都心部における人流変容メカニズムの解明と予測モデルの構築

Changes in visitor behaviour across COVID-19 pandemic: Unveiling urban visitation dynamics and non-linear relationships with the built environment using mobile big data

発表者・研究代表者　蕭 耕偉郎（東京大学・准教授）

【研究データエコシステム構築事業シンポジウム2024】2024年10月8日（火）@一橋講堂

Lang Yuan, Kojiro Sho *, Sunyong Eom, Hayato Nishi, Daisuke Hasegawa, Han Zhao, Takashi Aoki, Jiarui Zhu, Kaoru Matsuo, Akinobu Masumura

1.Introduction

- urban visitors are broadly defined as people coming from outside the city for purposes such as tourism, work, or services.
- Urban visitors are vital to a city's economy, society, and culture. During COVID-19, their activity was severely disrupted. This underscores their key role in urban recovery post-pandemic.
- This study explores urban visitor behavior before, during, and after COVID-19, analyzing the nonlinear relationship with built environment factors and capturing dynamic changes to inform urban resilience.

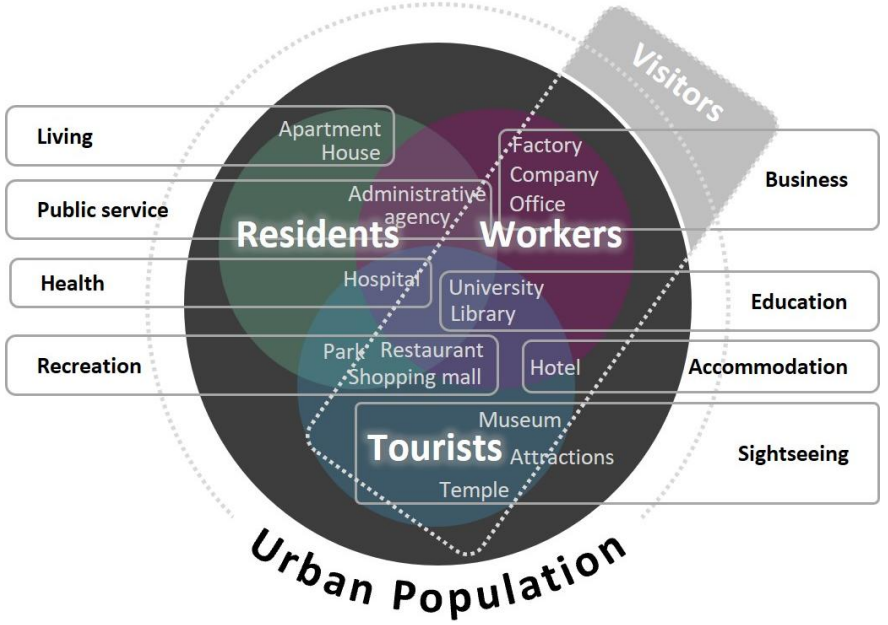


Fig. Definition of urban visitors

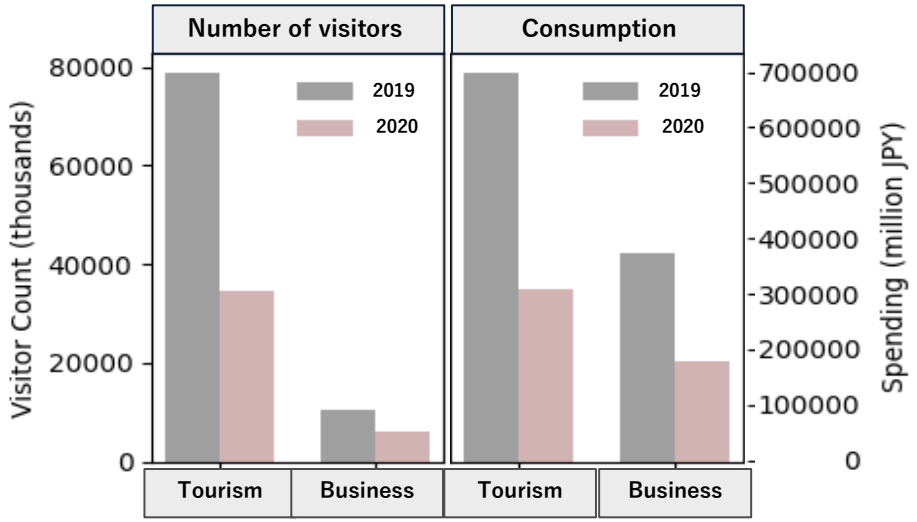


Fig. Transition of visitors in Fukuoka City(福岡県観光入込客推計調査, 2022)

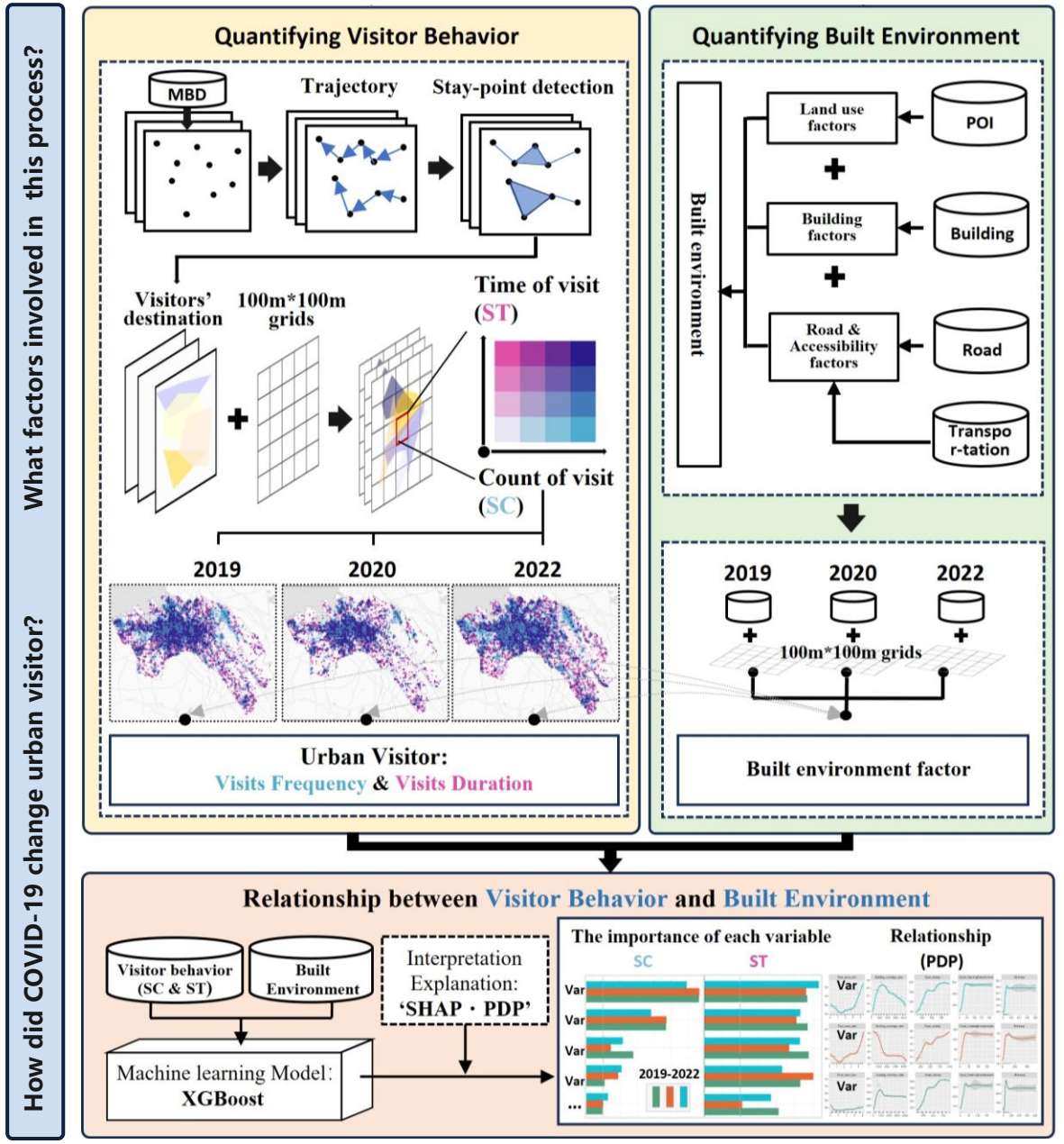
2. Methodology

2.1 Research Framework

- **Step1. Quantifying visitor behavior using mobile big data (携帯位置情報ビッグデータ)**

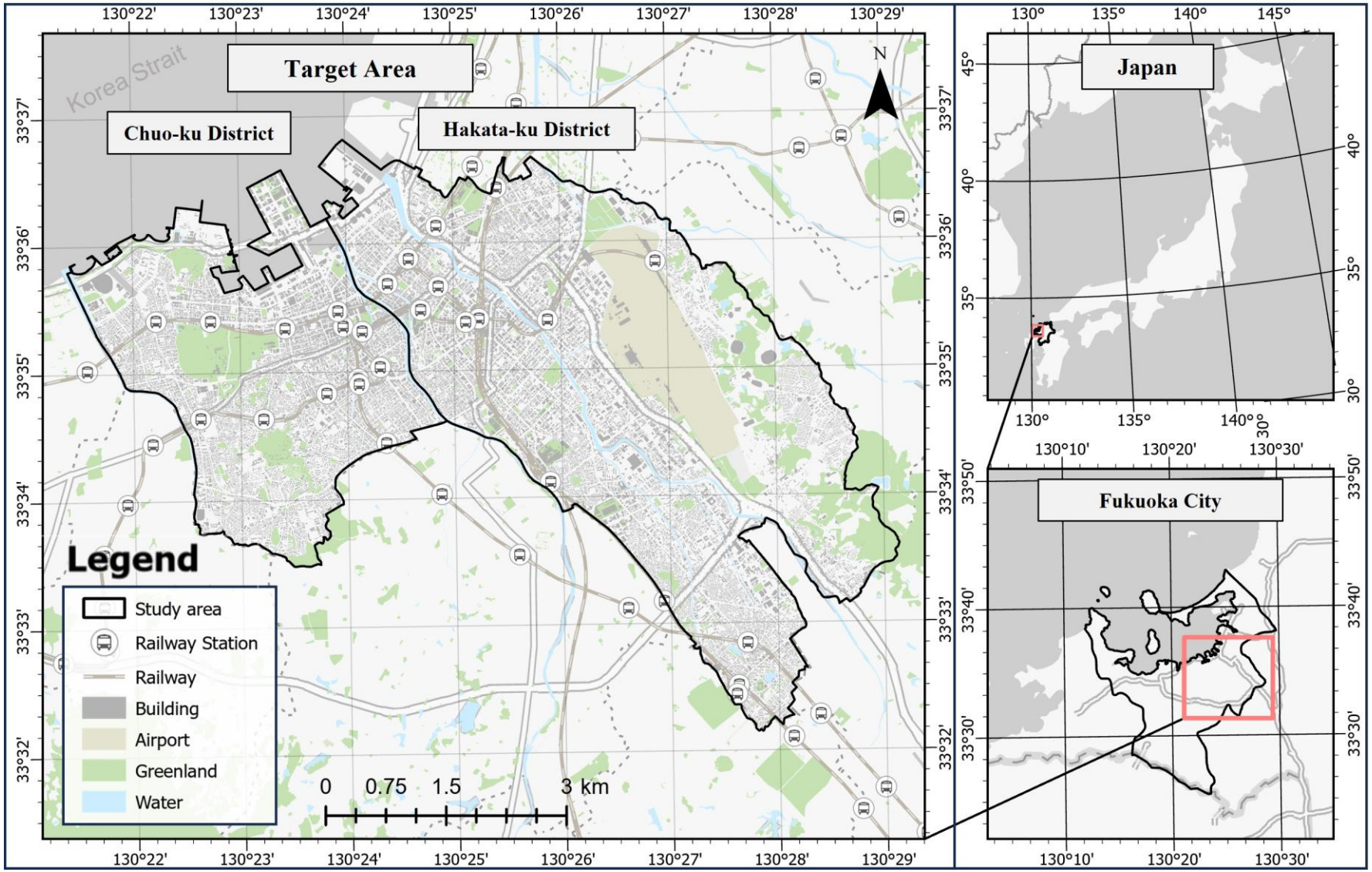
- **Step2. Measuring built environment from Land use, Building, Road & Accessibility.**

- **Step3. Investigating the relationship between visitor behavior and built environment and the changes**



2.2 Study Area

■ Chuo-ku and Hakata-ku of Fukuoka City



2.3 Data collection

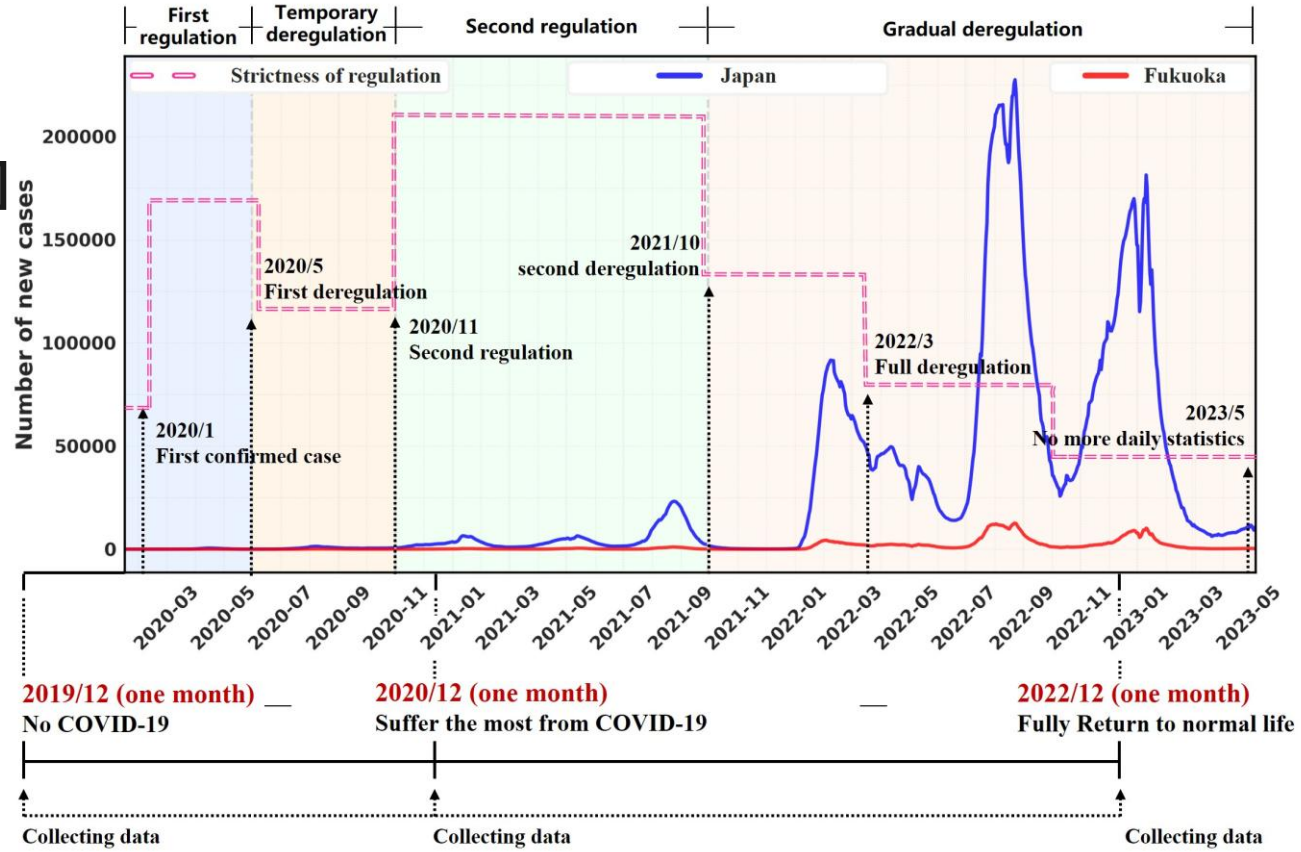
- Five types of data will be collected to represent visitors and quantify the built environment

- Data collection covers December of 2019, 2020 and 2022, representing before, during and after COVID-19

Table. Data source

	Data	Name	Sources
Urban visitor			
	Mobile phone big data	Point-type floating population data (Agoop ポイント型流動人口データ)	Agoop Corp.,2023
Built environment			
	POI data	Tele-Point-Pack! ゼンリン座標付き電話帳 DBテレポイント	ZENRIN InterMap Inc., 2023
	Road data	OpenStreetMap data (OSM data)	GEOFABLIC GmbH., 2023
	Building data	Current status map by building use (福岡令和３年度建物現況)	Fukuoka Government
	Land use data	Current Land Use Map (福岡令和３年度土地利用現況)	Fukuoka Government

Fig. Changing trend of COVID-19 cases and counter policy



2.4 Methods → Quantify visitor behavior through mobile phone big data

- **1.** **Local users** in Fukuoka City were excluded to focus on visitors.
- **2.** Each user's **trajectories** was mapped, and time spent was measured.
- **3.** Locations where visitors stay over 5 minutes within a 100-meter radius were marked as "**Stay points**" and the convex hull of stay point were marked as "visitor destination.

Step1. Criteria of Visitors' Destinations detection

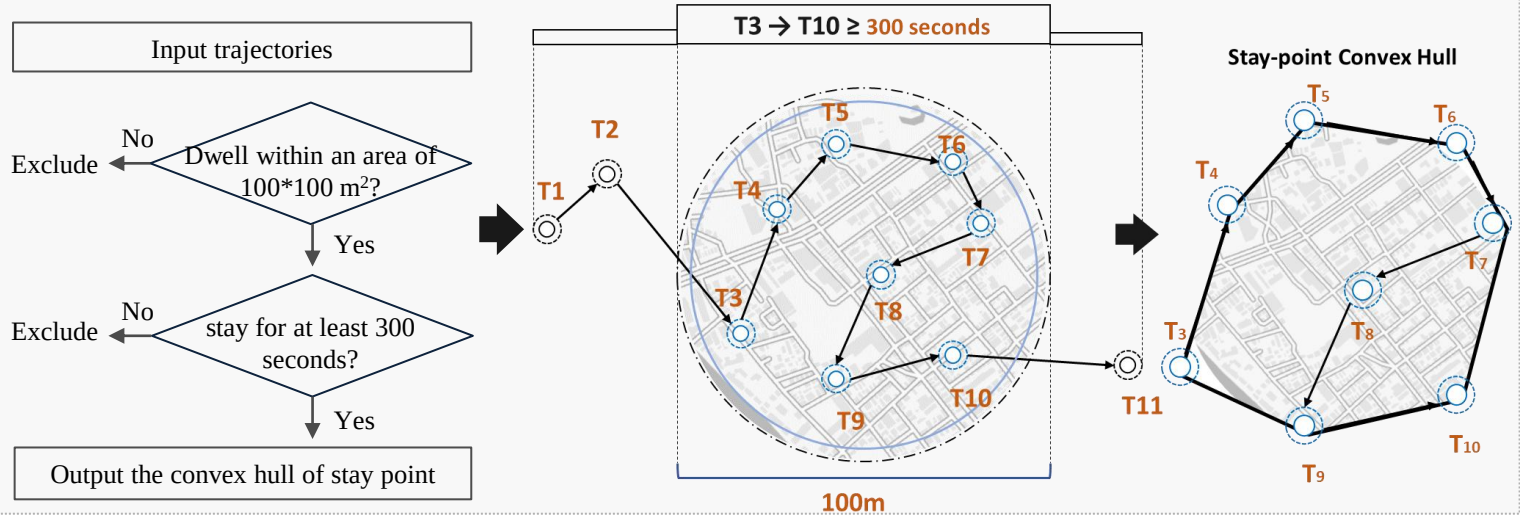
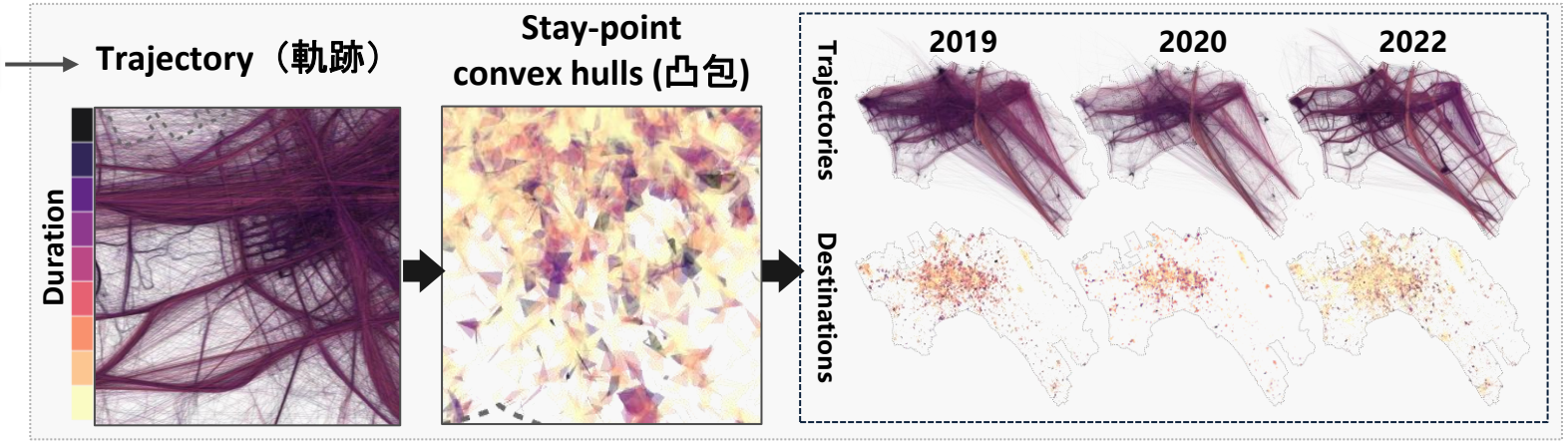


Table. Examples of Agoop MBD (人流データ)

Attributes	point 1	Point 2	Point 3
dailyID	060***72a2	060***72a2	060***72a2
latitude	33.583094	33.590097	33.59043
longitude	130.395348	130.422244	130.422136
home_country	Japan	Japan	Japan
home_citycode	42201	13123	42202
workplace_citycode	42201	14104	42202
Age	under19	20-29	50-59
time	2022/12/8 12:37	2022/12/8 12:40	2022/12/8 12:51

Step2. Generating Visitors' Destinations from users' Trajectories



*The collected data set contains 2,564,218 records from 2019, 1,016,233 records from 2020, and 6,612,239 records from 2022. There are 1,016,233 records from 2020 and 6,612,239 records from 2022.

2.4 Methods → Quantify visitor behavior through mobile phone big data

- The visit frequency (SC) within grid i is denoted SC_i as follows:

$$SC_i = n_i$$

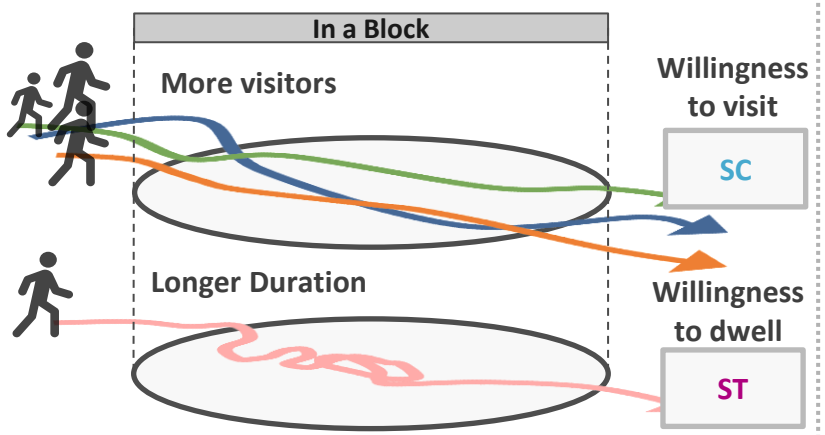
where n_i is the count of the stay-point convex hulls in grid i.

- The average visit duration (ST) of each grid was calculated as follows:

$$ST_i = {\sum_{j=1}^{n_i} T_{ij} w_{ij}}/{n_i}$$

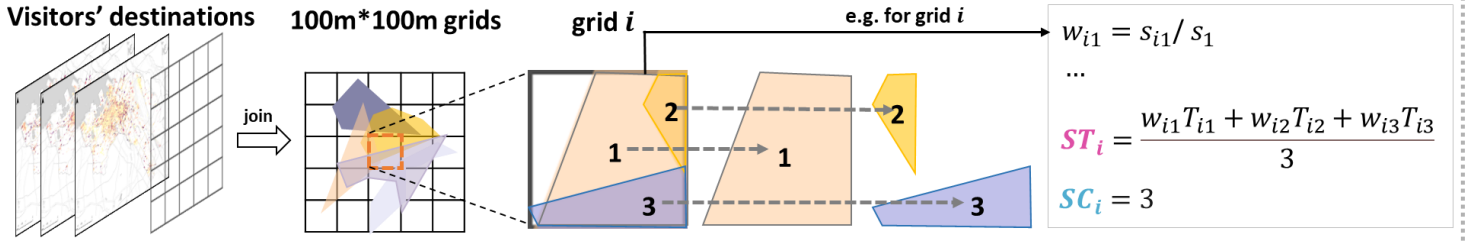
where T_{ij} is the visit duration of the j_{th} convex hulls in grid i, $w_{ij} = {S_{ij}}/{S_j}$. $T_{ij}w_{ij}$ is the weighted visit duration.

How to evaluate visitor behavior?



Step.3 Quantification of SC , ST

$$SC = n_i \qquad ST = \frac{\sum_{j=1}^{n_i} T_{ij} w_{ij}}{n_i}$$



T_{in} is the n_{th} convex hulls's time in block i, S_{in} is the area of the n_{th} convex hulls within this block.

2.4 Methods → Explore the relationship between visitors and the built environment

- We selected **19 indicators** to quantify the built environment as explanatory variables, in terms of land use, buildings, roads, and accessibility. And we used SC, ST as target variables, respectively.
- The **XGBoost machine learning model** was used to investigate the relationship. Built environment factors were the explanatory variables, while SC and ST were the dependent variables.
- To address this, **SHAP values (SHapley Additive exPlanations)** and **PDP (Partial Dependence Plots)** were used to explain the model's results.

Indicator	Description	Data	Unit	Mean*	SD*
Dependent variables					
SC	the visit frequency in each grid	MBD	count	81.06	207.43
ST	the average visit duration in each grid.	MBD	seconds	3443.88	5222.06
Explanatory variables					
land use factor					
Richness	Total number of POI type in the grid	POI data	count/10000m^2	10.06	27.57
SHDI* (Shannon's Diversity Index)	Degree of mixing of POIs in the grid*	POI data	count/10000m^2	0.56	0.63
Accommodation	Count of Accommodation POI in the grid	POI data	count/10000m^2	0.19	0.98
Business & Profession	Count of Business & Profession POI in the grid	POI data	count/10000m^2	2.04	6.67
Convenience Stores	Count of Convenience Stores POI in the grid	POI data	count/10000m^2	2.00	6.41
Catering & Entertainment	Count of Catering & Entertainment POI in the grid	POI data	count/10000m^2	0.33	1.27
Education	Count of Education POI in the grid	POI data	count/10000m^2	1.69	7.28
Leisure & Sightseeing	Count of Leisure & Sightseeing POI in the grid	POI data	count/10000m^2	0.07	0.35
Hospital	Count of Hospital POI in the grid	POI data	count/10000m^2	0.71	2.13
Industry	Count of Industry POI in the grid	POI data	count/10000m^2	1.03	2.78
Public Service	Count of Public Service POI in the grid	POI data	count/10000m^2	1.73	17.22
Shopping Mall	Count of shopping mall POI in the grid	POI data	count/10000m^2	0.17	1.54
Sports	Count of Sports POI in the grid	POI data	count/10000m^2	0.10	0.44
Building factor					
Building coverage ratio	Ratio of building footprint to grid area	Building data	m^2/10000m^2	2470.62	1543.24
Floor area ratio	Gross floor area divided by the gross building footprint in the grid	Building data	-	36.18	1069.38
Road and Accessibility factor					
Road density	Ratio of total road length in grid area	Road data	m/10000 m2	434.79	328.58
Intersection density	Ratio of number of road intersections in grid area	Road data	count/10000m^2	0.18	0.69
Bus stop density	Ratio of number of bus stops in grid area	Road data	count/10000m^2	0.08	0.33
Railway station density	Ratio of number of railway station in grid area	Road data	count/10000m^2	0.01	0.09

3. Results and Conclusion

3.1 Results 1→ Spatio-temporal patterns of visitor behavior during COVID-19

■ **SC and ST Distribution**:

Figure 7 shows SC (visit frequency) and ST (average visit time) distribution for 2019, 2020, and 2022, with pink for ST and blue for SC.

■ **Visitor Concentration**:

Key areas include Tenjin, Hakata, Nakasu, transport hubs (e.g., Hakata Station), and new complexes like Paypay Dome and LalaPort Mall.

■ **SC and ST Differences**:

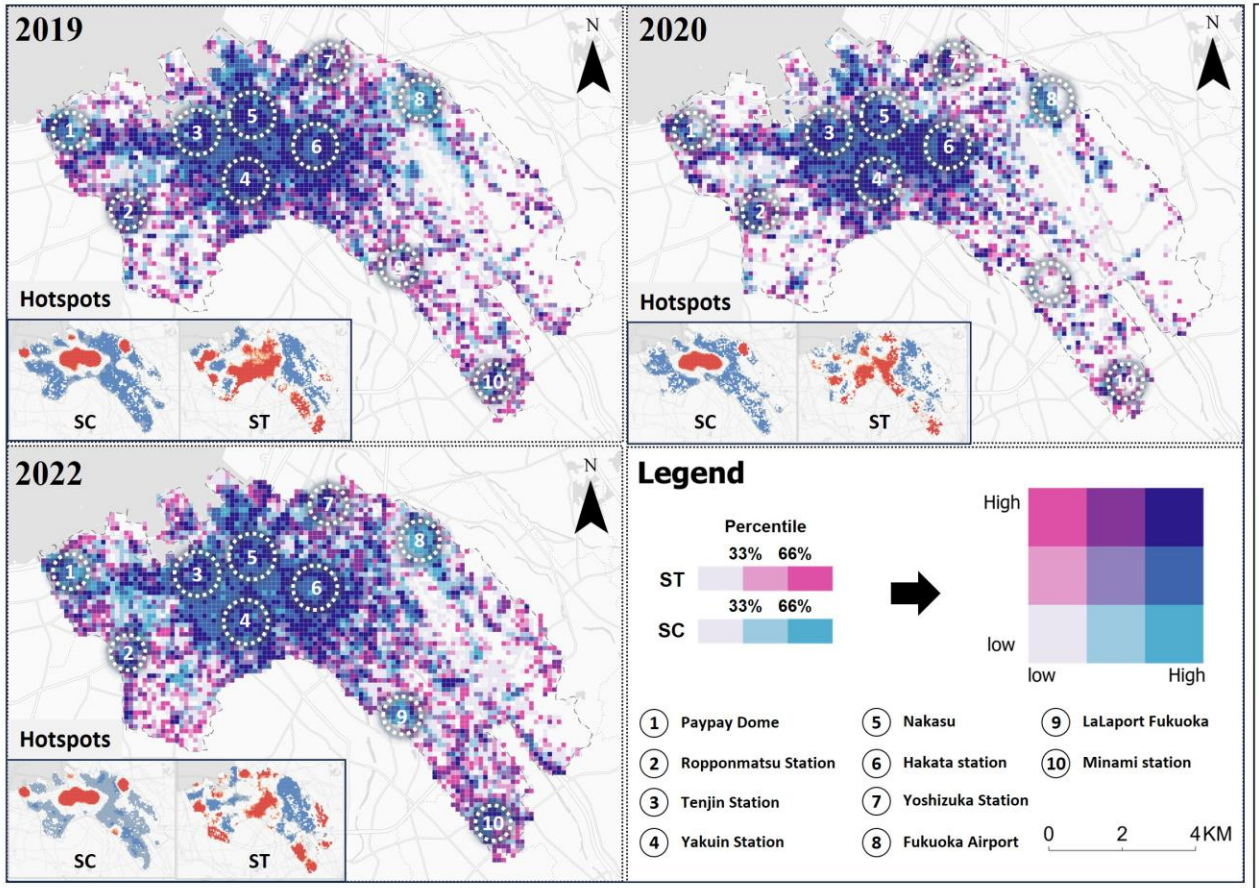
SC hotspots are around Tenjin, Nakasu-Kawabata, and Hakata stations. ST hotspots focus on southern Tenjin, Yakuin, and Ohori Park.

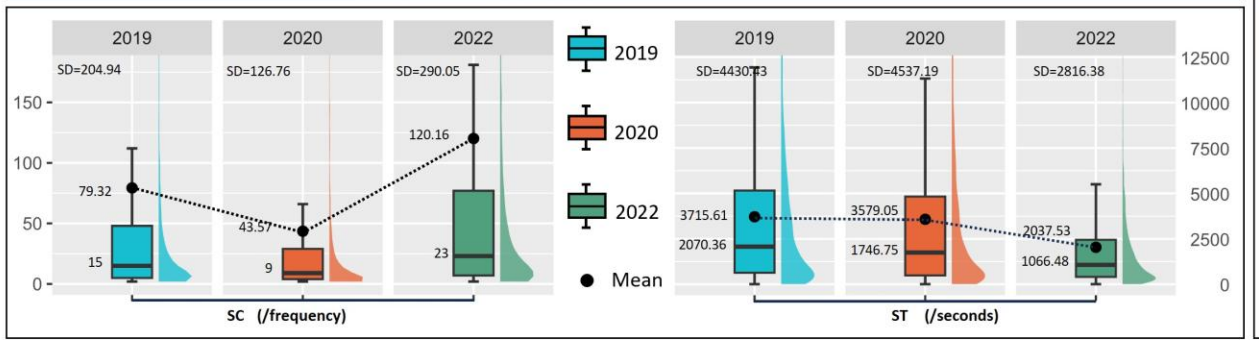
■ **Hotspot Trends**:

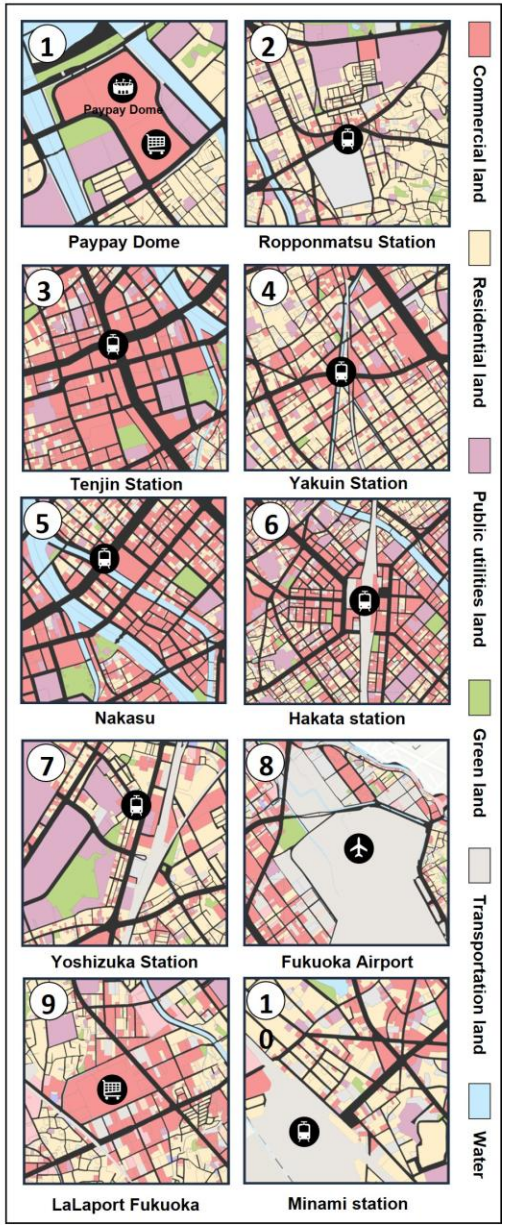
-2019: Paypay Dome and Fukuoka Conference Center were hotspots.

-2020: SC hotspots mostly disappeared; ST shrank by 28.6%.

-2022: SC hotspots recovered, LaLaPort Mall became a new hotspot, and ST further dispersed.





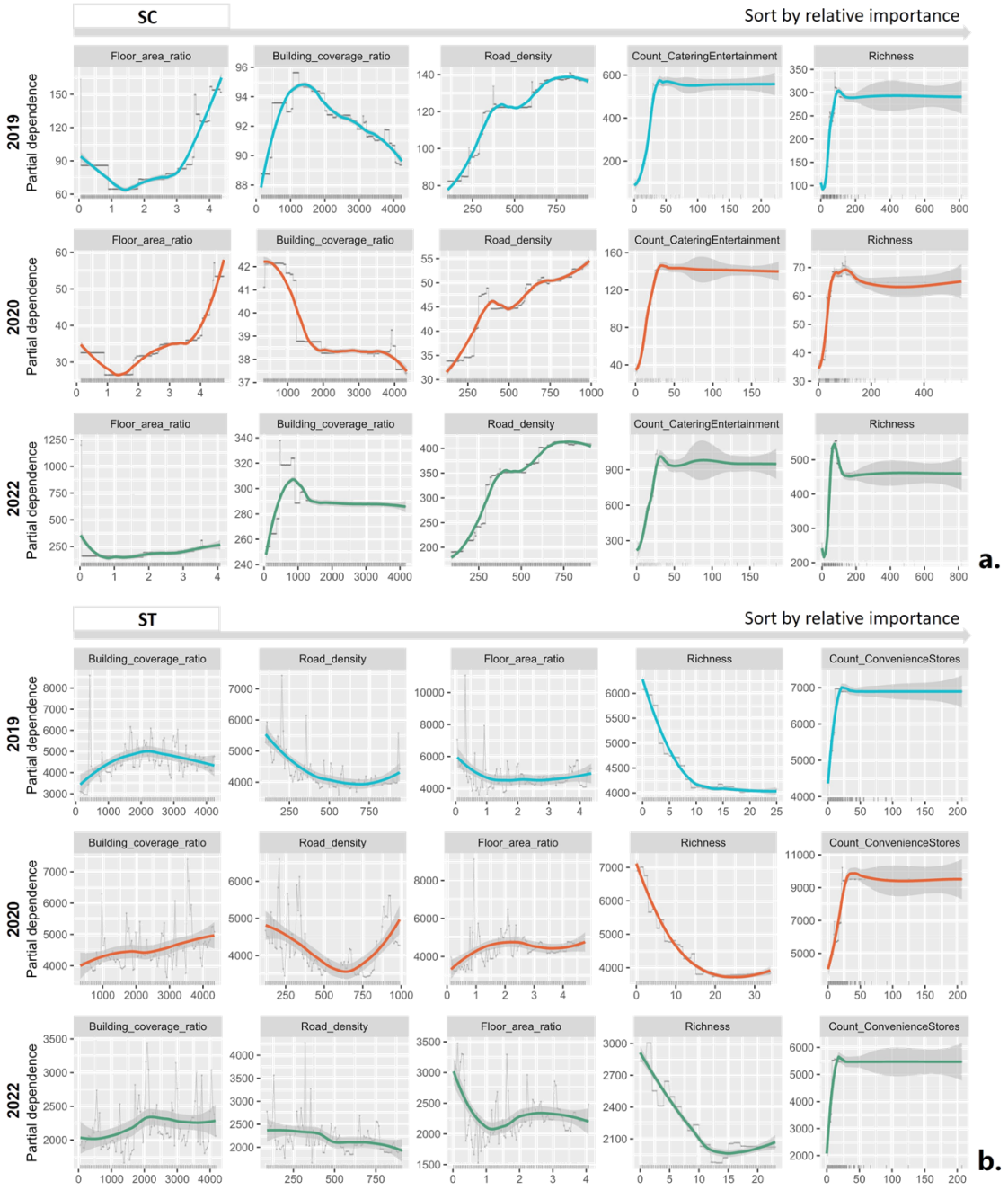


3. Results and Conclusion

3.1 Results 2 → The relationship between visitor behavior and built environment factors

- To clarify the results, PDPs were used. The PDPs reveal how key factors influence SC and ST by showing the effect of changing one variable while controlling for others, especially in nonlinear cases.

- **SC Relationship with the built environment (Fig. 10a)**
  - **Floor Area Ratio** : Negative correlation with SC below 1, positive above 2, weaker in 2022.
  - **Building Coverage Ratio** : Positive correlation below 1000 sqm/10,000 sqm, negative above; reversed in 2020.
  - **Road Density, Catering&Entertainment POIs, and POI Richness** : Positive correlation with SC, but diminishing returns.
- **ST Relationship with the built environment (Fig. 10b)**
  - **Building Coverage Ratio** : Negative correlation with SC below 1, positive above 2, weaker in 2022.
  - **Floor Area Ratio** : Positive correlation below 1000 sqm/10,000 sqm, negative above; reversed in 2020.
  - **POI Richness and Convenience Store POIs**: POI diversity shows a strong negative correlation with ST, while convenience store POIs have a positive correlation.



Thanks for your listening.