512

Analyses of Factors for Busy Streets with Space Configuration Indicators

KANEDA TOSHIYUKI 1, OTA AKIRA 2, ZHANG YUFAN 1, & KOBAYASHI JUNYA 3,

1 NAGOYA INSTITUTE OF TECHNOLOGY, JAPAN, 2 TOKYO CITY UNIVERSITY, JAPAN, SHIMIZU CORPORATION, JAPAN

ABSTRACT

Space syntax provides various indicators of the urban form on spatial characteristics and is becoming increasingly important in terms of accountability in policy analysis and design and planning assessments, as the number of pedestrians is now considered an effective policy variable for the creation of pedestrian spaces. The authors' Analysis of Factor of Busy Streets (AoFBS) is as follows: first, for the objective variable of pedestrian count data, a candidate explanatory variable including three factors (transportation accessibility factor, facility volume factor, and urban form indicator) is calculated by using multiple regression analysis and other methods, and the next step is to search for a factor model for data fitting and to compare the intensity of the factor variables based on the estimated parameters. It is characterized by the ordering of the strength of the three factors mentioned above. The authors consider our AoFBS to be a transitional period in three aspects: (1) the spread of cell phone location data, (2) spatial representation and SS indicators, and (3) extension to nonlinear models. In this paper, we focus on the Kanayama case, in which Agent Analysis was applied using a linear model, and the Nagoya CBD case, in which Segment Angular Analysis was applied using a log-log model, and also discuss especially on these three aspects.

KEYWORDS

Space Syntax, Pedestrian, Linear Model, Log-Log Model, Nagoya

1 INTRODUCTION -- SPACE SYNTAX AS ACCOUNTABILITY TOOL

The spatial maldistribution of pedestrians on street networks in middle-scale urban spaces, such as congestion and vacant streets, has long been considered to be related to public gaze, prospect, or “visibility,” but it has not been the subject of explicit quantitative investigation. The advancement of Space Syntax studies (SS) since Bill Hillier et.al.’s theory of Natural Movement in 1993 has made this possible. In today’s 21st century Japan, where the creation of pedestrian spaces with a “human dimension” is desired in urban centers, quantitative analysis of the spatial distribution of the number of pedestrians has great significance. This has been made possible by the widespread use of cell phone location data.

In this context, Space Syntax provides various indicators of urban form that characterize spatial modalities, and is becoming increasingly important in terms of accountability in policy analysis and planning and design assessments as the number of pedestrians is now regarded as a variable for policy effect. Because, if the number of outdoor pedestrians means just traffic movement, only two factors have traditionally been dominant (especially at the macro scale): the transportation accessibility factor as traffic distance resistance, and the facility volume factor, such as the size and agglomeration of facilities that attract visitors, thus the third factor, urban form factors such as street pattern, site and building shape, are the "latecomers".
The authors' Analysis of Factor of Busy Streets (AoFBS) is as follows. First, candidate explanatory variables including the three factors mentioned above are prepared for the objective variable of pedestrian count data. Next, a factor model for data fitting is explored using multiple regression analysis. Finally, a comparative study of the strength of the factor variables is conducted based on the estimated parameters. The unique feature of our approach is the ordering of the intensities of the three factors.

In the past, Desyllas et.al (2003) adopted the connectivity factor in SS theory as an indicator of urban form in addition to the transportation accessibility factor and the facility volume factors for the central London area, known as the first analytical case study to introduce SS indicators, but without ordering the intensity among the factors.

The authors have mainly focused on Nagoya, and since confirming the effect of space configuration indicators (Ota, Nakano, Kaneda, 2008) in 2008, we have attempted to analyze over 10 cases in Nagoya CBD, Sakae South, Sakae District, Osu District, Meieki District, Sakae Underground Mall, Meieki Underground Mall, Kanayama District, and central Okazaki City, and have confirmed the Space Configuration factors in most of them.

We have so far applied multiple regression analysis is the analysis of a linear model. While the linear model has the advantage of being simple and clear, and the factor structure is easy to extract and read, it has the disadvantage that the number of pedestrians, the number of counts, does not necessarily follow a normal distribution, and it is difficult to use for forecasting and other purposes.

The authors consider our factor analysis to be a transitional period from three aspects.

- Diffusion of cell phone location data
- Spatial representation and SS indicators
- Extension to nonlinear models

In this paper, the authors will discuss two case studies in Nagoya, which are still in the process of being researched, with a particular focus on the three issues.

2 RESEARCH DESIGN

2.1 Analysis of Factor of Busy Streets (AoFBS)

2.1.1 What is AoFBS?

We have already mentioned that of most interest to us is the comparison of the intensity of the following three factors.

(a) Space Configuration factor
(b) Facility volume (land use) factor
(c) Transportation accessibility factor

Hillier's Natural Movement Theory discusses (a) space configuration, and most indicators derived from space syntax, such as integration value, fall under this category. Hillier refers to (b) facility volume factor as the attractor factor, and it incorporates indicators of land use intensity such as the floor area for commercial use on the first floor and the floor-area ratio for each use. It is important to note that this is the OD end point (trip-end) for commuting trips. (c) transportation accessibility was developed to reflect the transportation structure of Japanese urban centers, where commuting trips rely heavily on mass transit systems such as subways and railroads, and are typically indexed by distance from transportation hubs such as subway and rail stations. However, the distance measurement is not limited to metric, but also includes visual step depth and other space syntax-derived measurements.

The authors' Analysis of Factor of Busy Streets (AoFBS) is as follows. First, candidate explanatory variables including the three factors mentioned above are prepared for the objective variable of pedestrian count data. Next, a factor model for data fitting is explored using multiple regression analysis. Finally, a comparative study of the strength of the factor variables is conducted based on the estimated parameters. The unique feature of this method is the ordering of the intensities of the three factors.

2.1.2 Analytical framework in linear model

The authors use multiple regression analysis for our AoFBS because the calculating procedure is simple, making it easy to gain insight into the factors, but an analytical framework is also meaningful for the transparency (elimination of arbitrariness) of the insight into the factors.
Our AoFBS may be loosely divided into three steps, as shown below.
STEP1: Examination of variables
STEP2: Model selection
STEP3: Exploration of the factor structure

The examination of variables in STEP1 is related to the preparation of the objective variable and candidate explanatory variables. Basic statistics such as mean and standard deviation as well as frequency distribution are mastered for all variables, and correlation coefficients and scatter plots between two variables are created (correlation matrix and scatter plot matrix). At this stage, suggestions for model form types are obtained.

Model selection in STEP2 is the process of selecting one model form that adopts a collection of explanatory variables by specifying the type of model form and applying Multiple Regression Analysis (MRA), such as stepwise regression. In Ota's framework (2016), the model form with the lowest Akaike Information Criterion (AIC) value is selected for each number of adopted variables, and the model form with the lowest AIC value is selected among them. When the model forms are arranged by the number of adopted variables, the AIC values show a single peak, the standardized partial regression coefficients of the adopted variables are order-preserving, and there is no inversion in the signs of the coefficients. The data set's characteristic structure is said to be "fine." Although it is necessary to specify the procedure in terms of reproducibility, it is not necessary to be overly concerned with this method; for example, selecting one explanatory variable from each of the three factor groups can be taken as the next best measure.

Furthermore, in order to avoid multicollinearity, there is a principle in multiple regression analysis that prohibits pairs of explanatory variables with high correlation coefficients. The authors' factor analysis in STEP3 is understood in the same method of interpretation as the causal path diagram of covariance structure analysis, and the model form is often interpreted by implicitly assuming zero internal correlation between each factor. It is desirable to define the VIF requirements for each candidate variable's adoption or rejection, which are sufficient to examine the factor ranking. Methods for estimating model forms that allow for internal correlation, such as PLS (Partial Least Square) regression analysis, should be explored as well.

The purpose of the factor structure in STEP3 is to comprehensively judge the goodness of fit of the model form based on the correlation between the predicted and actual values and the signs of the partial regression coefficients of the adopted explanatory variables, and to consider the factors based on the rank order of the strength (absolute value) of the standard partial regression coefficient in the explanatory factors. The standard partial regression coefficient is (marginal increase in the z-score of y) / (marginal increase in the z-score of x_i) and is understood as the slope between standardized variables, i.e., the partial differential coefficient. As a result, the standard partial regression coefficient is used in the case of linear model, but in the case of nonlinear model, the partial derivative of y depends on each x_i, so the case where each x_i is an average value may be illustrated.

2.1.3 Related works

As previously stated, the authors were motivated by the work of Desyllas et.al. in 2003, which was the first Space Syntax study to investigate the number of pedestrians in central London and analyze the determinants using multiple regression analysis.

Ota's analysis of prosperity factors used a linear model, expanding the objective variables to include not only pedestrian numbers but also rents and land prices (Ota 2016). One study looked at the factors that influenced pedestrian numbers in Nagoya's Sakae-south district in 2005 and 2011, using gate-counts pedestrian survey records as objective variables (Ota, Meziani, Kaneda 2015). This was at a time when mobile phone location data services were expensive and not widely available.

The results of another linear regression analysis using rents and land prices as the objective variables in the case study area of central Nagoya in 1935 and 1965 show that the space configuration indicators, as a separate factor from the facility volume factors and transportation accessibility factors, is effective in the central Nagoya area (Ota and Kaneda 2020). Although the hedonic approach was known for the linear regression equation of rents and land prices, and a high correlation between first-floor retail rents and pedestrian volumes was noted in the real estate business, the investigation of the causal path structure of the space configuration index → number of pedestrians → first-floor retail rents had to wait for the spread of big data.

Sharmin and Kamruzzaman (2018) did a meta-analysis of 1194 publications dealing with the Space Syntax index as a factor of pedestrian movement as a study of interest to the authors that emerged with the spread of big data in recent years. A detailed comparison of the coefficients was conducted for 14 major papers.

2.2 Three issues of today

2.2.1 Mobile phone location data

In the past, surveys have been used to collect data on pedestrian counts. In the gate count survey, the surveyor counts the number of pedestrians crossing a street section line during the measurement time. The Encounter Survey measures the number of pedestrians overtaking and passing while the surveyor is moving, and it allows for a smaller number of surveyors. This is known as the survey method and is used by Hillier in his discussion of natural movement (Hillier et.al 1993).

On the other hand, the three major mobile phone carriers in Japan have recently begun commercial and off-the-shelf services of mobile phone location data in urban areas, and the possibility of exploiting such data is expanding. The method's major advantages are its timeliness and availability of data, which more than compensate for the current lack of accuracy and monitor ratio. The point data is a chain of data that includes the location of each anonymized monitor's point in the city center at each time of day, as well as the speed and direction of the point, allowing the researcher to analyze the minute behavior of pedestrians. The case of the Saitama Station area is known as a report on the application of point data. Factor searches for both pedestrian counts and pedestrian speeds have been reported (Shimizu, Nishi, Kishimoto, 2019).

The small-areal data is aggregated data for each specified area obtained by the aggregation site for anonymization processing. Not only attributes of the monitors (gender and age) but also estimated attributes based on the aggregation of past history (residents, workers, visitors) can be used to obtain detailed aggregated data for periods ranging from one day to one year, and even by one hour time zone.

The authors use the KDDI Location Analyzer (KLA) data for the Kanayama case (Section 3.6) and the Nagoya CBD case (Chapter 4), described below. This is a website service that aggregates GPS location information obtained from anonymous monitors of au smartphones of KDDI, according to user queries, and outputs the estimations. While performance details are not disclosed, current mobile phone location point data such as KLA is considered to have a minimum interval of 5 minutes, a location accuracy (standard deviation of error) of 10-20 meters, and a standard aggregation service of 15 minutes. This more than makes up for the lack of location accuracy and monitor ratio (several tenths). In particular, in the analysis of primary movement lines, it is thought that the number of movement lines is aggregated by applying a process such as recounting the consecutive points of the same user ID within the measurement area to a single point. In this process, the estimated number of movement lines is not affected by the length of the street.

2.2.2 Space representations and Space Syntax indicators

To begin with, each aspect of urban form in our AoFBS is defined by its preconditioned spatial representation. And each indicator derived from the Space Syntax is defined by the spatial representation that is the precondition for its analysis. For relatively small urban spaces, Visual Graph Analysis (VGA) is often applied in cases of visibility surface or walkable surface as areal surfaces. For larger spaces, the network representation is based on street link (segment) data converted from GIS use, and in this case, Segment-Angular Analysis is expected to be used. In the authors' factor analysis, Space Syntax indicators are used as candidate variables, but the integration value (IV) have been adopted in most of the cases.

In addition, as a Space Syntax indicator in areal spatial representation, Agent analysis (AA) is a simulation that makes use of Turner and Penn's vision-driven agent EVA (Penn and Turner 2001, Turner and Penn 2002). They reported a 0.75 correlation with measurements in a department store simulation, while Omer and Kaplan (2017) recently examined AA in terms of MRA or AA. Other features such as the agent's direction of travel and the modification of the agent mechanism such as OD-Weighting have just recently begun and have not yet been thoroughly studied (Ferguson et.al 2012, Uyar et.al 2017, Kaneda et.al 2019). Chapter 3 describes the authors' detailed examination of AA.

2.2.3 Extension to nonlinear model

While the linear model is often used to determine the factor structure, the log-linear model and log-log model are also aimed at prediction, but the disadvantage is that the factor structure is difficult to read. In this paper, the Kanayama case (Chapter 3) uses a linear model, and the Nagoya CBD case (Chapter 4) uses a log-log model. Thus, we are going to try an extension of our analytic framework for log-log model.

The purpose of the factor structure in STEP3 in 2.1.2 is to comprehensively judge the goodness of fit of the model based on the correlation between the predicted and actual values and the signs of the partial regression coefficients of the adopted explanatory variables, and to consider the rank order of the strength (absolute value) of the standard partial regression coefficient in the explanatory factors. The standard partial regression coefficient is (marginal increase in the z-score of y) / (marginal increase in the z-score of x_i), and is understood as the slope between standardized variables, i.e., the partial differential coefficient. As a result, the standard partial regression coefficient is used in the case of linear model, but in the case of log-log model, the partial derivative of y depends on each x_i, so the case where each x_i is an average value is illustrated.

2.3 Structure of the paper

Recognizing that our AoFBS is transitional in above mentioned three aspects, this paper will describe two case analyses that we have worked on or are working on (Figure 1).

Case 1 Kanayama: The Kanayama area is a commercial district with a subcenter position as a transfer station in Nagoya City. We conducted an encounter survey in 2017 and conducted our factor analysis that included both Visibility Graph Analysis (VGA) and Agent Analysis (AA) of

Space Syntax (SS) indicators as urban form indicators under an aerial representation. We addressed this issue and discussed its character as factor variables in the two SS indicators. A follow-up analysis will also be conducted using the number of the movement lines in cell phone location data (KLA) in 2019.

Case 2 Nagoya CBD: The whole Nagoya CBD is a 635 hectares area that includes two central transfer stations, Meieki and Sakae. We conducted a factor analysis using the number of the movement lines based on KLA based on network representation using GIS street link (segment) data. In Chapter 4, we use a log-log model, in which the model equations are log-transformed and then subjected to multiple regression analysis. Segment Angular Analysis was applied as the urban form indicator.

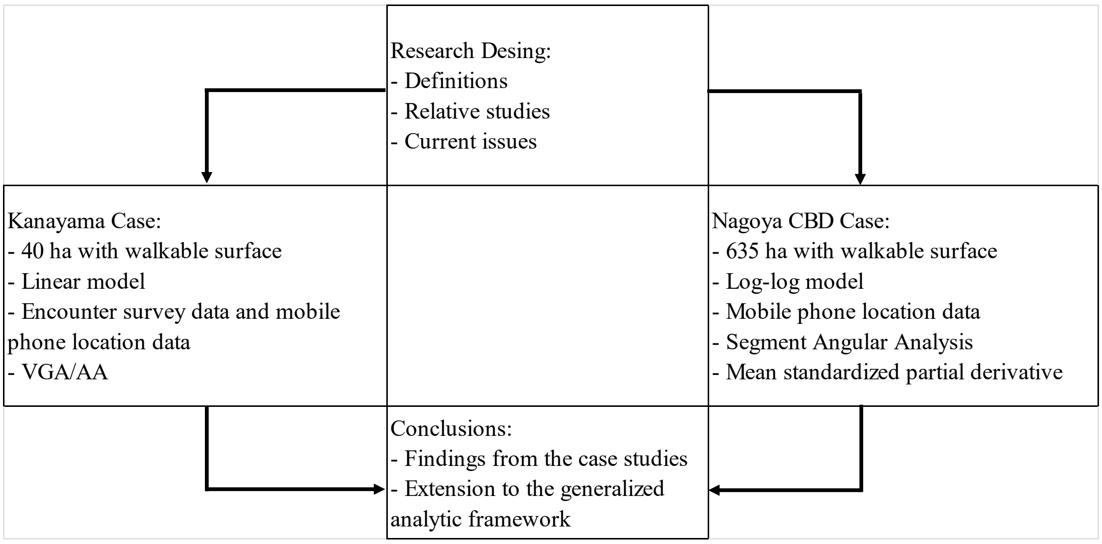


Figure 1: Structure of this paper

3 KANAYAMA CASE: FACTOR ANALYSES USING LINEAR MODEL WITH AGENT ANALYSIS INDICATORS

3.1 Kanayama Case

In Chapter 3, we conducted a multiple regression analysis using a visual graph analysis (VGA) indicator and an agent analysis (AA) indicator to explain the number of pedestrians on weekdays and holidays in the encounter survey data in the Kanayama area of Nagoya City, Aichi Prefecture. The potential of the AA indicator as a factor index of "liveliness" is investigated by comparing the model with the AA indicator and the model with the VGA indicator. A linear model is used in this chapter's multiple regression analysis.

The case study area in the Kanayama district in Nagoya City, Aichi Prefecture, which is based on the City of Nagoya's "Kanayama Station Area Community Development Plan," covers an area of approximately 40 hectares. Kanayama Station, located in the Kanayama area, serves as a terminal station for approximately 440,000 passengers per day (2016) on five train lines. Residential and office/school uses each account for more than 30% of the total, followed by commercial facilities, resulting in a mix of residential, business, and commercial uses in the area.

3.2 Encounter survey data

In this study, the encounter survey was used to determine the number of all-day pedestrians as an index of liveliness. Encounter surveys were conducted on September 23, 2017 (Saturday, a holiday), September 27, 2017 (Wednesday, a weekday, until 5:00 p.m. due to rain), and October 4, 2017 (Wednesday, a weekday, from 6:00 p.m.). The survey method consisted of counting the total number of pedestrians on each street by oncoming and overtaking their flows as the surveyor followed a predetermined path, as well as counting the number of pedestrians on each street by sight at intersections. Each survey started at 20 minutes past the hour and lasted approximately 20 minutes.

Since the number of pedestrians obtained by the encounter survey is influenced by the street length, the all-day pedestrian count (persons/m) per meter of street length was used as the analysis's objective variable (Y). We obtained 178 pedestrian count data for all days, aggregated by street(sidewalk) segment, for both weekdays (Y1) and holidays (Y2).

3.3 Linear regression model analysis with Visibility Graph Analysis (VGA) indicator

3.3.1 Candidate factor variables

In this study, candidate variables from the three candidate factor groupings of transportation accessibility, facility volume (land use), and space configuration are used.
(X1) distance from station entrances is the first potential factor (group), transportation accessibility. This is the shortest possible path distance between the appropriate street midpoint and the eight station entrances.

We prepared five variables as the second group of candidate factors for the facility volume: (X2) floor-area ratio of commercial facilities, (X3) floor-area ratio of offices and schools, (X4) floor-area ratio of accommodation facilities, (X5) floor-area ratio of cultural facilities, and (X6) floor-area ratio of first-floor commercial buildings.

While the transportation accessibility and facility volume factors indicate the OD of walking trips and departure and arrival points, the third candidate factor group, the space configuration factor group, is an index of the urban form. As part of the space configuration index, two VGA indicators, (X7) visible area and (X8) global integration value (GIV) of the whole area, were created. The walking space to be analyzed for making the VGA indicators includes both the crosswalks and the corridors inside the stations, in addition to the sidewalks and area outside the building sites. In this case, the roadway, the interior of adjacent buildings, and the railroad are all regarded as non-walkable spaces. The visible area is the total number of points visible from a given point. The GIV indicates the depth of the spatial connection. If a point has a high value for the GIV, it means it is shallow in depth compared to its surroundings and has a high centrality in the space. For the both indicators, DepthmapX software was used and a 1 m square grid was set up for measurement. Figure 2 shows the spatial distribution of GIVs.

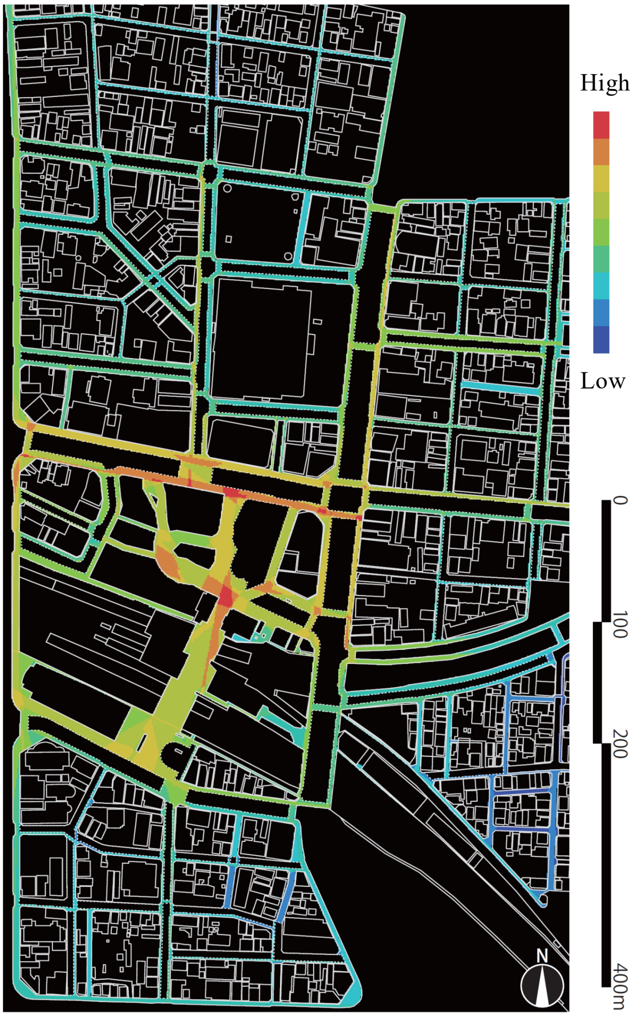


Figure 2: Spatial distributions of the global integration values (GIVs)

For edge effect, a grid of the similar size to this case was used to examine the differences when a 10% buffer was provided for the Mean Depth and GIV indicates. Since the effect of the buffer on the edges was only about 5%, it was not considered to have a significant impact on the conclusions, and therefore was not taken into account.

3.3.2 Examination of the selected multiple regression model

The stepwise variable increase and decrease method was used to select the model that minimized the Akaike Information Criterion (AIC) for the multiple regression analysis. To avoid multicollinearity, the VIF between the candidate factor variables was less than two, and all candidate factor variables were used for the analysis.

As a result, a three-variable model for weekdays and a six-variable model for holidays were chosen (Table.1). The weekday model had an AIC of 492.942 and a multiple correlation coefficient of 0.593 (coefficient of determination 0.352), whereas the holiday model had an AIC of 595.593 and a multiple correlation coefficient of 0.651 (coefficient of determination 0.424). When these two models are compared, it is possible to summarize that the holiday model has a higher value of the multiple correlation coefficient (coefficient of determination).

Table.1: Result of analysis of factor (multiple regression model selection) for the number of pedestrians on weekday and holiday (all day) using VGA indicators

[Weekday]

Multiple correlation coefficient : 0.593
Coefficient of determination : 0.352 AIC : 492.942

	Standard partial regression coefficient	Partial regression coefficient	t value	p value
Constant	—	0.496	1.140	—
(X1) distance from station entrances	-0.404	-0.004	-5.428	0.000
(X8) global integration value (GIV)	0.193	0.338	3.078	0.002
(X6) floor-area ratio of first-floor commercial buildings	0.173	0.011	2.350	0.020

[Holiday]

Multiple correlation coefficient : 0.651
Coefficient of determination : 0.424 AIC : 595.593

	Standard partial regression coefficient	Partial regression coefficient	t value	p value
Constant	—	0.189	0.314	—
(X1) distance from station entrance	-0.320	-0.005	-4.171	0.000
(X6) floor-area ratio of first-floor commercial buildings	0.289	0.026	3.905	0.000
(X8) global integration value (GIV)	0.132	0.325	1.913	0.057
(X4) floor-area ratio of accommodation facilities	0.121	0.005	1.993	0.048
(X7) visible area	0.108	0.000	1.618	0.107
(X3) floor-area ratio of offices and schools	0.099	0.001	1.569	0.119

The variables from each factor group of transportation accessibility, facility volume, and space configuration, were used without omission in both the weekend and holiday models, and it can be observed that all three groups contributed to the results in the Kanayama area. This result is consistent with the results of current studies and supports the validity of the VGA indicator as a space configuration index in the Kanayama area.

3.4 Linear regression model analysis using Agent Analysis (AA) indicator

3.4.1 Introducing agent analysis (AA)

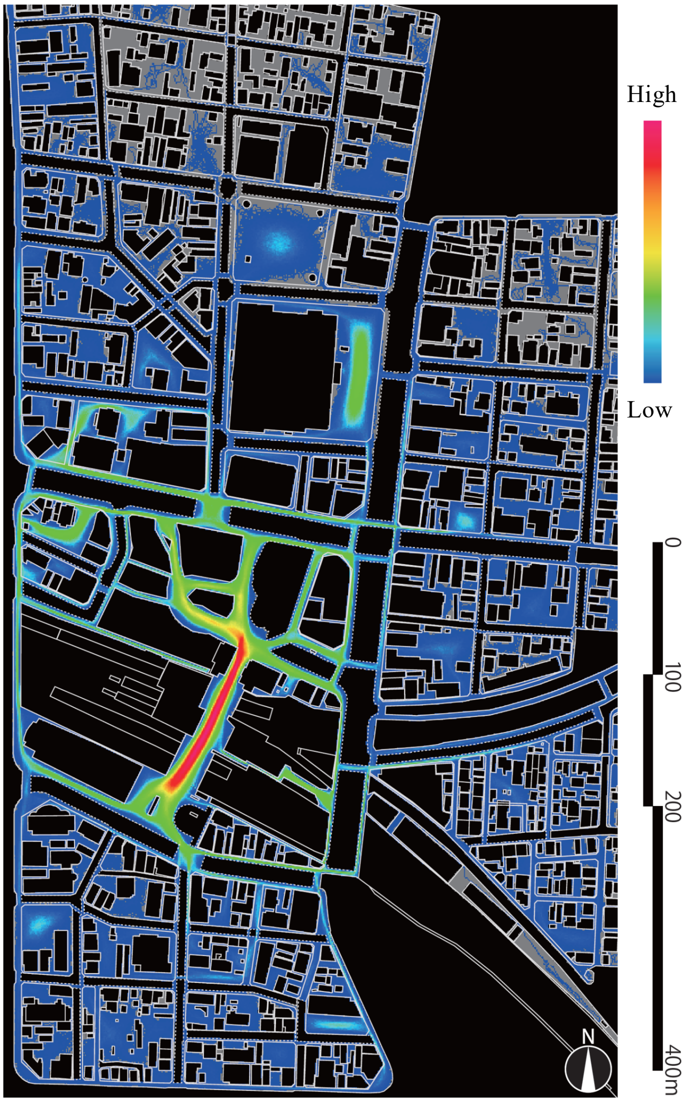


Figure 3: Spatial distributions of footprint counts (AA, station-generated)

We attempted to apply AA. (1) Random generation: Only the quantity of agents generated per step is determined, and agents are generated at random from various locations on the map (the same probability is given per grid cell). (2) Selected generation: where the number and ratio of agent generations are predetermined. In this study, we use the term "station-generated" because the selected generation's point of generation is the station.

-The number of agent generations: 2000

-Agent movement distance: uniform distribution from 0 to 1500 m (average 750 m)

-Points of generation of agent stations: 8 points in total at Kanayama Sogo Station entrance and exit
-Agent station generation ratio: cross-sectional pedestrian volume survey results (outflow from station to Kanayama area)

Furthermore, each pedestrian agent is assumed to leave one footprint per second in the walking space, and the number of footprints counted per meter of street length in each street space once the simulation is completed is used as an indicator. According to its definition, this dimension is [counts/m]. In this study, two types of footprints are used: the number of counted footprints (AA, randomly-generated) and the number of counted footprints (AA, station-generated). Figure 3 shows the spatial distribution of footprint counts (AA, station-generated). The longer the walking trajectory, the larger the walkable space area and the closer it is to the station. This could be owing to the agent's behavioral characteristics and the influence of the agent's point of generation.

3.4.2 Analysis of factor using the AA indicator

The same analysis was conducted for the number of pedestrians on weekdays and holidays by replacing (X7) visible area and (X8) GIV, which are VGA indicators among the candidate factor variables, with (X9) footprint counts (AA, randomly-generated) and (X10) footprint counts (AA, station-generated). To avoid multicollinearity, the VIF between the candidate factor variables was less than 2, and all candidate factor variables were used for the analysis.

The analysis with (X9) footprint counts (AA, randomly-generated) yielded a multiple correlation coefficient of 0.726, AIC 438.726 for the weekday model and a multiple correlation coefficient of 0.693, AIC 581.396 for the holiday model, whereas the analysis with (X10) footprint counts (AA, station-generated) yielded a multiple correlation coefficient of 0.850, AIC 341.577 for the weekday model and a multiple correlation coefficient of 0.737, AIC 554.142 for the holiday model. The analysis using (X10) footprint counts (AA, station-generated) had an improvement for both weekday and holiday in the multiple correlation coefficient compared to the two models using the (X7) visible area and (X8) GIV of the VGA indicators and the (X9) footprint counts (AA, randomly-generated). The weekday model, in particular, recorded a multiple correlation coefficient of 0.850, which is close to the level that can be applied as a prediction model.
Three variables were selected for the weekday model and five variables for the holiday model using (X10) footprint counts (AA, station-generated) (Table.2). In both the weekday and holiday models, variables from each of the groups of transportation accessibility, facility volume, and space configuration were adopted without omissions, and the structure is similar to that of the model using the VGA indicator.

When the model with (X10) footprint counts (AA, station-generated) was compared to the model with (X9) footprint counts (AA, randomly-generated) , the intensity of the (X10) footprint counts

(AA, station-generated) introduced as a space configuration index became larger than that of the (X9) footprint counts (AA, randomly-generated) (standard partial regression coefficients: from 0.480 to 0.705 on a weekday and from 0.309 to 0.461 on a holiday), but note that the intensity of (X1) distance from station entrances, which is a transportation accessibility indicator, decreased accordingly (standard partial regression coefficients: from −0.369 to −0.267 on a weekday and from −0.299 to −0.243 on a holiday). This is because the AA (station-generated) indicator is not only an indicator of the space configuration factors, but also an indicator of transportation accessibility factors.

Table.2: Result of analysis of factor (multiple regression model selection) for the number of pedestrians on weekday and holiday (all day) using AA (station-generated) indicator

[Weekday]

Multiple correlation coefficient : 0.850
Coefficient of determination : 0.723 AIC : 341.577

	Standard partial regression coefficient	Partial regression coefficient	t value	p value
Constant	–	1.117	8.793	–
(X10)footprint counts (AA, station-generated)[counts/m]	0.705	0.179	16.413	0.000
(X1) distance from station entrance	-0.267	-0.003	-5.988	0.000
(X3) floor-area ratio of offices and schools	0.087	0.001	2.083	0.039

[Holiday]

Multiple correlation coefficient : 0.737
Coefficient of determination : 0.543 AIC : 554.142

	Standard partial regression coefficient	Partial regression coefficient	t value	p value
Constant	–	1.084	3.369	–
(X10)footprint counts (AA, station-generated)[counts/m]	0.461	0.164	8.012	0.000
(X1) distance from station entrance	-0.243	-0.004	-3.546	0.001
(X3) floor-area ratio of offices and schools	0.156	0.002	2.856	0.005
(X6) floor-area ratio of first-floor commercial buildings	0.151	0.014	2.280	0.024
(X4) floor-area ratio of accommodation facilities	0.148	0.006	2.770	0.006

Furthermore, a comparison of the variables adopted in the two models and their intensities on weekdays and holidays shows that the intensities of (X6) floor-area ratio of first-floor commercial buildings (not adopted on a weekday, standard partial regression coefficients:0.151 on a holiday) and (X4) floor-area ratio of accommodation facilities (not adopted on a weekday, standard partial regression coefficients:0.148 on a holiday) are larger on a holiday than on a weekday. The model, like the ones in Section 3 -3, captures the characteristics of the liveliness of the Kanayama area on weekdays and holidays.

When the models were compared with the smallest AIC value for each selected number of variables, the models with the smallest AIC values are the three-variable weekday models and the five-variable holiday models without no major interruption. This supports the validity of the models selected. (Table.3)

Table.3: Selected multiple regression models for pedestrians on weekday and holiday (All day) by number of factors

Number of Factors	Model	Single correlation: Multiple correlation	Single correlation	Standard partial regression coefficient	AIC	t test
1	(X10)footprint counts (AA, station-generated)[counts/m]	0.800	0.800	0.800	383.975	***
2	(X10)footprint counts (AA, station-generated)[counts/m]	0.846	0.800	0.695	343.961	***
	(X1) distance from station entrance		-0.543	-0.294		***
3	(X10)footprint counts (AA, station-generated)[counts/m]	0.850	0.800	0.705	341.577	***
	(X1) distance from station entrance		-0.543	-0.267		***
	(X3) floor-area ratio of offices and schools		0.164	0.087		**
4	(X10)footprint counts (AA, station-generated)[counts/m]	0.852	0.800	0.705	342.195	***
	(X1) distance from station entrance		-0.543	-0.266		***
	(X3) floor-area ratio of offices and schools		0.164	0.087		**
	(X2) floor-area ratio of commercial facilities		0.350	0.011		

(***: 1%, **: 5%, * : 10%)

[Weekday]

[Holiday]

Number of Factors	Model	Single correlation: Multiple correlation	Single correlation	Standard partial regression coefficient	AIC	t test
1	(X10)footprint counts (AA, station-generated)[counts/m]	0.607	0.607	0.607	603.940	***
2	(X10)footprint counts (AA, station-generated)[counts/m]	0.707	0.607	0.468	564.374	***
	(X1) distance from station entrance		-0.556	-0.388		***
3	(X10)footprint counts (AA, station-generated)[counts/m]	0.717	0.607	0.482	561.018	***
	(X1) distance from station entrance		-0.556	-0.349		***
	(X3) floor-area ratio of offices and schools		0.225	0.127		**
4	(X10)footprint counts (AA, station-generated)[counts/m]	0.728	0.607	0.494	557.444	***
	(X1) distance from station entrance		-0.556	-0.323		***
	(X3) floor-area ratio of offices and schools		0.225	0.140		**
	(X4) floor-area ratio of accommodation facilities		0.155	0.125		**
5	(X10)footprint counts (AA, station-generated)[counts/m]	0.737	0.607	0.461	554.142	***
	(X1) distance from station entrance		-0.556	-0.243		***
	(X3) floor-area ratio of offices and schools		0.225	0.156		***
	(X6) floor-area ratio of first-floor commercial buildings		0.471	0.151		**
	(X4) floor-area ratio of accommodation facilities		0.155	0.148		***
6	(X10)footprint counts (AA, station-generated)[counts/m]	0.739	0.607	0.461	555.358	***
	(X1) distance from station entrance		-0.556	-0.235		***
	(X3) floor-area ratio of offices and schools		0.225	0.163		***
	(X4) floor-area ratio of accommodation facilities		0.155	0.142		***
	(X6) floor-area ratio of first-floor commercial buildings		0.471	0.141		**
	(X5) floor-area ratio of cultural facilities		0.020	0.023		

(***: 1%, **: 5%, * : 10%)

3.5 Replacing to Mobile Phone Location Data

In this section, our analysis of factors is performed in response to the analytical framework described in the previous chapter, by replacing the objective variable, the number of all-day pedestrians, with mobile phone location data from the encounter survey, and the results of the analysis are compared and discussed. The data is the KDDI Location Analyzer's primary

movement line data (KLA). KDDI, a Japanese mobile phone service company, uses GPS location data obtained from its smartphone users to make expanded estimates by using the official population statistics. From March 22, 2019 to March 21, 2020, the average value per day (5:00 am on the same day to 4:00 am on the next day) was used for each typical holiday.
Due to the omission of narrow-width streets and the integration of sidewalks on both sides of one street, the KLA primary movement line data is missing 9 data points out of 178 street samples and the 169 samples are used for the analysis.

In the analysis of the KLA data, a three-variable model for weekdays and a four-variable model for holidays were selected. The weekday model's multiple correlation coefficient was 0.740 for the encounter survey data and 0.710 for the KLA data, and the AIC was 355.812 for the encounter survey data and 371.081 for the KLA data. The multiple correlation coefficient of the holiday model was 0.712 for the encounter survey data and 0.736 for the KLA data, and the AIC was 372.037 for the encounter survey data had an AIC of 359.694. The model based on encounter survey data fit well during the weekdays, whereas the model based on KLA data fit well on holidays, and no significant difference was found.

The variables adopted were consistent with the results of the encounter survey data up to the third rank, and in the order of strength of the standardized partial regression coefficients, they were: (X10) footprint counts (AA, station-generated); (X1) distance from station entrances; and (X3) floor-area ratio of offices and schools up to the third rank for both weekday and holiday. The fourth rank in the four-variable model for holidays was selected as (X4) floor-area ratio of accommodation facilities for the encounter survey data and (X5) floor-area ratio of cultural facilities for the KLA data, and there was no significant difference in the ranks as well.

3.6 Chapter conclusion

Comparing VGA and AA, the model equation fits better for AA regardless of weekdays or holidays, especially on weekdays, with a multiple correlation coefficient of 0.850 (station-generated) versus 0.593.

The order of factor intensity is shown for weekdays as an example: VGA was transportation accessibility, GIV, and facility volume; AA (station-generated) was AA footprint, transportation accessibility, and facility volume. Furthermore, the AA (station-generated) indicator has the both characteristics of a transportation accessibility factor and the space configuration factor, thus it is suggested that the AA indicator is enough powerful to replace the VGA indicator.

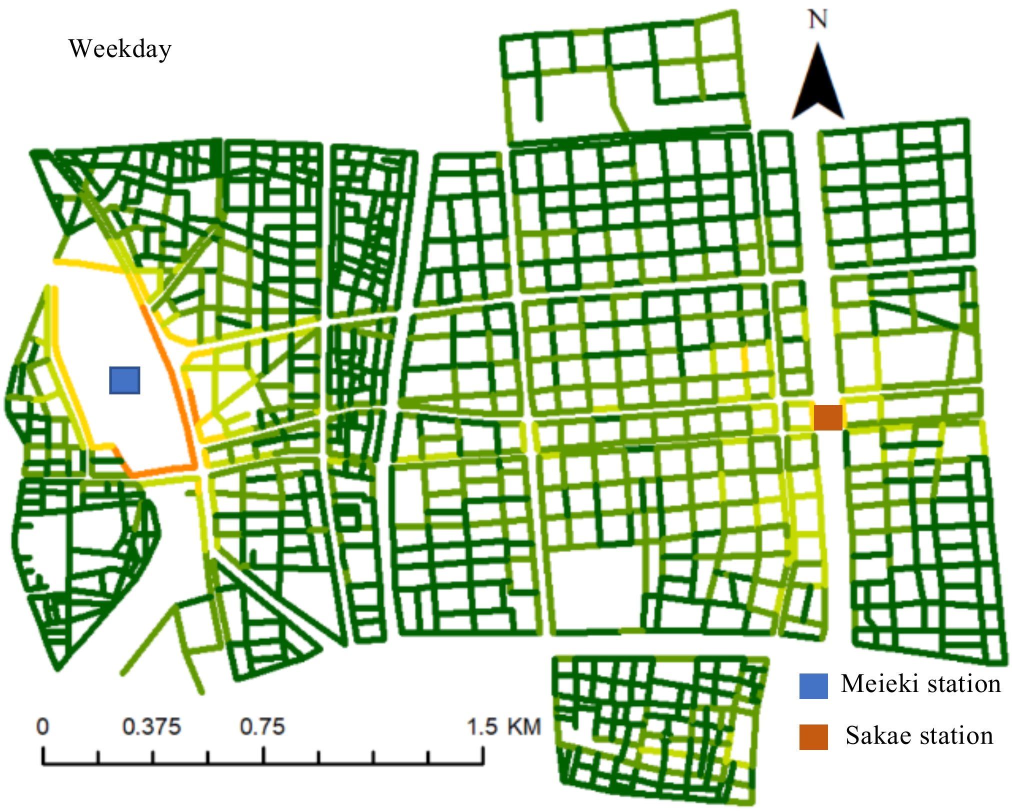
Additionally, we attempted to conduct our factor analysis by replacing the encounter survey data with primary movement line data from mobile phone location data and confirmed that the extracted factor structure was robust.

4 NAGOYA CBD CASE: ANALYSIS OF FACTORS USING LOG-LOG MODEL

4.1 Nagoya CBD case

In this chapter, we describe an analysis of the factors that affect the number of pedestrians in Nagoya CBD, with the number of pedestrians as the objective variable, based on mobile phone location data (small-areal aggregation data) and using a log-log model that introduces the Segment Angular Analysis indicator in Space Syntax theory.

The Nagoya CBD, which is the core of the Nagoya metropolitan area, one of Japan's, three largest metropolitan areas, contains two commercial and office centers around on Nagoya Station (hereinafter referred to as Meieki) and Sakae Station (hereinafter referred to as Sakae). In this study, the Nagoya CBD, which includes the Meieki and Sakae areas, is referred to as the study area and covers approximately 635 hectares.



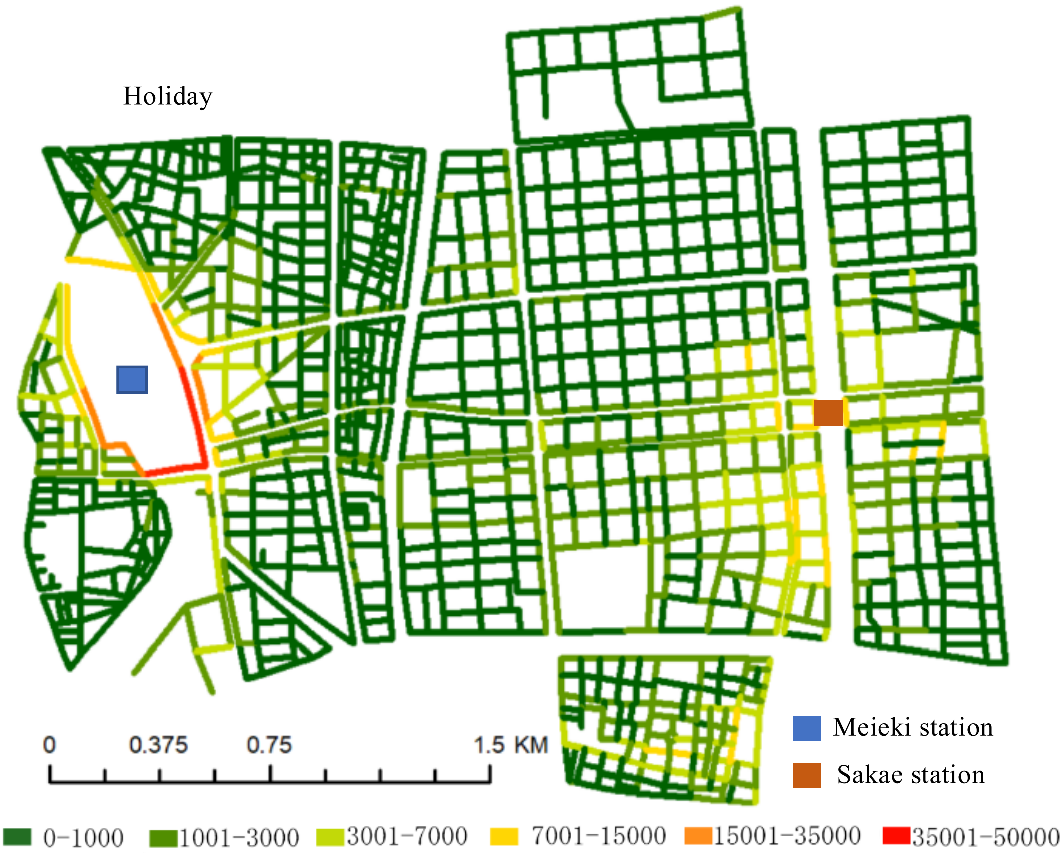


Figure 4: Spatial distribution of the pedestrians (persons/day)

4.2 Mobile Phone Location Data

In this study, the primary movement line data as well as the number of pedestrians were obtained from the KDDI Location Analyzer (KLA). The collection period for the data in this chapter was from March 22, 2019, to March 21, 2020 and the daily average was used. Weekday and holiday users must be over the age of 20 years old to only walk.

Observe the spatial distribution of pedestrians using GIS (Figure 4). The streets (sidewalks) around Nagoya Station had the highest number of pedestrians, both on weekdays and holidays, with 34,291persons per day on the streets with the highest number of pedestrians on weekdays and 46,703 persons per day on holidays.

When the spatial distribution of pedestrians on the streets on weekdays and holidays was compared in the Meieki area, there was no significant difference in the distribution range between weekdays and holidays.

On the other hand, we can find that the distribution of pedestrians in the Sakae area is much more widely on weekdays than on holidays.

4.3 Candidate Explanatory Variables

In this study, the number of pedestrians was used as the objective variable, while 15 variables were created as candidate explanatory variables. We categorized them into four groups this time around of factors: street attribute factor (4 variables), facility volume (land use) factor (6 variables), space configuration factor (2 variables), and transportation accessibility factor (3 variables).

As for the group of the street attribute factors, we created four potential explanatory factors: (Xa1) street width (m), (Xa2) sidewalk dummy (has no sidewalk: 0, has sidewalk: 1), (Xa3) trunk road dummy (less than four vehicle lanes: 0, more than or equal to four vehicle lanes: 1), and (Xa4) street length (m).

As for the group of the facility volume (land use) factors, we prepared six candidate explanatory factors: (Xb1) floor-area ratio of retails, (Xb2) floor-area ratio of entertainment, (Xb3) floor-area ratio of offices, (Xb4) floor-area ratio of retails (average within 400m), (Xb5) floor-area ratio of entertainment (average within 400m), (Xb6) floor-area ratio of offices (average within 400m).
As for the group of the space configuration factors, we prepared two candidate explanatory factors: (Xc1) integration value (R = 400m) and (Xc2) integration value (R = 1200m).
As for the group of the transportation accessibility factors, we prepared three candidate explanatory factors: (Xd1) distance from the nearest station (m), (Xd2) distance to Meieki (m), and (Xd3) distance to Sakae (m).

In this paper, the number of pedestrians (weekday (Y1), holiday (Y2)) and (Xd1) distance to the nearest station (m), (Xd2) distance to Meieki(m), and (Xd3) distance to Sakae (m) are logarithmically transformed for analysis by the log-log model.

4.3.1 Floor Area Ratio

We used data from the Nagoya City Basic Urban Planning Assessment 2016 survey of current conditions by building use data to calculate the floor area ratio of the candidate factor variables used in this study. We extracted data for the three categories of retail, entertainment, and office, and calculated the floor-area ratio for each use in each block in the Nagoya CBD after calculating the floor area for each use.

When calculating floor-area ratio by land use, when a street has a median strip, the floor-area ratio of one adjacent side block of the street is used as usual, and when there is no median strip, the average of the floor-area ratio of both adjacent side blocks of the street is used.
Since the number of pedestrians on a given street may affect surrounding streets, the center point of the street segment was chosen as the representative point of each street in this study, and the average of the floor-area ratio of all streets within a 400-meter radius was organized as the street's neighborhood floor-area ratios.

4.3.2 Space Configuration Index using Segment Angular Analysis

Using the Geospatial Information Authority of Japan's base map information, we created a segment map of the Nagoya CBD. When plotting the segment map, a buffer zone of 400m was placed around the periphery of the study area, taking into account edge effects. As for the space configuration index, we use the integration value from the UCL depthmap's segment angular analysis.

When performing the segment angular analysis, it is possible to specify the analysis area to meet the purpose, and this area is defined as the radius. In this study, two types of radius were analyzed: 400m (neighborhood) and 1200m (wide area).

The spatial distribution of the analysis results (Figure 5) shows that the integration value of the intersection around Meieki is high at radius = 400m, but this can be attributed to the short radial streets around the intersection. At Radius = 1200m, the integration value of streets in the Sakae area is high. In the case of radius = 1200m, the integration value of streets in the Sakae area is high, which is thought to be due to the fact that the streets in the Sakae area are grid-plan.

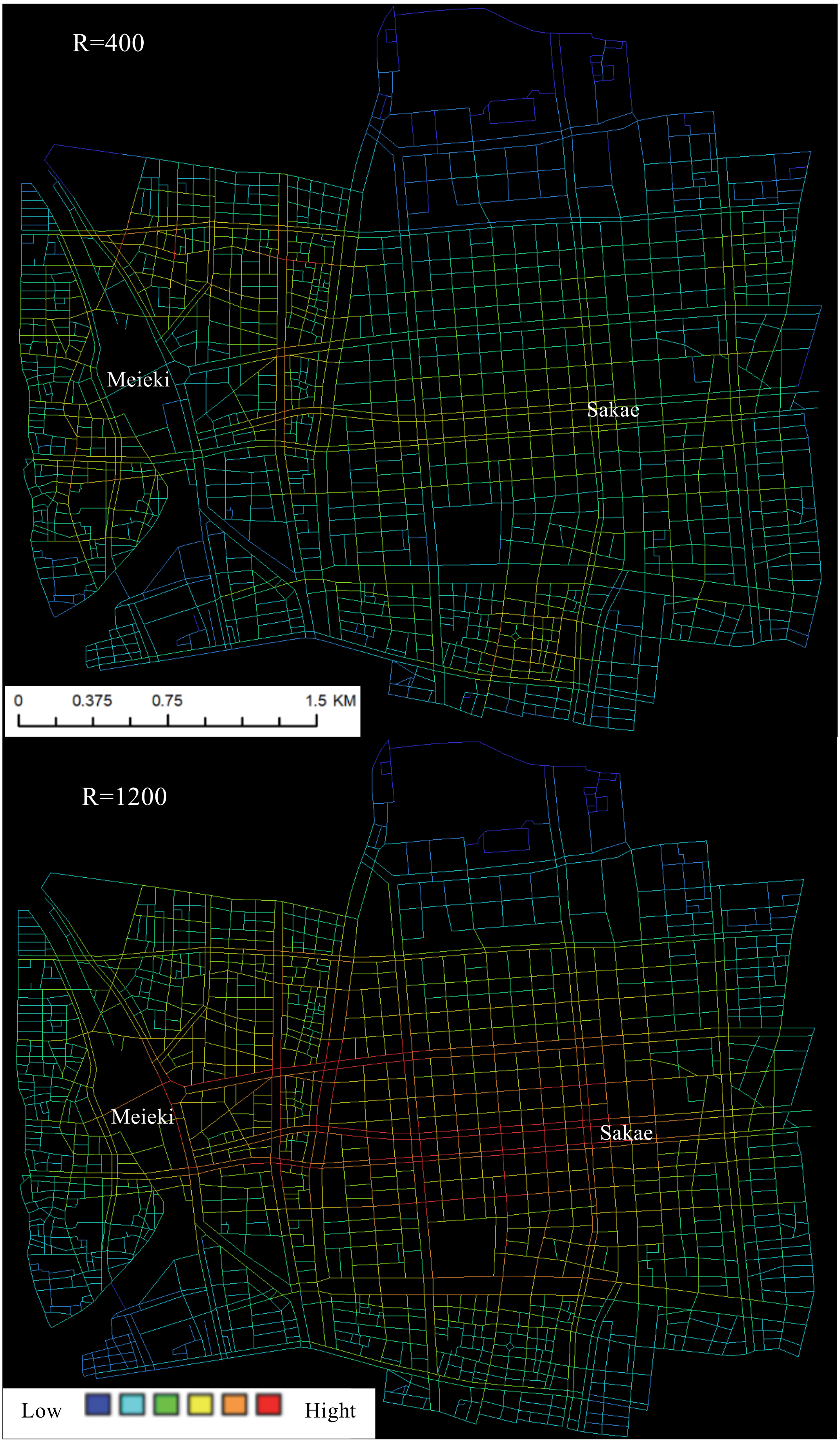


Figure 5: Integration value (R = 400, 1200) by segment angular

4.4 Log-log regression model analysis

4.4.1 Model Selection

The correlations between each candidate factor variable, had an absolute value of 0.7 or above for (Xa1) street width and (Xa3) trunk road dummy (0.880), (Xb5) floor-area ratio of entertainment (average within 400m) and (Xb6) floor-area ratio of offices (average within 400m) (0.725).

To examine multicollinearity, the VIF between each candidate factor variable is calculated; since combinations of candidate factor variables with a VIF greater than 2 are suspected of multicollinearity, only one of the candidate factor variables is included in the factor analysis. In the case of Xa1 street width (m) and (Xa3) trunk road dummy (VIF: 4.462), we adopted (Xa1) street width (m) and excluded (Xa3) trunk road dummy because the pedestrians in this study are more likely to be affected by street width than trunk road dummy. The (Xa3) trunk road dummy was excluded. Furthermore, (Xb5) floor-area ratio of entertainment (average within 400m range) and (Xb6) floor-area ratio of offices (average within 400m range) (VIF: 2.106) will be explored in detail for weekday and holiday separately.

Multiple regression analysis of log-log models is performed for the number of pedestrians on each weekday and holiday, using the candidate factor variables after VIF exclusion. In model selection, we used stepwise regression analysis with the minimum AIC in model selection.
The 10-variable model, the model with the smallest AIC excluding (Xb5) floor-area ratio of entertainment (an average within a 400m range), is (AIC: 3789.555). The model with the smallest AIC, excluding the (Xb6) floor-area ratio of offices, (an average within the 400m range), is the 11-variable model (AIC: 3842.135). The weekday model was selected by excluding (Xb5) floor-area ratio of the entertainment (an average within the 400m range).

In the holiday model, the model with the smallest AIC, excluding (Xb5) floor-area ratio of entertainment (an average within a 400m range), is the 10-variable model (AIC: 4556.39). The model with the smallest AIC, excluding the (Xb6) floor-area ratio of the offices (an average within a 400m range), is the 11-variable model (AIC: 4547.347). For the holiday model, the model excluding (Xb6) floor-area ratio of the offices (average within a 400m range) was selected.

4.4.2 Selection of a Five-Variables Model and its Implication

For the purposes of this chapter, we use a model with only five variables for both holidays and weekdays. Furthermore, the AIC minimum models were used, which confirm the order preserved. Table.4 displays the outcomes.

Table.4: Result of analysis of factor for pedestrians on weekday and holiday using log-log model

[weekday]

Multiple correlation coefficient: 0.750

Coefficient of determination:0.562 AIC : 3939.876

	Standardized Partial Differential Value	Partial regression coeffieient	t-value	p-value
Constant	-	9.676	29.367	-
(Xa2)sidewalk dummy	0.100	0.750	18.411	0.000
(Xb4) Retail uses (FAR)(average within 400m)	0.109	0.013	22.104	0.000
(Xb6) Office uses(FAR) (average within 400m)	0.096	0.004	18.272	0.000
(Xc1)Integration value(R=400)	0.055	0.010	10.581	0.000
Ln(Xd2)Distance from Meieki	-0.061	-0.710	-15.801	0.000

[holiday]

Multiple correlation coefficient: 0.701

Coefficient of determination:0.4914 AIC : 4674.083

	Standardized Partial Differential Value	Partial regression coeffieient	t-value	p-value
Constant		12.032	26.554	-
(Xa2)sidewalk dummy	0.071	0.792	17.185	0.000
(Xb4) Retail uses (FAR)(average within 400m)	0.125	0.022	31.151	0.000
(Xc1)Integration value(R=400)	0.038	0.010	9.116	0.000
Ln(Xd1)Distance to the nearest station	-0.018	-0.224	-6.421	0.000
Ln(Xd2)Distance from Meieki	-0.048	-0.832	-15.42	0.000

The multiple correlation coefficient for the weekday model is 0.750 (coefficient of determination 0.562) and the AIC is 3939.876, whereas the multiple correlation coefficient for the holiday model is 0.701 (coefficient of determination 0.491) and the AIC is 4674.083. When these two models are compared, the holiday model has a higher multiple correlation coefficient (coefficient of determination) and can be outlined as having more explanatory power.

The variables adopted in the weekday model were in the order of increasing absolute value of the standardized partial derivative at the mean value of x_i :(Xb4) floor-area ratio of retails (average within 400m) (partial derivative: 0.109), (Xa2) sidewalk dummy (0.100), (Xb6) floor-area ratio of offices (average within 400m) (0.096), (lnXd2) distance to Meieki (−0.061), (Xc1) integration value (R = 400m) (0.055).

In the holiday model, the following parameters were used in increasing order of increasing absolute value of the standardized partial derivative: (Xb4) floor-area ratio of retails (average within 400m) (0.125), (Xa2) sidewalk dummy (0.071), (lnXd2) distance to Meieki (−0.048), (Xc1) integration value (R = 400m) (0.038), and (lnXd1) distance to the nearest station (−0.018). (Xc1) Integration value (R = 400m) was adopted at 1% significance in both weekday and holiday scenarios, suggesting the effectiveness of the integration value as a space configuration factor for pedestrians in this study, which used log-log model analyses for the Nagoya CBD.

Comparing and examining each variable adopted and its intensity in the two models for weekday and holiday, it can be noted that (Xb6) floor-area ratio of offices (an average within a 400m range), which was adopted on a weekday, was not adopted on holiday, possibly due to the influence of office commuters on weekdays. (lnXd1) on weekdays the distance from the nearest station was not adopted, but on weekends and holidays, it was. The partial differential value of the (Xb4) floor-area ratio of retails (an average within a 400m range), increased from 0.109 on weekdays to 0.125 on holidays, suggesting that many visitors ride the subway to the store on holiday. This indicates that the model captures the characteristics of liveliness in the Nagoya CBD on weekdays and holidays.

4.5 Chapter Conclusion

In Nagoya CBD case, the variables were adopted without omission from each factor group of transportation accessibility, facility volume (land use), space configuration, and street attributes in both the weekday and holiday models of the log-log model. As an indicator of the space configuration, Integration Value (R=300m) obtained from Segment Angular Analysis was suggested to be effective.

Furthermore, the standardized partial differential value is used to compare the importance of each explanatory variable in the log-log model. This chapter results show that the log-log model with the addition of the space configuration indicators reflects on the characteristics of the liveliness on weekdays and holidays in the city center.

5 CONCLUSION -- TOWARD THE ESTABLISHMENT OF MORE GENERALIZED ANALYTICAL FRAMEWORK

The paper discussed in detail our on-going analytical framework for Analysis of Factor of Busy Streets (AoFBS), the three issues we face today, mobile phone location data, spatial representation and Space Syntax indicators, and extension to nonlinear models, and the two analytical cases we are currently working on.

Since a 1-meter fine grid can be built out, in the analysis of the 40-hectare Kanayama case described in Chapter 3, GIV in VGA was calculated, and the agent analysis was attempted. Furthermore, linear model analysis confirmed that a space configuration factor had a significant contribution. We primarily used encounter survey data, but we also confirmed that there was no significant difference in the results using small-areal aggregation data of mobile phone location data.

Additionally, a thorough study of the agent analysis was conducted in chapter 3. Following a series of analyses, it was concluded that the AA indicator not only replaces the VGA indicator as an indicator of the space configuration factors, but it also replaces the VGA indicator as an index of the transportation accessibility factor, particularly in the station-generated case.

In the analysis of the 645-hectare Nagoya downtown area in Chapter 4, since we used street segment data for GIS, we calculated the Integration Value in the Segment Angular Analysis and confirmed the contribution of the space configuration index by obtaining the standard partial differential equation in the log-log model analysis. We used the primary movement line data from the small-areal aggregation data of mobile phone location data as a data collection method. The results of the two analyses support the conclusions of our previous research, and therefore, we can conclude that the space configuration indicators influenced the number of pedestrians in the Nagoya downtown area.

The authors' analysis method simply extracts factors in the linear model analysis by comparing the strengths of the standardized partial regression coefficients of the model forms obtained in the multiple regression analysis. However, since the partial regression coefficient depends on the value of each x_i, in the log-log model analysis, the standard partial differential equation, which means a standardized slope, was obtained instead, and the mean standard partial differential value, which was substituted for the mean value of each x_i, was used to examine the factors' strength. Although multiple regression analysis has an inferential statistical basis, future research may be directed toward "limited interpretation" based on the accuracy characteristics of the selected data collection method or "active interpretation" based on covariance structure analysis, leaving the position of inferential statistics.

The multiple correlations and t-values of both the objective variable and the candidate explanatory variables in the log-linear model and the log-log model can be improved by transforming them to logarithms, and sometimes logarithmic transformation is desirable due to the meanings of the variables and models. This is an example of the latter when the objective variable, the number of pedestrians counted in a unit time, is assumed to have a Poisson distribution. The latter example is a gravity model in gravity model in the log-log model with facility volume factor x_1 and transportation accessibility factor x_2, and the model of $\ln y = \alpha \ln x_1 - \beta \ln x_2 + \gamma$ the gravity model $y = \gamma x_1{}^{\alpha} / x_2{}^{\beta}$.

Table.5: Generalized Framework

Model Family	Linear Model	Exponetial Model		
Model Equation		Log-linear Model or Log-log Model	Generalized Linear Model (GLM)	
Assumed Distribution of Objective Variable	(Normal Distribution)	Log-normal Distribution	Poisson Distribution	Negative Binominal Distribution
Calculation	Multiple Regression Analysis (MRA, OLS)	Multiple regression analysis after log transformation of the equation	Applying GLM by specifying the assumed distribution and the error structure (Maximum Likelihood Method)	
Factor Intensity	Standardized Partial Regression Coefficient	Mean Standardized Partial Derivative		

As our future work, there is a generalization of the analytical framework through the application of Generalized Linear Model (GLM) in Table 5. GLM is a statistical modeling with an exponential model, although we are unable to show an example of analysis in Meieki case due to space limitations. When a normal distribution is assumed for the explanatory variables, for an example, the error structure of the GLM is additive ($+\varepsilon$), whereas the error structure of the log-linear and log-log models is multiplicative ($\cdot e^{\varepsilon}$). Negative binomial distribution regression may be performed when over-dispersion is an issue (Stavroulaki et.al, 2019). However, the distributional assumptions and error structure are case-by-case, and while an additive error structure is suggested if the Poisson assumption is introduced for pedestrian counts per unit time, whether the errors associated with small-areal aggregation of cell phone location data support this is another question. In addition, when the number of pedestrians divided by the street length is log-transformed and used as the objective variable, a right-hand side transfer process, called the offset process, is also employed.

ENDNOTE

Definitions of Pedestrian Number Indicator

A pedestrian in the CBD is a person walking down the street. If we assume that a pedestrian draws a single movement line, the pedestrian number is equal to the number of pedestrian movements, which are referred to as pedestrian volume and pedestrian flow, respectively.
In Space Syntax, the street network is represented as a so-called movement line model, with streets without separated pedestrian paths represented by a single link, and streets with separated pedestrian paths divided into links corresponding to sidewalks. The link, classified in a manner consistent with the street segment, is the unit of aggregation for the number of pedestrians. Crosswalks are often viewed as linkages. The number of pedestrians is assigned to each street (sidewalk) link as a statistic.

Depending on the purpose of the analysis, the number of pedestrians might be indexed with different meanings. The following is a typical example:

(1) Number of pedestrians
(2) Number of pedestrians / street (sidewalk) area
(3) Number of pedestrians / street (sidewalk) width
(4) Number of pedestrians / street (sidewalk) length

The number of pedestrians in (1) was previously written to be equal to the number of movement lines, but this was done for simplification, and does not apply if the measurement time is short or the measurement area is small. The dimension of (2) coincides with pedestrian density, which is known as the level of service (LoS) concept explored by Fruin (LoS is the inverse of density). In this case, the level of service is the ratio of sidewalk supply to pedestrian demand, and if the measurement time is long enough and the number of movement lines can be regarded as the pedestrian demand, the indicator (3) is also valid. Additionally, when the length of a street (movement line) link is uniform, or when the measurement time is limited, as in an encounter survey, the number of people present can be considered proportional to the street space area, and the indicator (4) with the correction per unit length may be necessary (Hillier et.al 1993, Kaneda, Ota and Kobayashi 2020).

Mean Standardized Partial Differential Value, (MSPDV)

Given a linear model $\ln y = \beta_0 + \beta_1 x_1 + \beta_2 \ln x_2$, where y and x_2 are logarithmically transformed. Since, $y = \mathrm{e}^{\beta_0} \cdot \mathrm{e}^{\beta_1 x_1} \cdot {x_2}^{\beta_2}$, the partial derivative of y for x_1 is

$$\frac{\partial y}{\partial x_1} = \mathrm{e}^{\beta_0} \cdot \left(\mathrm{e}^{\beta_1 x_1}\right)' \cdot {x_2}^{\beta_2}$$
$$= \beta_1 \cdot \mathrm{e}^{\beta_0} \cdot \mathrm{e}^{\beta_1 x_1} \cdot {x_2}^{\beta_2}.$$

Here, the partial differential equation $\left(\frac{SD_{x_1}}{SD_y}\right)\frac{\partial y}{\partial x_1}$ modified by using the standard deviation of x_1 and y is called the standardized partial differential equation (SPDE). Since the specific value depends on each x_i, the value when the average value is substituted for each is called the x-mean standardized partial differentiation coefficient MSPDV for x_1.

$$\left(\frac{SD_{x_1}}{SD_y}\right)\frac{\partial y}{\partial x_1}\bigg|_{x=\overline{x_i}\ for\ all\ i} = \left(\frac{SD_{x_1}}{SD_y}\right) \cdot \beta_1 \cdot \mathrm{e}^{\beta_0} \cdot \mathrm{e}^{\beta_1 \overline{x_1}} \cdot \overline{x_2}^{\beta_2}$$

Note that the MSPDV for x_2 is

$$\left(\frac{SD_{x_2}}{SD_y}\right)\frac{\partial y}{\partial x_2}\bigg|_{x=\overline{x_i}\ for\ all\ i} = \left(\frac{SD_{x_2}}{SD_y}\right) \cdot \beta_2 \cdot \mathrm{e}^{\beta_0} \cdot \mathrm{e}^{\beta_1 \overline{x_1}} \cdot \overline{x_2}^{\beta_2 - 1}.$$

ACKNOWLEDGMENT

This research is sponsored by Japan Society for the Promotion of Science, KAKENHI Grant Number: 18H03825.

REFERENCES

Desyllas, J. et.al. (2003) 'Pedestrian demand modelling of large cities: an applied example from London', (CASA Working Papers 62). Centre for Advanced Spatial Analysis (UCL): London, UK.

Ferguson, P., Friedrich, E. and Karimi, K. (2012) 'Origin-Destination weighting in agent modelling for pedestrian movement forecasting', Paper presented at 8th International Space Syntax Symposium, Santiago, Chile, 3-6th January 2012.

Hillier, B. et.al. (1993) 'Natural Movement: Or, Configuration and Attraction in Urban Pedestrian Movement', Environment and Planning B: Planning and Design, 20(1), pp. 29–66.

Kaneda, T. et.al. (2019) 'Modeling and simulation of vision-driven pedestrian agent." Paper presented at 12th space syntax symposium. Beijing, China, 8 -13 July 2019.

Kaneda, T., Ota, A. and Kobayashi, J. (2020) 'A study on analyses of factor of pedestrians' spatial distribution in Kanayama district, Nagoya - By applying agent analysis in space syntax studies-', Journal of Architecture and Planning, 85 (767), pp.121-129.

Lerman, Y., Rofe, Y. and Omer, I. (2014) 'Using Space Syntax to Model Pedestrian Movement in Urban Transportation Planning', Geographical Analysis,46(4) p.392–410.

Okamoto, K. et.al. (2013) 'Correlation Analyses between Underground Spatial Configuration and Pedestrian Flows by Space Syntax measures – A Case Study of Underground Mall Complex in Nagoya Station –', Paul Sanders, Mirko Guaralda & Linda Carroli, 20th International Seminar on Urban Form Volume 1, Brisbane, Australia, 17-20 July, pp. 116-128.

Omer, I. and Kaplan, N. (2017) 'Using space syntax and agent-based approaches for modeling pedestrian volume at the urban scale', Computers, Environment and Urban Systems,64, pp. 57-67.

Ota, A. et.al. (2008) 'Factor Analysis of the Pedestrian Number in a Downtown Area Using Visibility Indicators: A Case Study of the Sakae-South District, a Downtown Area of Nagoya', Ray Lucas & Gordon Mair(eds.), Sensory Urbanism Proceedings 2008, 6 Jan 2009, pp.158-164.

Ota, A., Meziani, R. and Kaneda, T. (2015) 'Comparative Analyses on Factors of the Pedestrian Numbers in a Downtown Area Using Space Syntax Indicators – Case comparisons of Sakae-South, Nagoya CBD in Japan between 2005 and 2011', Paper presented at 10th International Space Syntax Symposium, London, UK, 13-17 July 2017.

Ota, A. Takahashi, H. and Kaneda, T. (2015) 'Factor Analysis of Rent Using Space Syntax Measures: Comparative Analysis by Building Use around Shibuya Station', Yasushi Asami, Yoshiro Higano, Hideo Fukui (eds), Frontiers of Real Estate Science in Japan, New Frontiers in Regional Science: Asian Perspectives,29(16), pp.237-257.

Ota, A. (2016) 'A Study on Factors of Prosperity in Nagoya CBD by Using Space Syntax Measures', Doctoral thesis (Doctor of Engineering), Nagoya Institute of Technology.

Ota, A. et.al. (2017) 'Analysis of the correlation between underground spatial configurations and pedestrian flows by using space syntax measures: A case study of the Sakae district underground mall complex', 15th International Conference on Computers in Urban Planning and Urban Management, Adelaide, Australia. 11-14 July 2017.

Ota, A. et.al. (2018) 'A Data Analysis Study on Factors of the Pedestrian Flows in Two Different Underground Malls Using Space Syntax Measures: Case Comparisons in Nagoya, Japan', Innovative Approaches in Agent-Based Modelling and Business Intelligence, pp.173-187.

Ota, A. and Kaneda, T. (2020) ‘A Comparative Study of Factors of Land Price Index by Space Syntax Measures in Nagoya CBD Between 1935 and 1965: Case Study on Nagoya Downtown Area During Pre- and Post-war Period’, Kaneda Toshiyuki (eds), Agent-Based Social Systems book series vol.16, Downtown Dynamics, pp.85-102.

Özer, Ö. and Kubat, A. (2007) ‘Walking initiatives: a quantitative movement analysis’, Paper presented at the 6th international space syntax symposium, Istanbul, Turkey,12-15 June 2007.

Penn, A. and Turner, A. (2001) ‘Space syntax based agent simulation’, In: Schreckenberg, Michael and Sharma, Som Deo, (eds.) Pedestrian and Evacuation Dynamics, pp. 99-114.

Sharmin, S. and Kamruzzaman, Md. (2018) ‘Meta-analysis of the relationships between space syntax measures and pedestrian movement’, Transport Reviews, 38(4), pp.524-550.

Shimizu, K., Nishi, H. and Kishimoto, T. (2019) ‘Estimation of pedestrian density and speed on street network using smartphone spatio-temporal data’, Paper presented at 12th International Space Syntax Symposium, Beijing, China, 8 -13 July 2019.

Stavroulaki, G. et.al. (2019) ‘Statistical modelling and analysis of big data on pedestrian movement’, presented at 12th International Space Syntax Symposium, Beijing, China, 8 -13 July 2019.

Turner, A. and Penn, A. (2002) ‘Encoding natural movement as an agent-based system: an investigation into human pedestrian behaviour in the built environment’, Environment and planning B: Planning and Design, 29(4), pp.473-490.

Uyar, G., et.al. (2017) ‘Using an analytical process to contextualise architectural prototypes for metro stations’, Paper presented at 11th International Space Syntax Symposium, Lisbon, Portugal,3-7 July 2017.