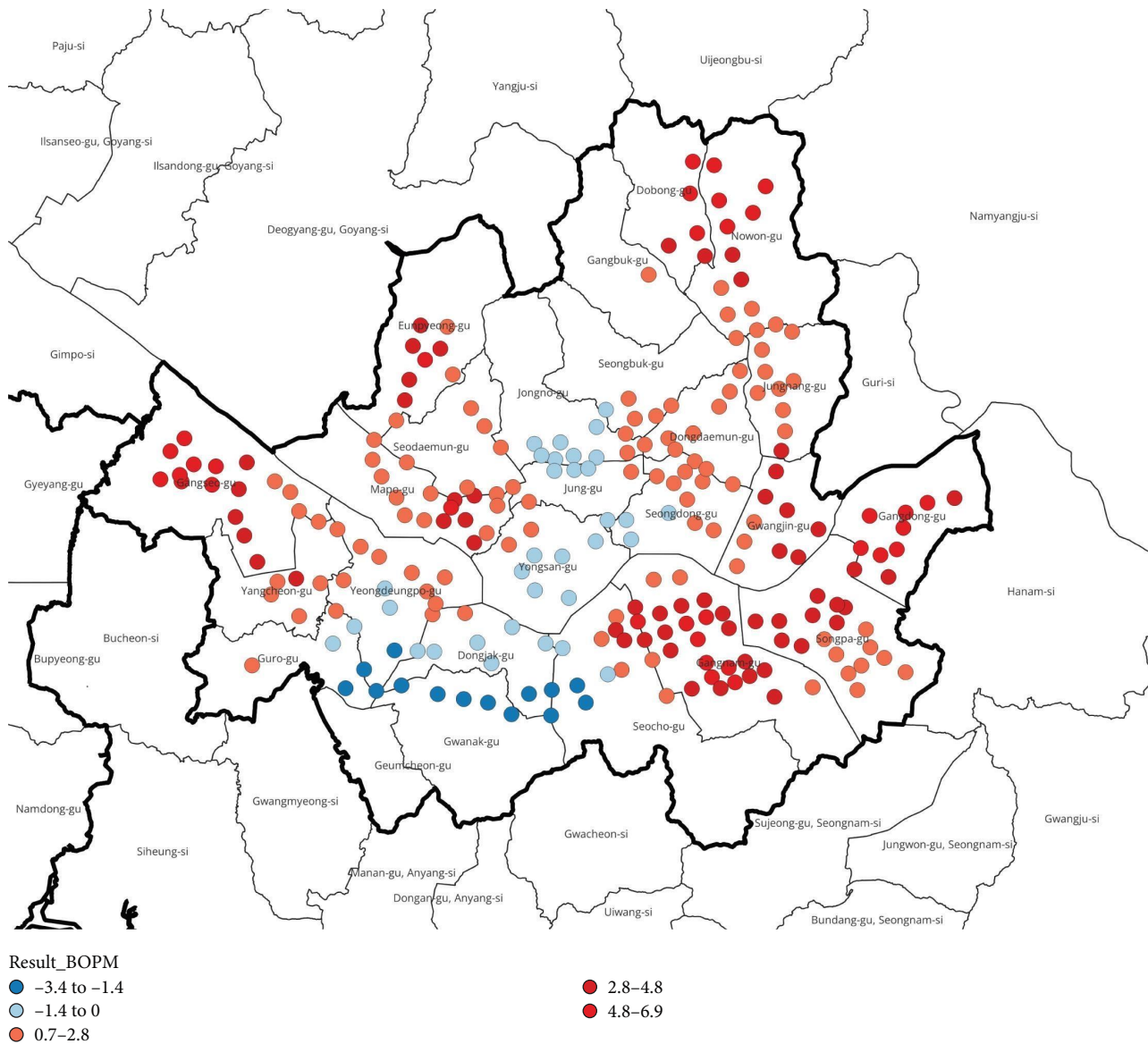


FIGURE 12: BOPMR—NOLINE.

TABLE 12: Estimation results of the GWR model (ALAMR).

| Variable                    | Min      | 1st Qu   | Median   | 3rd Qu   | Max      | Global  |
|-----------------------------|----------|----------|----------|----------|----------|---------|
| Intercept                   | 11.98035 | 22.83279 | 30.82313 | 38.78014 | 60.24013 | 32.1887 |
| Area of business facilities | -0.04742 | 0.000725 | 0.007168 | 0.013903 | 0.04432  | 0.0067  |
| Number of households        | -9.4728  | -1.44427 | -0.77413 | -0.20575 | 2.292747 | -0.9777 |
| Number of general hospitals | -5.08556 | 0.614635 | 2.038598 | 4.183085 | 20.47491 | 2.77    |
| Population of 20 and 30s    | -2.98885 | -0.11203 | 0.473943 | 1.448976 | 12.26711 | 0.7079  |

—Number of observations: 267. —Adjusted  $R^2$ : 0.7469.

5.2. GWR Model. The same variables as in the linear regression model were used to estimate the GWR model. The estimation results indicated that the GWR model exhibited an overall better model fit than the linear regression model.

In particular, for the model analyzing the afternoon peak-hour alighting concentration rate, the model fit significantly improved when the GWR model was used compared with the linear regression model. To ensure visual clarity in the

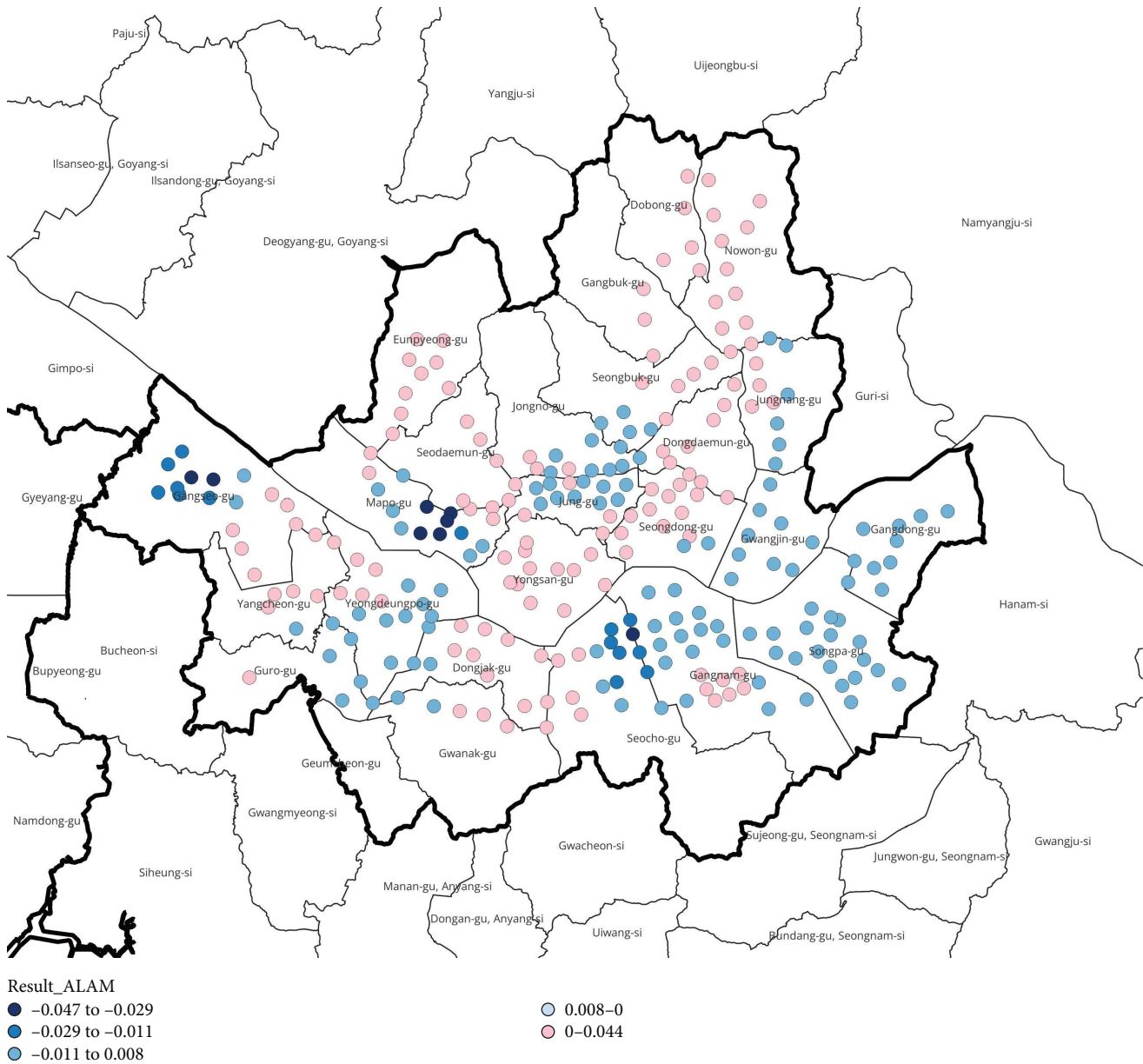


FIGURE 13: ALAMR—CBAREA.

diagrams, railway stations where the respective variables had a positive influence are marked in red, whereas those with a negative influence are marked in blue. The comparison results of goodness-of-fit between the linear regression model and the GWR model are presented in Table 9.

5.2.1. BOAMR. The estimation results of the GWR model for the morning peak hour boarding concentration rate revealed that the magnitude and direction of the impact of the independent variables on this rate varied significantly across railway stations. Specifically, in terms of the commercial facility area within the influence zone, certain railway stations in the southeastern and northeastern regions of Seoul with larger sales facility areas within the influence zone exhibited a significant reduction in the morning peak-hour boarding

concentration rate compared with other stations. Moreover, for stations located in the northeastern region of Seoul, an increase in the number of households within the influence zone resulted in the opposite trend, reducing the morning boarding concentration rate. For railway stations along the western boundary of Seoul, an increase in the number of hospitals within the influence zone resulted in a more substantial decrease in the morning peak-hour boarding concentration rate compared with those in other stations. Furthermore, for railway stations located along the western boundary of Seoul and the Gangnam area, stations with a higher number of transfer routes experienced a more significant decrease in the morning peak-hour concentration. The model estimation results can be found in Table 10, and the estimates for each station are presented in Figures 5–8.

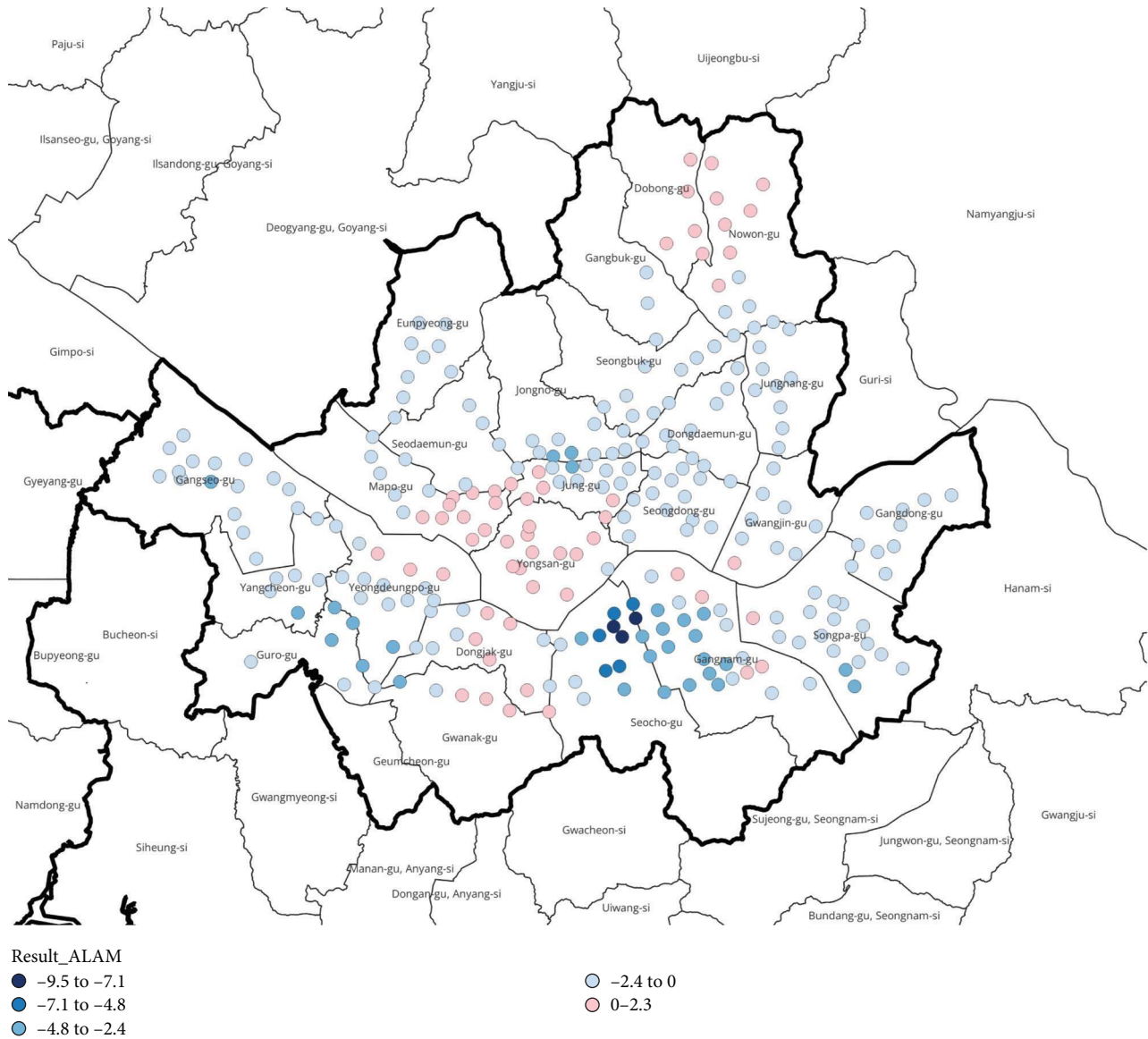


FIGURE 14: ALAMR—NOHH.

The analysis of spatial heterogeneity of factors influencing AM peak boarding concentration rate revealed that while the magnitude of impact varied across factors, the directionality of influence did not significantly differ based on station location. This suggests that during the morning peak hours, boarding primarily comprises commuters heading to work, which is mandatory, hence the consistent directionality of influence across stations. Moreover, it was observed that most stations in the southeastern and central regions exhibited distinct tendencies compared to other areas. In cases where commercial area coverage was influenced, stations in the southeastern region appeared to be more affected compared to others, while for the influence of hospital density, stations in the central region showed differing directionality compared to stations in other areas. This differentiation is likely due to the demographic structure and land use characteristics mentioned

earlier in the study area description, indicating distinctive features of these regions compared to others.

**5.2.2. BOPMR.** The afternoon peak-hour alighting concentration rate exhibited a relatively distinct pattern based on the location of railway stations compared with the morning peak-hour boarding concentration rate. Some stations in the southeastern region showed an inverse relationship, where an increase in the business facility area within the influence zone reduced the afternoon peak-hour alighting concentration rate. In addition, stations along the western boundary of Seoul demonstrated a trend in which an increase in the number of households within the influence zone significantly reduced the afternoon peak-hour alighting concentration rate. Moreover, some stations in the downtown and north-eastern regions of Seoul exhibited a contrasting trend

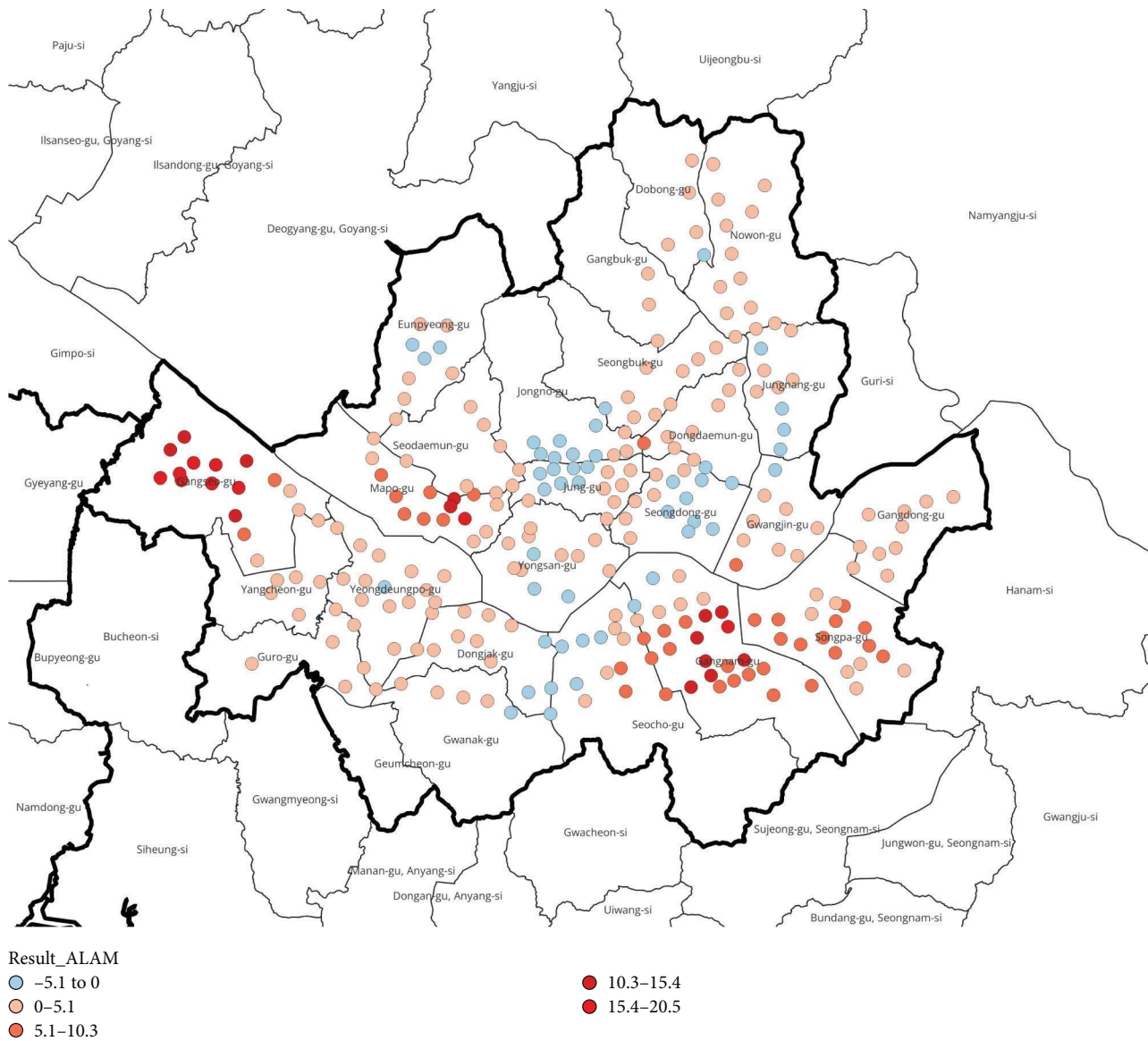


FIGURE 15: ALAMR—NOGHOSP.

compared with most stations, where an increase in the number of hospitals within the influence zone reduced the afternoon peak-hour alighting concentration rate. Stations along the western boundary exhibited a trend in which an increase in the number of hospitals within the influence zone substantially increased the afternoon peak-hour alighting concentration rate compared with those in other stations. Furthermore, stations in downtown Seoul demonstrated a contrasting trend with other stations, in which more transfer routes reduced the afternoon peak-hour concentration rate. Stations around the southeastern region of Seoul, forming a clock-like pattern, exhibited a trend in which more transfer routes significantly increased the afternoon peak-hour alighting concentration rate. The model estimation results can be found in Table 11,

and the estimates for each station are presented in Figures 9–12.

In the case of afternoon peak-hour boarding concentration, it was found that the directionality and magnitude of factors influencing boarding concentration differed depending on the station location, compared to the morning peak-hour boarding concentration. While the area of office facilities generally exhibited relatively homogeneous influence across station locations, the impact of factors such as the number of hospitals within the catchment area and the number of transfer routes varied significantly based on station location. Concerning the number of hospitals within the catchment area, it was observed that the influence on several stations located in the southwest of Seoul exhibited patterns

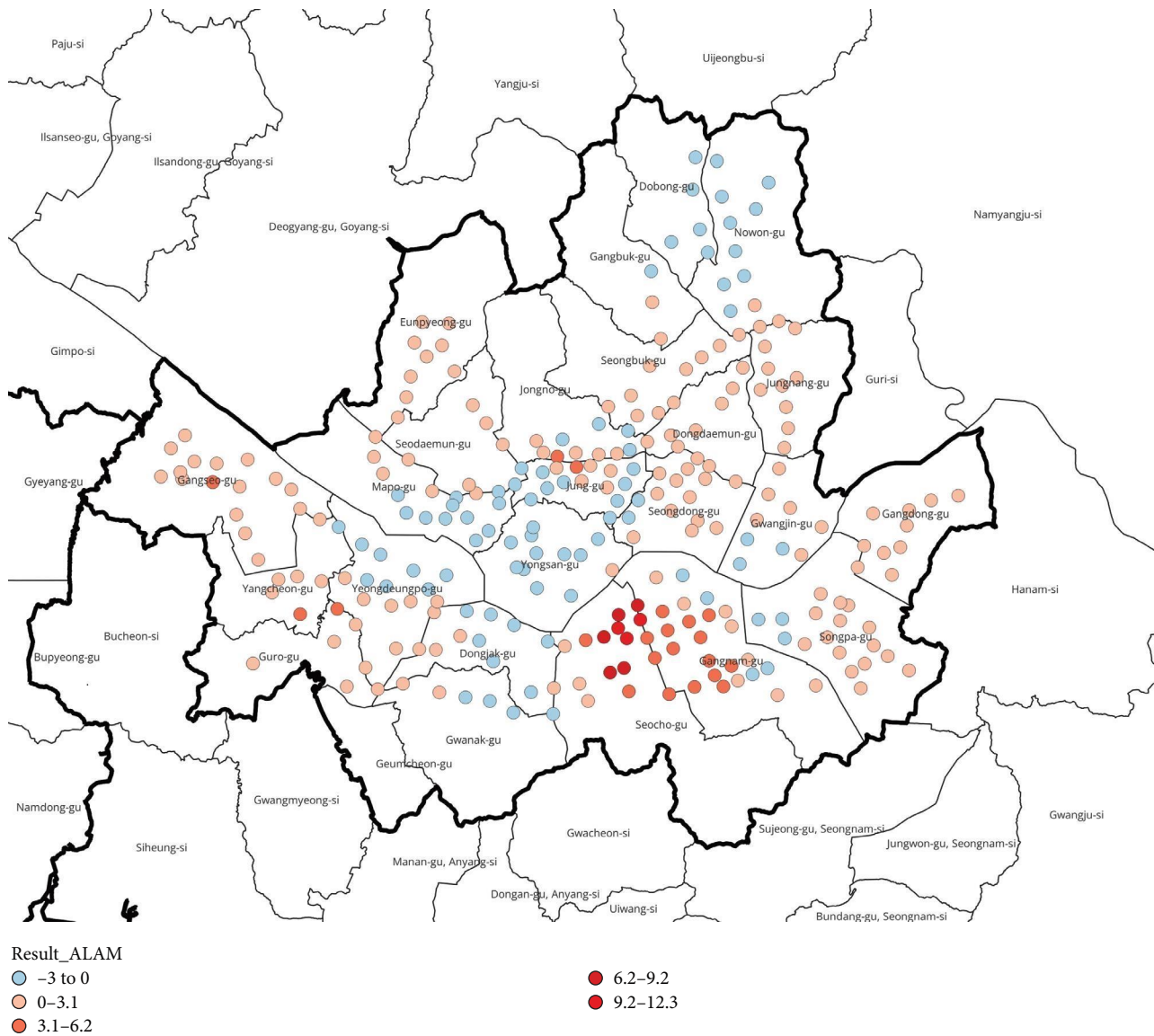


FIGURE 16: ALAMR—POP20 + POP30.

TABLE 13: Estimation results of the GWR model (ALPMR).

| Variable  | Min      | 1st Qu   | Median   | 3rd Qu   | Max      | Global  |
|---|----------|----------|----------|----------|----------|---------|
| Intercept                                       | -26.4325 | 5.845377 | 13.39001 | 20.37492 | 47.38807 | 13.3878 |
| Area of sales facilities                        | -0.33038 | -0.05704 | -0.02755 | -0.01082 | 0.131516 | -0.0309 |
| Number of households                            | -0.34061 | 0.186216 | 0.316496 | 0.448695 | 2.176658 | 0.3294  |
| Number of hospitals                             | -0.35329 | -0.04516 | -0.01152 | 0.007872 | 0.117845 | -0.022  |
| Concentration rate of alighting during weekends | -0.75961 | -0.00368 | 0.308296 | 0.572221 | 1.843474 | 0.3112  |

—Number of observations: 267. —Adjusted  $R^2$ : 0.8716.

markedly different from the influence on other stations, particularly those in the central and northeastern regions, showing contrasting patterns compared to stations in other areas. In Seoul, hospitals tend to be located in areas where a certain

level of floating population is secured, leading to the clustering of hospitals around stations. However, considering the distinct patterns exhibited by stations in the mentioned regions compared to others, it can be speculated that the

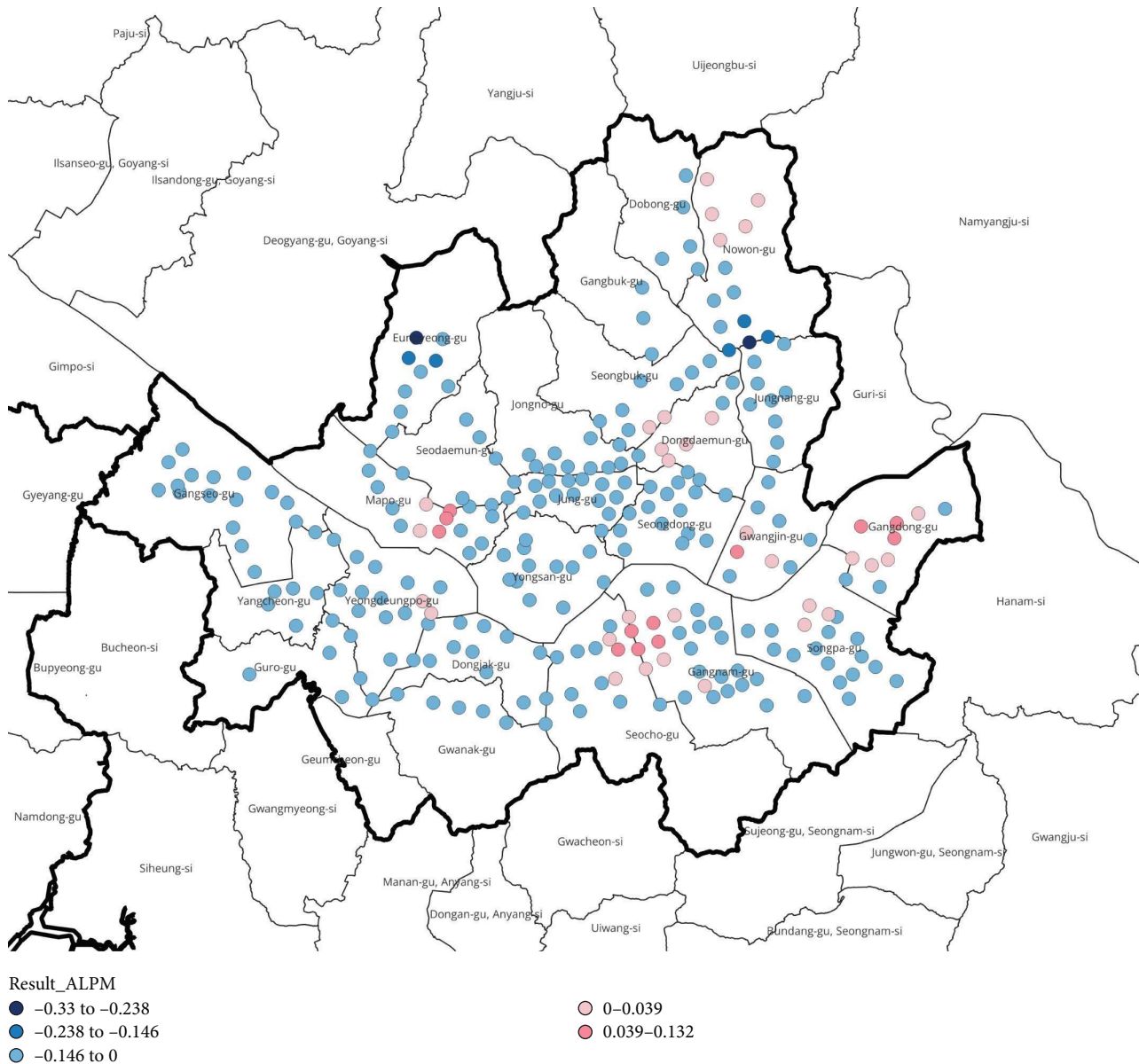


FIGURE 17: ALPMR—SALEAREA.

criteria for hospital location might differ in these areas compared to others.

5.2.3. *ALAMR*. The model estimation results revealed that the influence of the business facility area within the impact zone on the morning peak-hour alighting concentration rate varied significantly depending on the railway station location. In addition, some railway stations in the central part of Seoul exhibited the opposite trend, in which more households within the influence zone increased the morning peak-hour alighting concentration rate, in contrast to other stations. Stations in the southeast and along the western boundary demonstrated that more comprehensive hospitals within the influence zone substantially increased the morning peak-hour alighting concentration rate. Furthermore,

stations in the southeastern region exhibited a trend in which a higher concentration of young people within the influence zone significantly increased the morning peak-hour alighting concentration rate. The model estimation results can be found in Table 12, and the estimates for each station are presented in Figures 13–16.

The morning peak-hour alighting concentration also revealed a wide range of influences in terms of magnitude and directionality of various factors depending on the station location. Most alighting during the morning peak hours is related to commuting, and stations located in areas with a relatively high proportion of residential areas tend to exhibit negative effects due to the higher number of households within the catchment area. Conversely, stations located in areas with a low proportion of residential areas tend to show positive effects.

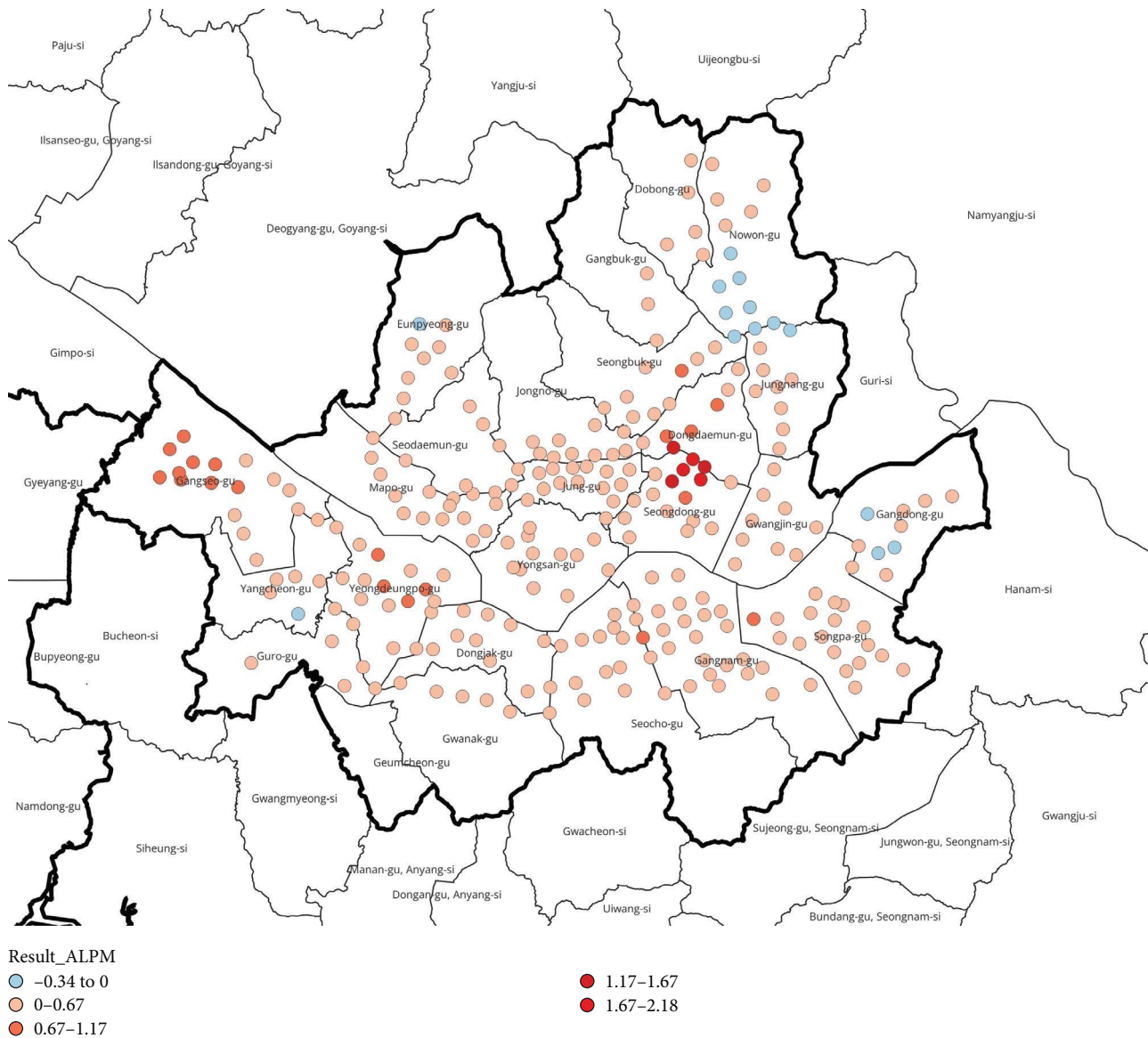


FIGURE 18: ALPMR—NOHH.

Additionally, it was found that the morning alighting concentration at stations in the southeastern region is significantly influenced by the young population, particularly as this region tends to have relatively more office facilities even within the southeastern region. This suggests that the high proportion of residential facilities for single households in this area may have contributed to this influence.

5.2.4. ALPMR. A larger retail facility area reduced the afternoon peak-hour concentration rate. However, some stations exhibited the opposite pattern. Moreover, the influence of the number of households and hospitals within the impact zone generally exhibited homogeneous patterns; however, certain exhibited contrasting trends. In addition, stations in the southeastern and downtown areas of Seoul demonstrated that stations with higher weekend alighting ratios than weekdays had lower afternoon peak-hour alighting concentration rates.

The model estimation results can be found in Table 13, and the estimates for each station are presented in Figures 17–20.

The trips involving alighting during the afternoon peak hours mostly occur for commuting purposes, so the proportion of commercial facility area tends to have predominantly negative effects, while the influence of household numbers within the catchment area tends to have positive effects, which aligns with common sense. However, in some stations where there is a positive correlation between the proportion of retail facility area within the catchment area and the afternoon alighting concentration, they are deemed to be adjacent to major downtown areas within Seoul. Moreover, considering that stations with low weekend alighting concentrations are generally located in areas with relatively fewer recreational facilities, it is reasonable to conclude that stations located in the outskirts of Seoul, especially in the northern outskirts, are negatively influenced by weekend

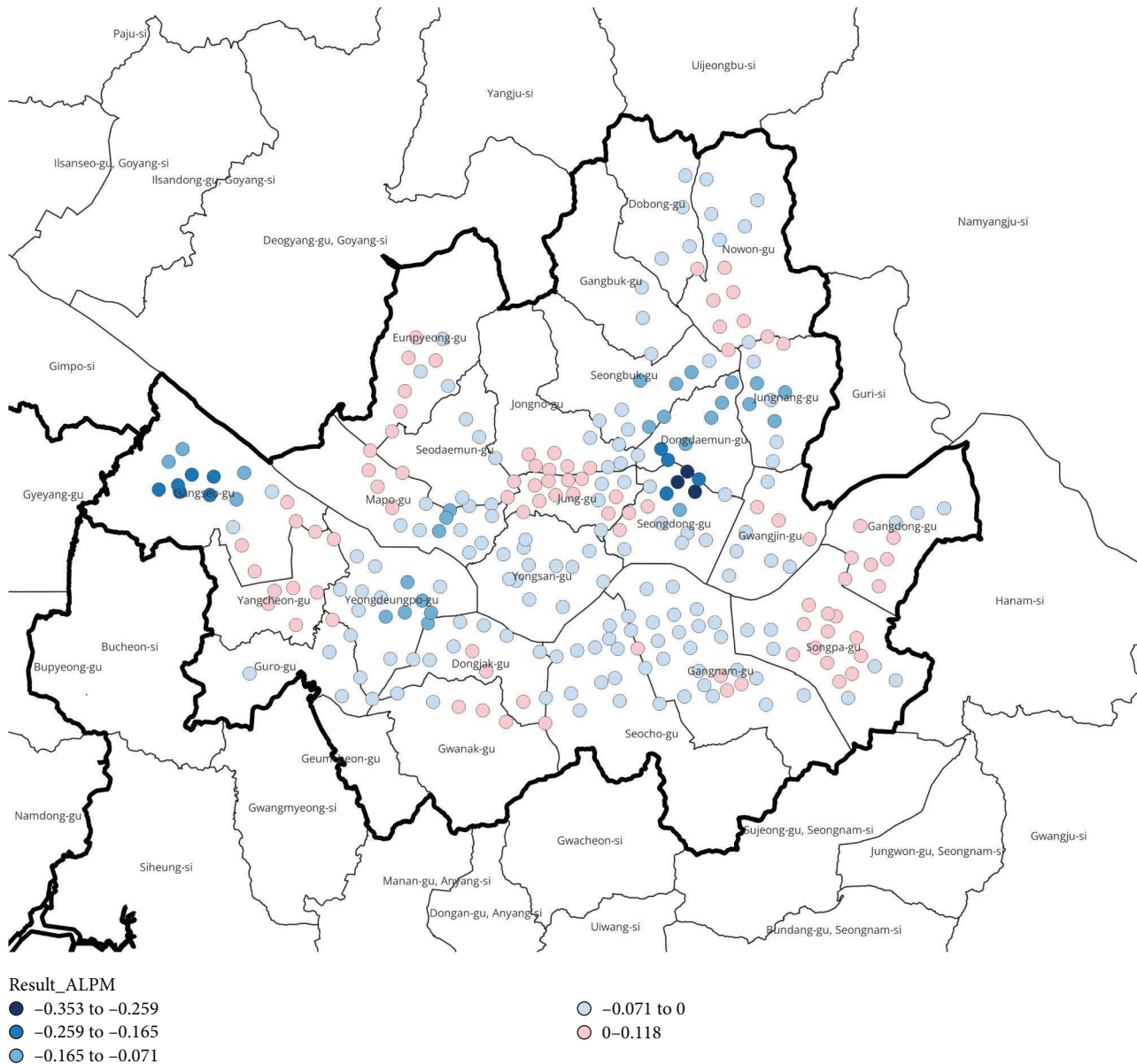


FIGURE 19: ALPMR—NOHOSP.

alighting concentrations. In this context, stations located in the southeastern region, having relatively more recreational facilities, exhibit a relatively greater impact of weekend alighting concentrations on afternoon peak-hour alighting concentrations compared to stations in other areas.

### 6. Conclusion

The expansion of metropolitan areas and the development of megacities are prominent global urban planning directions. Initially proposed as an alternative to balanced regional development, a growing negative perception of this outlook has emerged. One of the anticipated side effects of nurturing megacities is increased investment and population concentration in major cities, which can further strain urban infrastructure. This may result in user inconvenience and potentially

lead to safety incidents. Developing additional urban railways is necessary to address this problem. However, failing to consider the temporal variations in boarding and alighting patterns according to the location of railway stations in estimating the optimal station size can lead to user inconvenience and increased congestion. The factors influencing the patterns of boarding and alighting showed regional differences in broad terms, but within the same region, certain stations exhibited distinct patterns compared to other stations within that region. Moreover, the concentration of boarding and alighting at urban railway stations was observed to vary in direction depending on the time of day, even when influenced by the same variables.

Therefore, this study aimed to elucidate the influence of railway station location and impact zone characteristics on the temporal variability of boarding and alighting patterns in



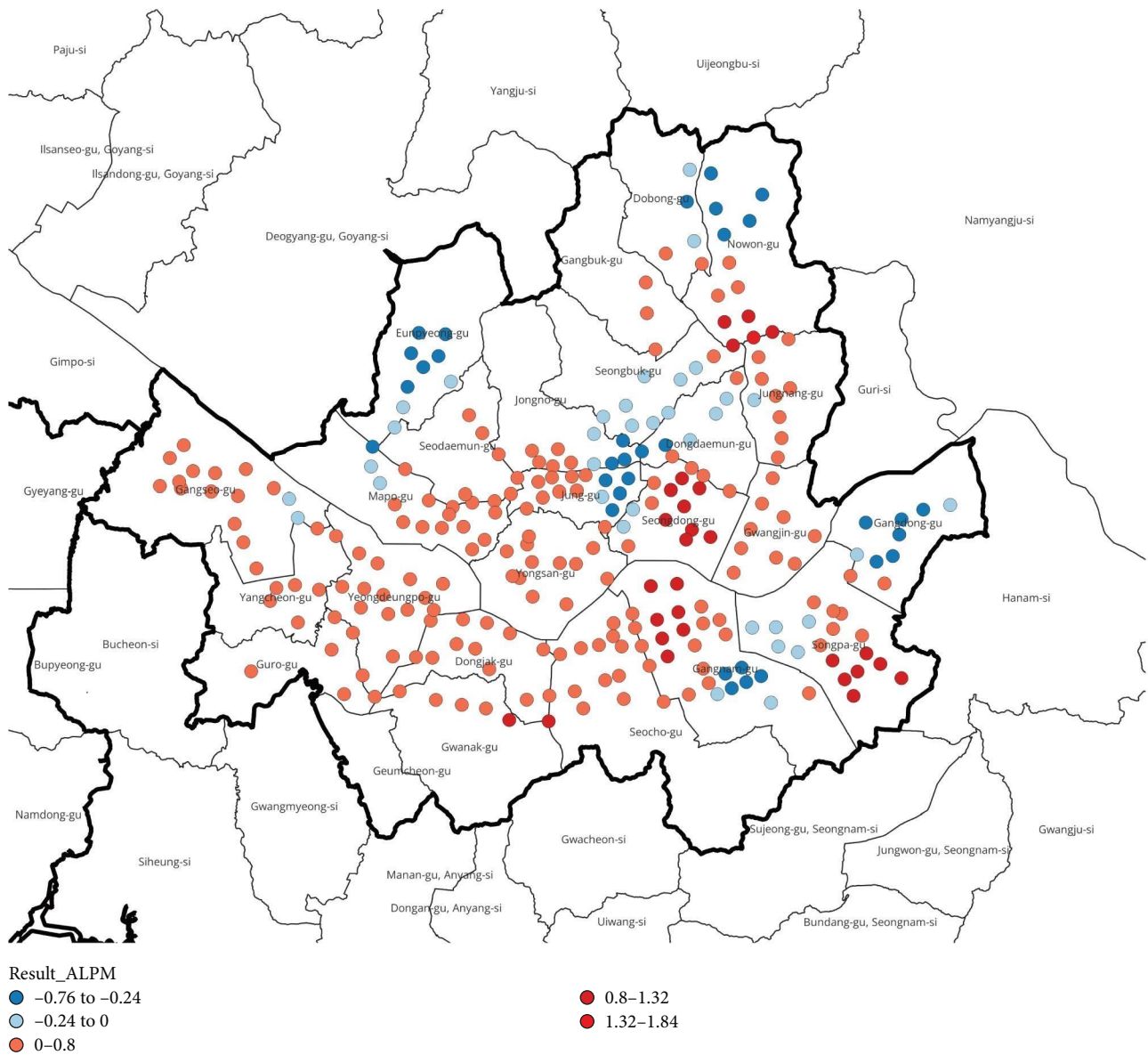


FIGURE 20: ALPMR—ALWER.

railway stations. This study focused on 276 urban railway stations in Seoul, South Korea, and individually estimated linear regression and GWR models to analyze the spatial heterogeneity of these impacts. The results revealed that stations in the southeast, downtown, and western boundaries of Seoul often exhibited different patterns compared with those in other regions. This was attributed to the distinct characteristics of each region.

These findings emphasized that even with similar attributes of the surrounding potential railway station construction sites, boarding, and alighting patterns may differ depending on the characteristics of travelers and travel patterns. Therefore, incorporating this spatial heterogeneity due to railway station location in estimating railway station size, training/station operation plans, and other design/operation activities before

the actual construction is crucial. This should be done through quantitative and qualitative analyses, considering the growth process of the city and its identity.

There are various factors that affect the optimal scale of urban railway stations. For example, urban railways operational attributes such as timetable, headway during peak and nonpeak hours, land use attributes around urban railway stations, including the number and floor area of facilities by land use type, and socioeconomic attributes of residents within the catchment area or stations, such as age, gender, and income level, can affect the optimal railway stations. This study focused on peak and nonpeak boarding/alighting concentration rates among the various factors and designated them as the dependent variables. In this context, this work examined the temporal variability and spatial heterogeneity of the

impact of several demographic and land use-related attributes surrounding railway stations on-boarding/alighting concentration rates. The findings of this study are expected to contribute to estimating the optimal scale of railway stations. Moreover, it is anticipated to collaborate with the expansion of railway in South Korea, as a part of ongoing transport policy such as “Achieving Commute Times in 30-min range,” contributing to congestion alleviation and reduction in waiting times.

However, as mentioned earlier, owing to limitations in data collection, this study utilized only the demographic characteristics, some land-use-related attributes of the surrounding areas, and some infrastructure-related characteristics as independent variables. In this context, it seems that this study has limitations in that it did not consider a wider range of variables that could influence the boarding/alighting patterns of urban railway stations. These limitations stem primarily from the nature of the data primarily utilized in this study. The public transportation SCD, which was central to this study, is distributed after encrypting the personal information of actual travelers due to personal information protection issues. Therefore, it is not possible to consider the characteristics of actual travelers at each station. Additionally, there were many obstacles to obtaining more diverse and granular land use data within the catchment area of the urban railway that led to the exclusion of such data from the study, which can be considered one of the limitations. In the future, if SCD data, POI data, and similar datasets are more widely utilized and issues such as encryption of personal information are resolved, applying such data to the methodology presented in this study could significantly enhance the explanatory power of the model. If we could consider the characteristics of actual travelers, we could analyze how the travel patterns vary depending on the stations where actual travelers with similar attributes board and alight. This analysis would be highly beneficial for establishing appropriate station sizes. Moreover, if we could classify land use around stations into more detailed categories and build datasets containing attributes of buildings and floor areas by land use type, we could expect to observe more pronounced temporal variability and spatial heterogeneity than the macroscopic research results presented in this study.

### Data Availability

The datasets used and/or analyzed during the current study are available from the corresponding author.

### Conflicts of Interest

The authors declare that they have no conflicts of interest.

### Acknowledgments

This research was supported by a grant from the R&D program (Development of Digital Operation Platform Technology for Enhancing KTX speed, PK2402C4) of the Korea Railroad Research Institute.

## References

- [1] Y. Ahn, T. Kowada, H. Tsukaguchi, and U. Vandebona, “Estimation of passenger flow for planning and management of railway stations,” *Transportation Research Procedia*, vol. 25, pp. 315–330, 2017.
- [2] H. Yang and Y. Tang, “Managing rail transit peak-hour congestion with a fare-reward scheme,” *Transportation Research Part B: Methodological*, vol. 110, pp. 122–136, 2018.
- [3] J. Teng, W. Pan, and C. Zhang, “Quantitative modeling of congestion in metro station based on passenger time perceptions,” *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2674, no. 5, pp. 270–281, 2020.
- [4] X. Li, G. K. Sinniah, R. Li, and X. Li, “Correlation between land use pattern and urban rail ridership based on bicycle-sharing trajectory,” *ISPRS International Journal of Geo-Information*, vol. 11, no. 12, Article ID 589, 2022.
- [5] M. Li, X. Zhou, Y. Wang, L. Jia, and M. An, “Modelling cascade dynamics of passenger flow congestion in urban rail transit network induced by train delay,” *Alexandria Engineering Journal*, vol. 61, no. 11, pp. 8797–8807, 2022.
- [6] Y. Wang, M. Li, J. Zhou, and H. Zheng, “Sudden passenger flow characteristics and congestion control based on intelligent urban rail transit network,” *Neural Computing and Applications*, vol. 34, pp. 6615–6624, 2022.
- [7] H. Jiao, S. Huang, and Y. Zhou, “Understanding the land use function of station areas based on spatiotemporal similarity in rail transit ridership: a case study in Shanghai, China,” *Journal of Transport Geography*, vol. 109, Article ID 103568, 2023.
- [8] H. Liang, Z. Tian, R. Liu, G. Sun, and F. He, “Managing peak-hour congestion in urban rail transit with the sub-train price adjustment strategy,” *Mathematical Problems in Engineering*, vol. 2022, Article ID 3723630, 14 pages, 2022.
- [9] L. Yu, Q. Chen, and K. Chen, “Deviation of peak hours for urban rail transit stations: a case study in Xi’an, China,” *Sustainability*, vol. 11, no. 10, Article ID 2733, 2019.
- [10] Z. Huang, M. Zhang, and X. Liu, “Estimating light-rail transit peak-hour boarding based on accessibility at station and route levels in Wuhan, China,” *Transportation Planning and Technology*, vol. 40, no. 5, pp. 624–639, 2017.
- [11] L. Li, Y. Wang, G. Zhong, J. Zhang, and B. Ran, “Short-to-medium term passenger flow forecasting for metro stations using a hybrid model,” *KSCE Journal of Civil Engineering*, vol. 22, pp. 1937–1945, 2018.
- [12] L. Wang, X. Yan, and Y. Zhang, “Coordinative passenger flow control during peak hours in urban rail stations,” in *Proceedings of the 15th COTA International Conference of Transportation Professionals*, pp. 1153–1164, ASCE, Beijing, China, July 2015.
- [13] G. Gulhan, H. Ceylan, M. Özuysal, and H. Ceylan, “Impact of utility-based accessibility measures on urban public transportation planning: a case study of Denizli, Turkey,” *Cities*, vol. 32, pp. 102–112, 2013.
- [14] J. Gutiérrez, O. D. Cardozo, and J. C. García-Palomares, “Transit ridership forecasting at station level: an approach based on distance-decay weighted regression,” *Journal of Transport Geography*, vol. 19, no. 6, pp. 1081–1092, 2011.
- [15] Z. Yang, X. Li, Y. Guo, and X. Qian, “Understanding active transportation accessibility’s impacts on polycentric and monocentric cities’ housing price,” *Research in Transportation Economics*, vol. 98, Article ID 101282, 2023.
- [16] X. Tang, R. Bi, and Z. Wang, “Spatial analysis of moving-vehicle crashes and fixed-object crashes based on multi-scale

- geographically weighted regression,” *Accident Analysis & Prevention*, vol. 189, Article ID 107123, 2023.
- [17] J. Tang, F. Gao, C. Han, X. Cen, and Z. Li, “Uncovering the spatially heterogeneous effects of shared mobility on public transit and taxi,” *Journal of Transport Geography*, vol. 95, Article ID 103134, 2021.
- [18] Y.-C. Chiou, R.-C. Jou, and C.-H. Yang, “Factors affecting public transportation usage rate: geographically weighted regression,” *Transportation Research Part A: Policy and Practice*, vol. 78, pp. 161–177, 2015.
- [19] H. Y. Saputra and I. F. Radam, “Accessibility model of BRT stop locations using Geographically Weighted regression (GWR): a case study in Banjarmasin, Indonesia,” *International Journal of Transportation Science and Technology*, vol. 12, no. 3, pp. 779–792, 2023.
- [20] A. Jaber, L. A. Baker, and B. Csonka, “The influence of public transportation stops on bike-sharing destination trips: spatial analysis of Budapest City,” *Future Transportation*, vol. 2, no. 3, pp. 688–697, 2022.
- [21] A. Özbil Torun, K. Göçer, D. Yeşiltepe, and G. Argın, “Understanding the role of urban form in explaining transportation and recreational walking among children in a logistic GWR model: a spatial analysis in Istanbul, Turkey,” *Journal of Transport Geography*, vol. 82, Article ID 102617, 2020.
- [22] O. D. Cardozo, J. C. García-Palomares, and J. Gutiérrez, “Application of geographically weighted regression to the direct forecasting of transit ridership at station-level,” *Applied Geography*, vol. 34, pp. 548–558, 2012.
- [23] A. Hadayeghi, A. S. Shalaby, and B. N. Persaud, “Development of planning level transportation safety tools using geographically weighted poisson regression,” *Accident Analysis & Prevention*, vol. 42, no. 2, pp. 676–688, 2010.
- [24] M. F. Dziauddin, N. Powe, and S. Alvanides, “Estimating the effects of light rail transit (LRT) system on residential property values using geographically weighted regression (GWR),” *Applied Spatial Analysis and Policy*, vol. 8, pp. 1–25, 2015.