# Nighttime Vitality and Its Relationship to Urban Diversity: An Exploratory Analysis in Shenzhen, China

Junwei Zhang ⓘ, Xintao Liu ⓘ, Xiaoyue Tan, Tao Jia, Ahmad M. Senousi ⓘ, Jianwei Huang ⓘ, Ling Yin ⓘ, and Fan Zhang

*Abstract*—Relationship between urban diversity and urban vitality is imperative for guiding better design in urban development, although existing frameworks are not able to efficiently examine the relationship at multiple scales. In this article, we propose a new framework to integrate nighttime light (NTL) imagery and multisource urban data into multiscale geographically weighted regression (MGWR) models to examine the varying relationship between diversity and vitality across space and time. NTL is used as a proxy for urban nighttime vitality. Public transport, taxi transit, and points of interest data are used to derive three aspects of urban diversity indices: ridership diversity, spatial interaction diversity, and built environment diversity. By comparing the models in holiday and nonholiday weeks in Shenzhen, China, the NTL-based vitality proxy was found to be strongly correlated with the urban diversity indices, given by the satisfactory goodness of fit (*r*-squared = 0.9) of the MGWR models. The spatially varying relationships between diversity indices and nighttime vitality were observed and patterns discussed. The analysis of the coefficients revealed the importance of stable public transport and fluctuating taxi trips for nighttime vitality. The new index proposed for the diversity of spatial interaction (DSI) is a strong indicator for nighttime vitality, adding to existing vitality indicators. Furthermore, this study found that DSI and density of catering have less temporal variation, indicating their robustness in measuring nighttime vitality.

Junwei Zhang and Xiaoyue Tan are with the Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong (e-mail: lsjacob.zhang@connect.polyu.hk; xiaoyue.tan@connect.polyu.hk).

Xintao Liu is with the Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong and also with the Research Institute for Smart Cities, The Hong Kong Polytechnic University, Hong Kong (e-mail: xintao.liu@polyu.edu.hk).

Tao Jia is with the School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430072, China (e-mail: tao.jia@whu.edu.cn).

Ahmad M. Senousi is with the Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong and also with the Geomatics Engineering Lab, Civil Engineering Department, Faculty of Engineering, Cairo University, Cairo 12613, Egypt (e-mail: ahmad.m.ahmad@connect.polyu.hk).

Jianwei Huang is with the Institute of Space and Earth Information Science, The Chinese University of Hong Kong, Hong Kong (e-mail: jianwei.huang@link.cuhk.edu.hk).

Ling Yin and Fan Zhang are with the Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China (e-mail: yinling@siat.ac.cn; zhangfan@siat.ac.cn).

Digital Object Identifier 10.1109/JSTARS.2021.3130763

This study provided empirical insights into how nighttime vitality is related to urban diversity, demonstrating new applications of NTL for intracity studies.

*Index Terms*—Multiscale geographically weighted regression (MGWR), nighttime light (NTL), spatiotemporal variation, urban diversity, urban vitality.

## I. INTRODUCTION

URBAN vitality has become an essential concept in assessing the quality of urban development [1]–[3]. It addresses a vigorous figure of socioeconomic activity and abundant qualitative research has indicated its strong ties with "diversity" in urban design [4]–[6]. In recent decades, developing quantitative measures of urban vitality has become more important for evidence-based evaluation, as the lack of empirical observations of city dynamics could lead to serious planning failures. These failures could, for example, result in ghost cities—large and well-built residential areas within which few people live [7]. It is still urgent to adopt the state-of-art method and integrate multisource urban big data to quantify urban vitality and its relationship to key factors in urban diversity.

Nighttime light (NTL) remote sensing imagery explicitly captures the spatial distributions of artificial radiance from human settlements [10]. After removing the background noises, it is generally agreed and reported that the light on the ground surface is significantly related to the density of population, infrastructures, and socioeconomic activity over space and time [9]–[11]. NTL data are widely applied by the scientific community in various research fields, such as socioeconomic development assessment [12]–[14], conflict and disaster detection [15], energy consumption estimation [16], and light pollution assessment [17]. The recent development of NTL composites equipped with Visible Infrared Imaging Radiometer Suite (VIIRS) provides higher spatial–temporal resolution compared to previous technologies [18]. VIIRS data are widely documented as an efficient representation of socioeconomic activities [10], [19]. Light intensity extracted from VIIRS imageries is a good source to capture human activities, as recent studies have demonstrated its strong correlation to other human activity data, such as social media check-ins in both regular and holiday time [20], [21]. Benefit from high temporal resolution (daily) of VIIRS data, most recent study also applies it in assessing the impact of

Covid-19 pandemic on urban dynamics [22], [35]. Although existing literature show that NTL captures urban dynamics efficiently, the use of such light intensity for a further application requires a deeper understanding of the relationship between NTL and domain-related factors (such as economic or environmental). However, very few studies have discovered how NTL is associated with factors in the context of evaluating urban vitality. One study compared NTL with their urban vitality model developed by other urban big data [8], while as mentioned, the relationship between NTL and vitality-related factors is barely unknown. This gap motivates our study to systematically examine the relationship between NTL and diversity-related factors, thus aiming to answer whether and how NTL can be applied in urban vitality research.

Relying on recent advent of urban sensing data, previous studies quantified urban vitality from two main perspectives. One stream of studies developed environment-based indicators, such as infrastructure density, land use diversity, and catering density, which, in theory, have the potential to improve urban vitality [7], [23], [24]. The main drawback of environment-based indicators is that the result is relatively static over time, which could fail to reflect how human activity may vary temporally, even if a place is well built. Another stream of studies has developed mobility-based indicators to directly represent vitality, such as travel intensity and temporal variation [25]. The main drawback of this approach is that it may fail to illuminate the factors to explain why a place has vitality, which can make it difficult to make further insights and develop policies to enhance specific types of urban environments. The research gap in these early vitality research indicates that integrative evaluation considering both human activity and urban environment data is necessary, which also motivates our study to integrate NTL with other urban big data in vitality evaluation.

The drawbacks of single-type indices have inspired a growing body of research to integrate the urban environment and human mobility to capture more aspects of urban vitality [26]–[30]. Such a scheme produces models not only to quantify urban vitality but, more importantly, to understand the relationship between a vitality proxy and environment indicators [26], [29]. However, most existing methods for constructing relational models do not consider spatial heterogeneity very well [30] if using a linear regression model. To examine this locally varying relationship, geographically weighted regression (GWR) can be used to develop an empirical model, but very few studies have adopted GWR in urban vitality evaluation [31]. More importantly, GWR-based vitality evaluation face methodological drawbacks of capturing multiscale characteristics of different factors due to the use of fixed bandwidth. To fill this methodological gap, the recent development of GWR, multiscale geographically weighted regression (MGWR) [32], is a promising model to provide a deeper understanding of the relationship between vitality and key indicators.

This article proposes a new framework to examine the spatially varying relationship between the proxy of urban vitality and vitality indices. The proposed framework contributes to the vitality literature by combining new remote sensing data (NTL) with social sensing data (e.g., urban environment and travel card records) via an improved technique, i.e., MGWR. This framework also makes contributions to NTL literature on how NTL intensity is associated with different aspects of urban diversity, extending the application field of NTL data. Comparative experiments are conducted between regular and holiday weeks in Shenzhen, China. Further explanation of methods and results are given in the following sections.

## II. STUDY AREA AND DATASET

### A. Study Area and Time

This study selects Shenzhen, a southern city of mainland China, as the study area (see Fig. 1). Covering 2000 km$^2$ with a population of 15 million, Shenzhen is one of the top-tier cities in China. Its special geographic location on the north side of Hong Kong coupled with China's open-door policy has proved advantageous for Shenzhen's rapid urban and economic development. The prosperity of the city has attracted many migrants from all over the country, many of whom are now becoming permanent residents. These characteristics have made the city diverse and vibrant; thus, we choose Shenzhen as a case study to obtain a better understanding of nighttime vitality and its relationship to urban diversity. The results will provide empirical insights into its success for other Chinese cities.

The period for data collection and processing covers three weeks, from January 13 to February 2, 2017. The third week was the Spring Festival holiday week in China, so the first week is considered a regular week and the second week a prefestival week when some migrants leave the city to visit their hometowns. It would be interesting and significant to see how vitality models may change in different temporal contexts.

### B. Datasets

In recent decades, numerous satellites and platforms have been launched to obtain nocturnal artificial light records. Among them, the Visible Infrared Imaging Radiometer Suite Day/Night band (VIIRS/DNB), carried on the Suomi National Polar-Orbiting Partnership (Suomi NPP) satellite, produces a widely used dataset with higher spatial resolution and better radiometric quality than its predecessors [33], [34]. Nightly VIIRS/DNB imagery offers rapid and timely artificial light observations, which are promising for monitoring daily human activities and socioeconomic dynamics [35]–[37]. Unlike mobile phone data observing on human activity in real time, NTL data represent intensity and dynamics of human activity by capturing spatial configuration of infrastructures (associated with diverse types of urban activity that are not limited in the overpass time of data recording) and change of light intensity (associated with dynamic changes of human behavior and local context). Therefore, the overpass time of VIIRS (1:30 A.M.) does not influence much for proxying night vitality because the vitality of the city is not the instant activity given a time snapshot, instead, is associated with socioeconomic activities. This research uses the processed pixel values as a proxy of urban vitality at local locations. The data processing methods will be introduced in Section III.
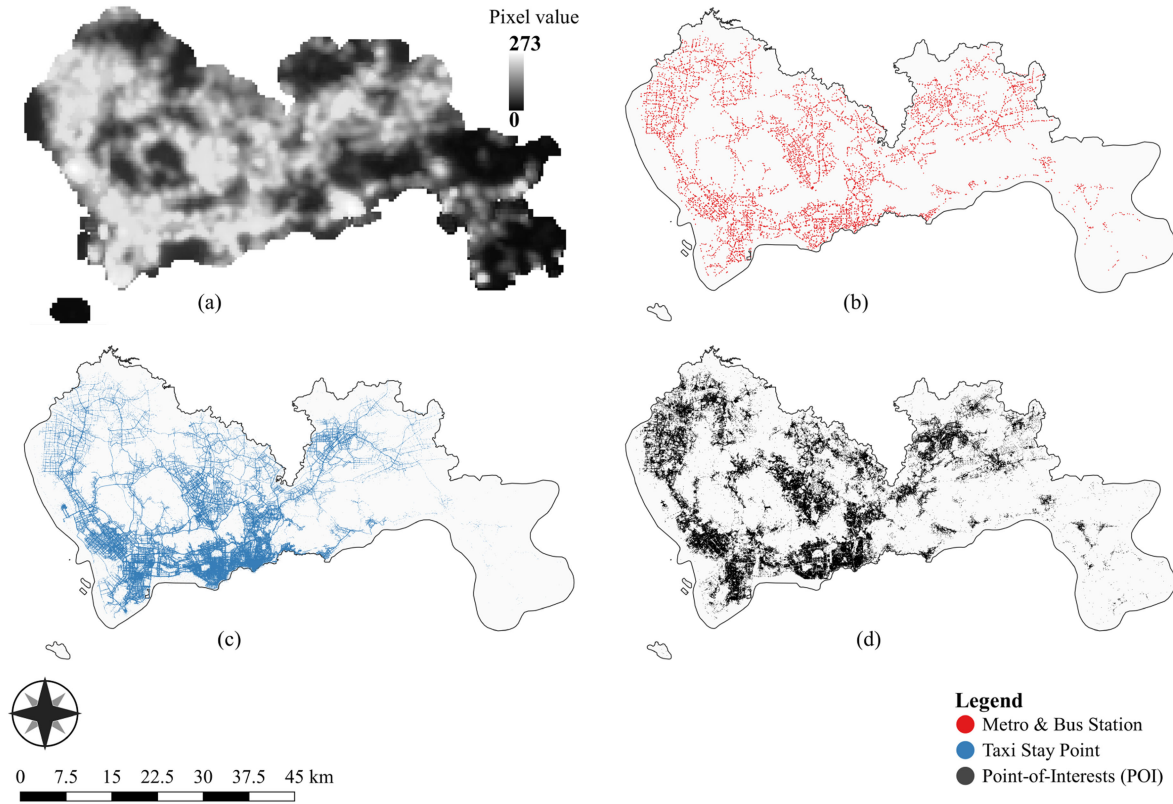
Fig. 1. Study area and spatial distribution of datasets. (a) NTL imagery. (b) Stations of public transport. (c) Pick-up and drop-off Locations of taxi. and (d) Point of Interests (POI).
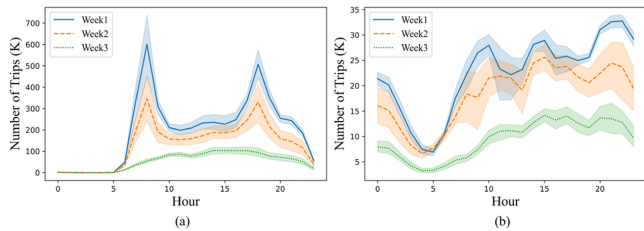


Fig. 2. Hourly variation of trips. (a) Public transport. (b) Taxi. Note that week 3 is the holiday week, week 2 is the migration week when people may leave Shenzhen to go back to their hometowns, and week 1 is regular week.

Public transport data (i.e., metro and bus trips), taxi transit data, and POIs are the other sources of urban big data used in this study. Anonymous public transport and taxi data are remotely accessed from our research collaborators' server at the Shenzhen Institute of Advanced Technology. At original places, transit data are collected by pricing machines deployed at metro/bus stations and in taxis. The origin–destination (OD) locations and timestamps are recorded for each trip, and travelers' IDs are desensitized. The temporal resolution of transit data in this study is at the minute level. NTL data capture the overall vitality across the city, and transit data capture more detailed patterns of trip intensity, temporal variation, and spatial interaction. In Fig. 2, the aggregated transit volumes show the obvious differences between different transport modes in different weeks. This demonstrates that the weeks in the selected period represent

different temporal contexts and that the multisource transit data capture diverse aspects of traveling behavior within the city.

POI data are used to derive indices for the urban built environment relating to urban vitality. The high spatial resolution and wide coverage of the POIs make this data unique for quantifying the relevant aspects of the urban environment [38], [39]. In this study, we develop a web crawler program and 479220 POIs in Shenzhen are collected from Amap, a major map service provider in China. The main attributes gathered from the POIs for use in this study are the coordinates and service type of each entity (see Fig. 3).

The spatial units of this study are grid cells with a size of 500 × 500 m, and the coordinate system used is the Xian 1980 cartesian system. The grid cells are used as the unified unit to extract and align indices from different data and serve as the entry points in the MGWR model. The 500-m size is in accordance with the spatial resolution of the NTL data used in this study. Furthermore, the size is suitable for urban vitality evaluation because it is considered an acceptable walking distance from transport stations, which enables residents to participate in activities, thus promoting the vitality of the city [4], [40]. The temporal scheme for calculating and aligning indices and calibrating MGWR models is weekly. We choose this timeframe for two reasons. First, the vitality proxy represented by the maximum value of NTL data of each week (7 days) is more reliable than a single-day observation. Second, the calculation of trip variation requires an input of time-series data in consecutive days so that we evaluate
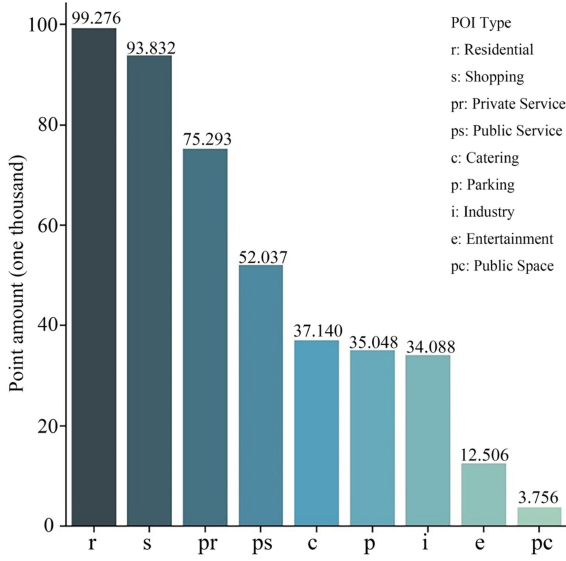
Fig. 3.    Distribution of POI types.

it every 7 days for alignment with vitality proxy. These indices will be introduced in more detail in the following sections.

## III. METHODOLOGY

### A.  Describing MGWR Model

MGWR extends the traditional regression model to effectively depict the local relationship between explanatory variables and a response variable. Similar to the GWR model, MGWR enables the inference of spatially varying coefficients rather than global coefficients. As opposed to GWR, the MGWR used in this study is more suitable as it allows for the consideration of different spatial scales when inferring local coefficients for different explanatory variables [32]. Several studies have also suggested that the flexible scales cause MGWR to overperform GWR in terms of handling collinearity issues and increasing the overall model performance [41], [42]. In this study, we conduct comparative experiments to examine whether the MGWR-based vitality model is better than GWR and global linear regression. For this, the mgwr module of the pysal python package is used to implement and infer the specific vitality models.

A general MGWR model can be expressed as

$$y_i = \sum_{j=0}^{m} \beta_{bwj} \left(u_i v_i\right) x_{ij} + \varepsilon_i \tag{1}$$

where $x_{ij}$ is the $j$th explanatory variable of observation $i$ at location $(u_i v_i)$, $\beta_{\mathrm{bwj}}(u_i v_i)$ is the coefficient of $j$th variable inferred by using the bwj bandwidth, $\varepsilon_i$ is the error term, and $y_i$ is the response variable.

The scale indicates the size of the kernel bandwidth for data borrowing in the process of finding optimal parameter estimates. To obtain adaptive bandwidth, the MGWR treats the model as a generalized additive model (GAM) [43] and proceeds with calibration via a back-fitting algorithm. The $\beta_{\mathrm{bwj}} x_i$ defined in

MGWR is treated as the $j$th additive term $f_i$ in GAM-MGWRs

$$y = \sum_{j=0}^{m} f_i + \varepsilon \,. \tag{2}$$

In the model training process, GWR provides the initial estimates of the parameters. For each iteration, the model fitting process proceeds with the GAM, in which one term, i.e., $f_i$, is considered as a variable while the rest of the terms are constants. The coefficient $\beta_j$ and the optimal bandwidth bwj are obtained by the GWR model. The fitting process continues to iterate through the term $f_i$ and ends at a GAM model that has achieved parameter estimates and bandwidths for all variables. However, the outside iterations continue to generate multiple GAM models. The consecutive GAM models are compared by the change in the GWR smooth function (SOC-f). The whole iteration is terminated when the criterion converges, or more specifically when SOC-f is smaller than $10^{-5}$ between consecutive GAM models.

### B.  Vitality Proxy and Vitality Indices

Some studies have attempted to make compound indices to quantify urban vitality [23], [28], [57], while no "true" formula of urban vitality is supported by all researchers. In this study, we follow the fundamental and most widely used principle to obtain a vitality proxy: more intense socioeconomic activity indicates a place with more urban vitality [8], [24], [30]. NTL provides unique and effective observations of the intensity of socioeconomic activity based on the light of city infrastructure at night [59].

This research uses gap-filled NTL radiance retrieved from the Daily Lunar BRDF-Adjusted Nighttime Lights dataset (Black Marble—VNP46A2) to generate an urban vitality proxy. To reduce the bias introduced by the geographic mismatch of NTL imageries over days, the NTL radiance was averaged using a $3 \times 3$–pixel window as recommended in the literature [44]. However, despite the state-of-art correction strategy used to generate the Black Marble products, uncertainties remain in the NTL time series due to the relatively lower accuracy in nighttime atmospheric correction and cloud detection [45]. Therefore, the weekly maximum value, instead of the average value, of each pixel is extracted as the nighttime vitality proxy. The maximum value is more reliable compared to the weekly average, which is more influenced by clouds. Therefore, the maximum value in this case captures the relative strength of economic activities in different places during the week.

Beyond a vitality proxy, this study focuses on proposing and implementing models to quantify and explain the local relationship between the vitality proxy and vitality indices (Fig. 4).

The selection of vitality indices depends on "diversity," an essential principle of urban design for facilitating urban vitality [4], [46]. This study proposes a new framework to include the following aspects of the diversity of the city:

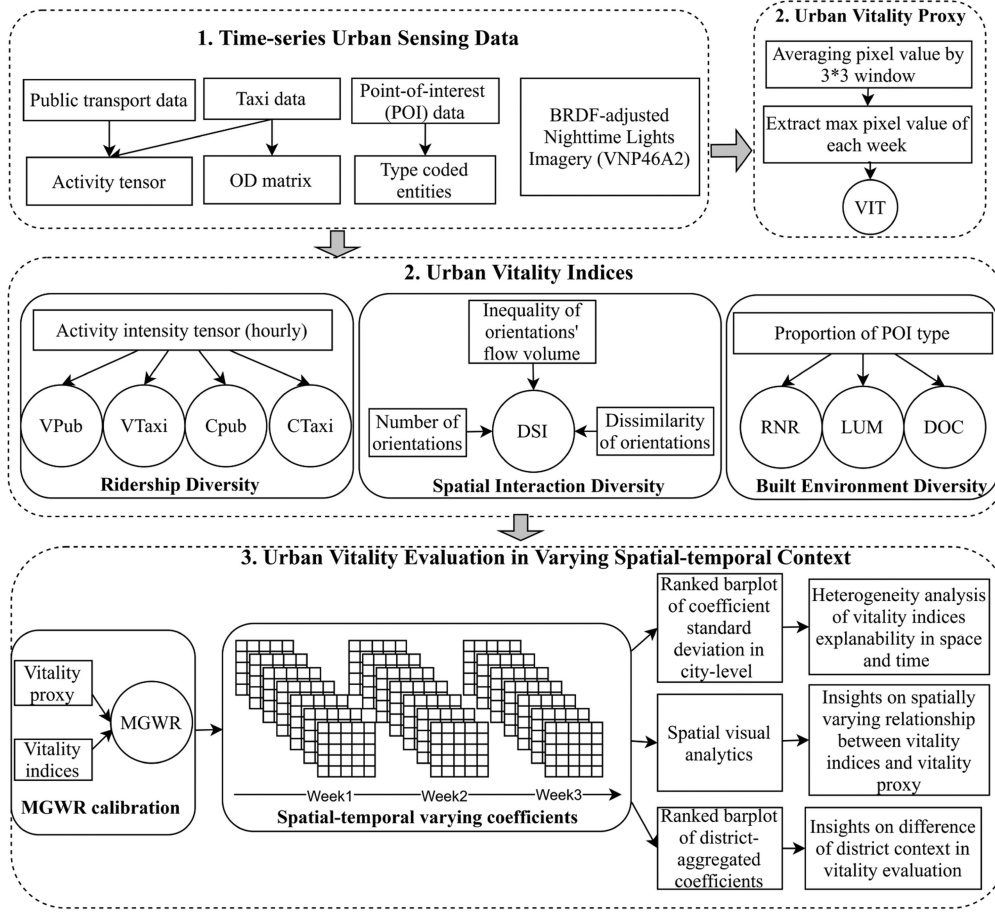1) "Ridership Diversity" that reflects nonregular travel patterns in the time-series ridership volumes;

Fig. 4. Framework of vitality evaluation proposed in this study.

2) "Built Environment Diversity" that reflects the built environment components that facilitate diverse use of the land;
3) "Spatial Interaction Diversity" that reflects how a place attracts trips from diverse places, which is a novel contribution of this study.

Overall, the vitality proxy is the dependent variable, whereas the vitality indices are the explanatory variables in the MGWR-based vitality model (see Table I). The details of the vitality indices are elaborated in the following sections.

### C. Ridership Diversity

In this study, two indices addressing the diversity of the temporal variation of trips are implemented [47], [48]. The main reason for this is that places with higher urban vitality should be associated with more nonregular ridership variation patterns, suggesting that these places attract travelers with different travel purposes. Ridership diversity is measured by variability ($V$) and consistency ($C$), where $V$ measures the day-to-day temporal variation of ridership volume and $C$ measures the within-day nonregularities.

For each spatial unit and week, we calculate the variability of public transport and taxi transit as

$$V = \frac{\Sigma_i^n \Sigma_j^n \left(1 - \text{Corr}\left(\text{TS}_i, \ \text{TS}_j\right)\right)}{\left(n \times (n-1)/2\right)} \qquad (3)$$

where $V$ is the variability index given by a set of time-series data. The numerator is calculated as the summation of dissimilarity ($1 - \text{Corr}$) of every possible time series pair $(\text{TS}_i, \ \text{TS}_j)$, which is then normalized by the number of pairs. $n$ is the number of TS to be compared, in our weekly model, equal to 7 (days). Pearson correlation is used as our Corr() function. A higher $V$ means more nonregular temporal changes from day to day.

The consistency index is used to measure ridership diversity based on the extreme nonregular volume within each day. A higher consistency value indicates more hourly peaks possibly caused by special events, suggesting a place with more a diverse context. In practice, this is implemented based on the outliers of within-day ridership time-series volume using median absolute deviation (MAD) as shown in the following:

$$C = \sum_{j=1}^{N} \left( \frac{\sum_{i=1}^{n} \left| \frac{\text{OutD}_i}{\text{MAD}} \right|}{n} \right)_j \qquad (4)$$

where $C$ represents the consistency index given by the time-series data of $N$ days (equal to 7 in this study). In day $j$, $n$ is the number of outliers, and $\text{OutD}_i$ is the deviation from the $i$th outlier to the median. The larger the difference between the intensity of the activity peak (e.g., morning/evening rush) and the median intensity, the higher the $C$ value will be.

## D. Spatial Interaction Diversity

Spatial interaction has always been an important perspective from which to understand human mobility and urban dynamics [49], though very few studies have implemented interaction indices in urban vitality evaluation [27]. To the best of our knowledge, no study has implemented an index to measure the diversity of spatial interaction (DSI) as a component of the vitality model. To address this gap, we propose and calculate the DSI in our framework.

Spatial interaction can be observed in OD data as people move from place to place. Such interaction in geospace is abstractly similar to the interaction in information space and, in particular, similar to the citations among academic papers. The DSI in geospace is therefore inspired by an index in the informatics field that measures the diversity of interdisciplinary journal citations [50], [51]. We analogize a place (i.e., spatial unit) to a journal, and OD trips to the "citations" among places. The DSI of a place is determined by the trips to the place from different origins, and consists of three index components as shown in the following:

$$\text{DSI} = \text{Varity} \times \text{Balance} \times \text{Disparity} \quad (5)$$

where variety describes the number of places oriented to the destination; for example, mix-use land may attract people from places all over the city, resulting in a large Variety value.

Balance evaluates whether the trip volume is evenly distributed among the oriented places. For example, if a place attracts trips from many origins, but the trip volume from a few places in particular greatly outweighs the others, this is counter to the idea of diverse interaction that requires balanced trip volumes from all places.

Disparity measures the differences embedded in the orientations themselves or, in other words, how these orientations are different from each other. In the city context, a vibrant place is likely to attract trips from different types of originated places. The overall DSI increases when the disparity is higher. It should be noted that we conduct multiple experiments using various disparity functions, such as $1-\text{Cos}$, Euclidean, and geographic distance. Geographic distance is found to be the most effective in terms of the MGWR model performance.

A massive number of taxi OD points in each week provides large-scale observations of spatial interaction, which serve as the major data for calculating DSI. The taxi OD data provide activity points with better spatial coverage and finer spatial resolution than metro OD data, which are limited to the station level. Specifically, we implement DSI as an improved form of the Rao–Stirling index [51]. For each grid and each week, the DSI is calculated as follows:

$$\text{DSI}_d = \left(\frac{n_o}{N}\right) \times \text{Gini}_o * \left(\frac{\sum_{i \neq j}^{n_o} d_{ij}}{n_o * (n_o - 1)}\right) \quad (6)$$

where $\text{DSI}_d$ is the diversity of spatial interaction of grid $d$, $n_o$ is the number of the originating grids, $N$ is the total number of grids in the study area, $\text{Gini}_o$ is the $1-\text{Gini}$ index of the trip volumes from the originating grids, and $d_{ij}$ is the disparity between the $i$th and $j$th originating grids.

## E. Built Environment Diversity

The diversity of the built environment is mentioned in most previous studies on urban vitality evaluation [2], [6], [23], [25]. How a place is planned and built can significantly impact urban vitality by attracting travelers with different purposes [25]. In this study, three indices of the built environment are implemented as explanatory variables for urban vitality: land use mixture (LUM) based on Shannon entropy [52], residential–nonresidential (RNR), and density of catering (DOC) business. The LUM index is calculated by an entropy function as

$$\text{LUM} = -1 \left(\frac{\sum_{i=1}^{n} pi \times \ln(pi)}{\ln(n)}\right) \quad (7)$$

where $pi$ refers to the proportion of the POI types $i$ in the spatial unit and $n$ is the total number of POI types, which equals 9 in this study. The higher the LUM, the higher degree of LUM it indicates.

The RNR index, a ratio index based on the number of POIs, is calculated as

$$\text{RNR} = 1 - \left|\frac{R_i - \text{Non}R_i}{R_i + \text{Non}R_i}\right| \quad (8)$$

where $R_i$ refers to the residential proportion and $\text{Non}R_i$ refers to the nonresidential proportion of each spatial unit. A high RNR value will be obtained when the residential proportion is close to the proportion of nonresidential, which indicates a place that serves both residential and general purposes simultaneously.

Recent studies have associated the DOC index with urban vitality [24], [53]. The reasoning for this index is that catering services are normally established before other businesses because catering is the most basic demand of people. Therefore, a higher DOC indicates more activities, suggesting better urban vitality. The DOC index is calculated as follows:

$$\text{DOC}_i = N(\text{Catering POI})_i \quad (9)$$

where DOC in grid $i$ is the number of catering POIs ($i \in \{1, 2, \ldots, 9009\}$).

## IV. RESULTS

### A. Characteristics of the Vitality Proxy and Indices

Based on the NTL data, the processed pixel values represent the intensity of the urban vitality proxy at specific locations. It should be reminded that indices and models are obtained for three weeks: week 3 was the festival week, week 1 was the normal week, and week 2 was the prefestival week. The spatial distribution of the vitality proxy is shown in Fig. 5. Higher pixel values are demonstrated by the red color and lower values are in green. Overall, the locations with the highest values of the vitality proxy are clustered within districts, such as Nanshan, Luohu, and Futian, revealing the multicenter urban structure. The vitality proxy tends to gradually decrease in value with increased distance from these centers. The hotspot structures appear similar, although the pixel values may change in different manners from week to week.

The temporal variation of the vitality proxy in different grids reveals human dynamics during regular and holiday weeks. For each pixel, we subtract values in consecutive weeks, namely
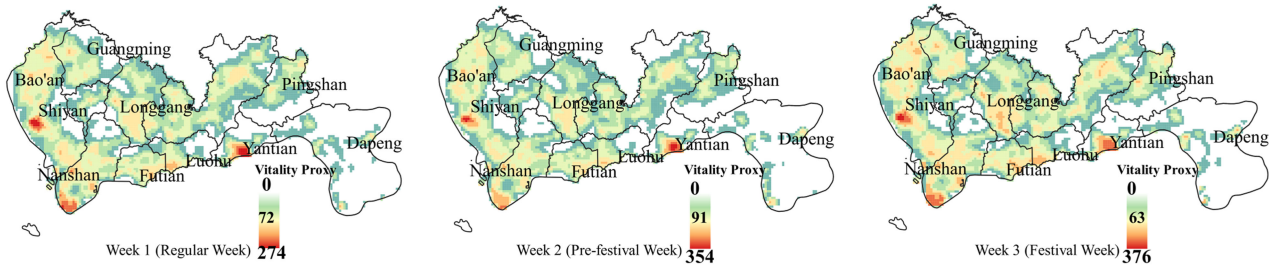
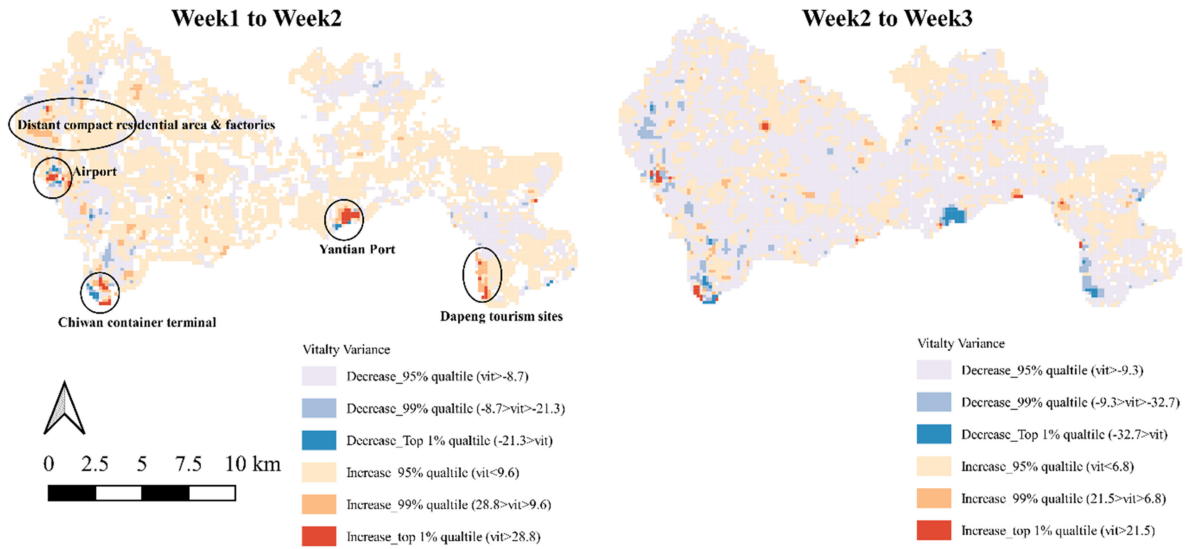Fig. 5.   Spatial distribution of urban vitality proxy.



Fig. 6.   Spatial distribution of changes between weeks.

Week2 − Week1 and Week3 − Week2 [see Fig. 6]. From the normal week (i.e., the first week) to the prefestival week (i.e., the second week), the increase in the vitality proxy is found in a wide range of locations, indicating socioeconomic activities becoming more vibrant across the city during the prefestival week due to various types of events, such as family reunions and tourism [62], [63]. During the holiday week, many locations with high vitality return to lower levels, though some locations in less central areas (e.g., North of Shenzhen) increase in vitality in week 3. Distant locations are known to have clusters of factories in Shenzhen. Many migrant workers return to the city in week 2, sustaining and even increasing the intensity of vitality in these distant locations. These results show that NTL-based values efficiently capture the overall structure of vitality as well as the detailed urban dynamics over time at a local scale.

Vitality indices are the variables constructed to explain urban vitality using various aspects of diversity. Within each grid cell, we calculate four ridership diversity indices (i.e., VPub, CPub, VTaxi, and CTaxi), three built environment diversity indices (i.e., LUM, RNR, and DOC), and a spatial interaction diversity index (i.e., DSI). All indices are normalized to a 0–1 scale for comparison and regression model fitting. The statistical distribution of vitality indices and the proxy is examined using letter-value plots [see Fig. 7], which overcomes the shortcomings of traditional boxplots when handling big data and presents information outside the normal quantiles in a clearer way [54].
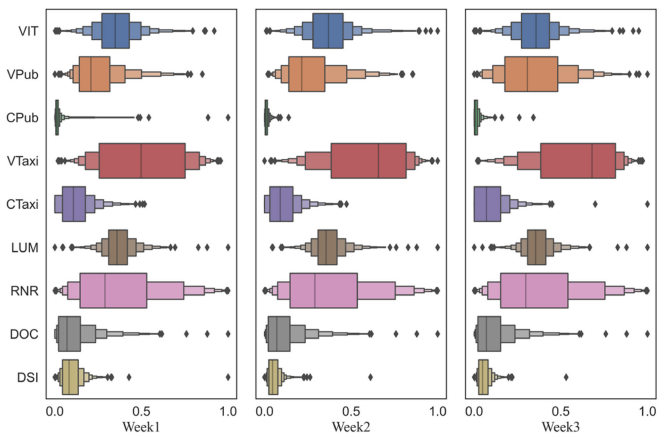


Fig. 7.   Statistical distribution of vitality proxy and indices.

Vitality proxy (VIT) presents a long right tail and some outliers, meaning that most locations have a relatively low or medium vitality proxy while few locations have high vitality. This result agrees with the common understanding of the power law of human mobility [55]. The distribution of ridership diversity indices reveals the basic characteristics of public transport and taxi mobility. In particular, CPub is very strongly right-skewed, indicating that within-day public transport trips are not diverse (i.e., are quite stable). VTaxi is the only index that is

<div style="text-align:center">

TABLE I
SUMMARY OF MODEL VARIABLES

</div>

| Role | Variable Category | Variable Name |
|---|---|---|
| Response Variable | Vitality Proxy | VIT |
| Explanatory Variables | Ridership Diversity | VPub |
| | | CPub |
| | | VTaxi |
| | | CTaxi |
| | Built Environment Diversity | LUM |
| | | RNR |
| | | DOC |
| | Interaction Diversity | DSI |

VIT = urban vitality proxy derived from nighttime light data, VPub = variability metric of ridership amount of public transport across days, CPub = consistency metric of ridership amount of public transport across hours, VTaxi = variability metric of ridership amount of taxi across days, CTaxi = consistency metric of ridership amount of taxi across hours, LUM = mixture of land use metric, RNR = residential and nonresidential ratio, DOC = density of catering business, DSI = diversity of spatial interaction.

left-skewed, indicating that day-to-day trip variation patterns of taxi trips are common in many locations. Travel behavior in different transport modes could greatly differ [56]. Overall, the vitality proxy and indices calculated in this study capture various distributions of city characteristics.

### B. MGWR Model Performance

MGWR models are inferred for each week to examine the local relationship between the vitality proxy and vitality indices. To obtain reliable evidence for which indices better explain the vitality proxy in varying spatial and temporal contexts, model performance is a key aspect to be assessed.

Table II shows the diagnosis of the overall goodness of fit and the significance level of the bandwidth inference. The overall performance of all weekly models is satisfactory, given by the r-squared values of week 1 ($R^2 = 0.931$), week 2 ($R^2 = 0.908$), and week 3 ($R^2 = 0.918$). Bandwidths are the spatial ranges inferred by (M)GWR to "borrow" data points to fit local regressors. In this study, the bandwidth inference for all variables is statistically significant, given by the Adj alpha lower than 0.05. It should be mentioned that the bandwidth inferences for DSI, the new index proposed in this study, have a significance level even lower than 0.01.

The differences in the bandwidths of the variables represent an advantage of MGWR over traditional GWR. The flexible bandwidths indicate that the explanatory variables explain the response variable at multiple scales. In this study, noticeable differences are found in the bandwidths of the vitality indices. In week 1, for example, VPub, CPub, CTaxi, RNR, and DOC show relatively small-scale bandwidths (i.e., less than 100 m), while the bandwidths of VTaxi and DSI are at a medium scale (i.e., less than 300 m), and that of LUM is at a large scale (i.e., more than 1000 m). Smaller scales mean that the variables can confidently explain the vitality proxy at a finer resolution. The results in Table II show that no vitality indices explain vitality in a fixed scale. The bandwidths of most indices vary between

<div style="text-align:center">

TABLE II
BANDWIDTH AND OVERALL DIAGNOSIS

</div>

| | Bandwidth Diagnosis | | | | Overall Diagnosis | | | |
|---|---|---|---|---|---|---|---|---|
| | Bandwidth | ENP_j | Adj t-val(95%) | Adj alpha | R2 | AICc | Residual sum of squares | Model |
| Constant | 10 | 307.187 | 3.782 | 0.000*** | | | | |
| VPub | 96 | 26.886 | 3.118 | 0.002*** | | | | |
| CPub | 48 | 41.233 | 3.243 | 0.001*** | | | | |
| VTaxi | 294 | 4.037 | 2.505 | 0.012* | | | | |
| CTaxi | 82 | 30.119 | 3.152 | 0.002*** | 0.931 | 1939.408 | 88.280 | Week1 |
| LUM | 1272 | 1.255 | 2.057 | 0.04* | | | | |
| RNR | 50 | 48.176 | 3.288 | 0.001*** | | | | |
| DOC | 59 | 37.435 | 3.215 | 0.001*** | | | | |
| DSI | 110 | 16.54 | 2.971 | 0.003*** | | | | |
| Constant | 10 | 314.835 | 3.788 | 0.000*** | | | | |
| VPub | 300 | 8.3 | 2.751 | 0.006*** | | | | |
| CPub | 338 | 7.542 | 2.719 | 0.007*** | | | | |
| VTaxi | 189 | 6.615 | 2.676 | 0.008*** | | | | |
| CTaxi | 110 | 20.886 | 3.043 | 0.002*** | 0.908 | 1905.237 | 114.473 | Week2 |
| LUM | 1054 | 1.778 | 2.198 | 0.028* | | | | |
| RNR | 179 | 12.791 | 2.891 | 0.004*** | | | | |
| DOC | 51 | 42.96 | 3.255 | 0.001*** | | | | |
| DSI | 97 | 19.233 | 3.018 | 0.003*** | | | | |
| Constant | 10 | 285.072 | 3.764 | 0.000*** | | | | |
| VPub | 210 | 11.226 | 2.85 | 0.004*** | | | | |
| CPub | 666 | 4.083 | 2.509 | 0.012* | | | | |
| VTaxi | 290 | 3.446 | 2.448 | 0.015* | | | | |
| CTaxi | 34 | 63.267 | 3.365 | 0.000*** | 0.918 | 1768.030 | 94.906 | Week3 |
| LUM | 1158 | 1.29 | 2.069 | 0.039* | | | | |
| RNR | 52 | 43.183 | 3.257 | 0.001*** | | | | |
| DOC | 902 | 1.903 | 2.225 | 0.026* | | | | |
| DSI | 94 | 17.788 | 2.994 | 0.003*** | | | | |

*** represents significance level of 1%.
* represents significance level of 5%.

<div style="text-align:center">

TABLE III
EVALUATION OF LOCAL MULTICOLLINEARITY OF MGWR MODELS

</div>

| | metrics | Week 1 | | | Week 2 | | | Week 3 | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | mean | std | max | mean | std | max | mean | std | max |
| Local variation decomposition proportions (VDP) | VPub | 0.109 | 0.150 | 0.809 | 0.053 | 0.083 | 0.568 | 0.063 | 0.091 | 0.612 |
| | CPub | 0.228 | 0.261 | 0.975 | 0.033 | 0.050 | 0.392 | 0.014 | 0.021 | 0.155 |
| | VTaxi | 0.215 | 0.245 | 0.848 | 0.494 | 0.299 | 0.945 | 0.321 | 0.264 | 0.867 |
| | CTaxi | 0.208 | 0.253 | 0.943 | 0.244 | 0.258 | 0.954 | 0.408 | 0.320 | 0.991 |
| | DSI | 0.450 | 0.342 | 0.977 | 0.574 | 0.292 | 0.986 | 0.553 | 0.333 | 0.981 |
| | DOC | 0.262 | 0.249 | 0.959 | 0.223 | 0.265 | 0.939 | 0.023 | 0.046 | 0.465 |
| | LUM | 0.006 | 0.008 | 0.075 | 0.013 | 0.019 | 0.165 | 0.009 | 0.018 | 0.184 |
| | RNR | 0.259 | 0.267 | 0.963 | 0.079 | 0.115 | 0.748 | 0.186 | 0.240 | 0.944 |
| Local condition number (CN) | | 3.702 | 0.984 | 7.983 | 3.684 | 1.271 | 9.945 | 3.643 | 1.407 | 11.802 |

the regular and holiday weeks. Two indices should be given special attention: 1) Mixture of land use (LUM) is considered an important aspect related to vitality. However, we found that it is only effective at a large scale in all weeks of this study, which is similarly mentioned by Yue *et al.* [25]. 2) Spatial interaction diversity, the new index implemented in this study, presents a small bandwidth in all weeks.

Besides this diagnosis, we also examine the collinearity issue of MGWR models in this study, although the collinearity of (M)GWR is not more problematic than traditional regression models [32]. In this article, we use the local condition number (CN) and local variation decomposition proportions (VDP) as collinearity indicators for MGWR [32]. VDP is inferred for each variable, while CN is obtained at the model level. Evaluations of all three weeks' models are shown in Table III. CN in all weeks is below the rule of thumb of 30, and the VDP of most variables is below the rule of thumb of 0.5 in terms of the mean value. These results suggest that the collinearity issue is acceptable and not problematic in this study's MGWR models.

TABLE IV
LOCAL PERFORMANCE COMPARISON AMONG OLS, GWR, AND MGWR

|  | Week1 | | | Week2 | | | Week3 | | |
|---|---|---|---|---|---|---|---|---|---|
|  | ols | gwr | mgwr | ols | gwr | mgwr | ols | gwr | mgwr |
| R2 | 0.216 | 0.729 | 0.931 | 0.216 | 0.697 | 0.908 | 0.176 | 0.680 | 0.918 |
| Adj-R2 | 0.211 | 0.639 | 0.884 | 0.211 | 0.603 | 0.859 | 0.170 | 0.590 | 0.870 |
| aic | 3323 | 2591 | 1242 | 3248 | 2640 | 1434 | 3085 | 2478 | 1253 |
| aicc |  | 2807 | 1939 |  | 2827 | 1905 |  | 2621 | 1768 |

TABLE V
STATISTICS OF MGWR COEFFICIENTS

|  | Min | Median | Max | Mean | SD | Model |
|---|---|---|---|---|---|---|
| Constant | -1.974 | 0.130 | 2.815 | 0.168 | 0.787 | |
| VPub | -0.165 | -0.014 | 0.629 | -0.005 | 0.108 | |
| CPub | -0.812 | 0.022 | 1.237 | 0.094 | 0.308 | |
| VTaxi | -0.339 | -0.106 | 0.012 | -0.129 | 0.118 | |
| CTaxi | -0.200 | 0.013 | 0.206 | 0.013 | 0.075 | Week1 |
| LUM | -0.039 | -0.037 | -0.036 | -0.038 | 0.001 | |
| RNR | -0.462 | -0.026 | 0.365 | -0.032 | 0.108 | |
| DOC | -0.161 | 0.087 | 0.451 | 0.093 | 0.100 | |
| DSI | -0.096 | 0.098 | 0.473 | 0.117 | 0.119 | |
| Constant | -1.951 | 0.222 | 3.045 | 0.233 | 0.797 | |
| VPub | -0.071 | -0.024 | 0.118 | -0.020 | 0.030 | |
| CPub | -0.175 | -0.010 | 0.028 | -0.028 | 0.048 | |
| VTaxi | -0.496 | -0.086 | 0.190 | -0.102 | 0.212 | |
| CTaxi | -0.140 | 0.014 | 0.260 | 0.021 | 0.087 | Week2 |
| LUM | -0.048 | -0.017 | -0.003 | -0.022 | 0.014 | |
| RNR | -0.134 | -0.023 | 0.159 | -0.026 | 0.064 | |
| DOC | -0.116 | 0.088 | 0.681 | 0.111 | 0.115 | |
| DSI | -0.163 | 0.099 | 0.397 | 0.115 | 0.105 | |
| Constant | -2.256 | 0.218 | 3.045 | 0.255 | 0.851 | |
| VPub | -0.049 | -0.003 | 0.273 | 0.020 | 0.059 | |
| CPub | -0.098 | -0.048 | -0.005 | -0.045 | 0.023 | |
| VTaxi | -0.309 | 0.020 | 0.122 | -0.033 | 0.135 | |
| CTaxi | -0.654 | 0.015 | 1.430 | 0.046 | 0.270 | Week3 |
| LUM | -0.035 | -0.033 | -0.026 | -0.032 | 0.002 | |
| RNR | -0.315 | -0.041 | 0.275 | -0.037 | 0.094 | |
| DOC | 0.065 | 0.096 | 0.109 | 0.090 | 0.015 | |
| DSI | -0.113 | 0.199 | 0.572 | 0.215 | 0.155 | |

Based on the same feature sets, the ordinary least squares (OLS) regression and classic GWR are also inferred for comparison (see Table IV). MGWR shows the best goodness of fit compared to the global regression (OLS) and GWR, given by the largest *r*-squared and lowest AIC. This information also supports that the MGWR-based vitality model explains the local relationship between the vitality indices and the vitality proxy in a better way than the fixed scale model (GWR) and global relational model (OLS).

### C. Vitality Evaluation Based on Multiscale Analysis of Varying Coefficients

Based on the good model performance illustrated in the previous section, the MGWR models can explain the vitality–diversity relationship. Therefore, we use model coefficients as the main source to examine urban vitality and its underlying dominant factors. We analyze the coefficients at both city and district levels for statistical features and spatiotemporal patterns.

*1) Relationship Between the Vitality Proxy and Urban Diversity:* The insights delivered by the coefficients are twofold. On the one hand, the positive and negative values indicate whether increasing values of the vitality indices predict an increasing or decreasing vitality proxy. On the other hand, the magnitude of the coefficient indicates to what extent the vitality indices can explain vitality proxy.

At the city level, we calculate the statistical features of the model coefficients in Table V using a process to aggregate pixel coefficients at a global (city) level. Although the combinations slightly vary between different weeks, we found that the vitality proxy is negatively related to ridership diversity of public transport (e.g., VPub and CPub) and built environment diversity (e.g., LUM and RNR). However, the vitality proxy is positively related to within-day taxi ridership diversity (e.g., CTaxi), DOC, and DSI.

The relationship between the vitality proxy and ridership diversity is not the same as the relationship presented by Sulis *et al.* [48] but reasonably reflects the situation in Shenzhen. In the previous study, variability (*V*) of trips (i.e., day-to-day trip volume variation) is positively related to vitality, and consistency (*C*) is negatively related to vitality. In our study, higher variability (*V*) of both public transport and taxi trips are negatively related to vitality, while the consistency of taxi trips is positively related to vitality. The possible reason for the contrast is due to the differences in the overall city context. Our empirical results suggest that, in Shenzhen, less day-to-day trip variability reflects better urban vitality; in other words, the long-term stability of these travel modes is more important. Meanwhile, the good

within-day ridership diversity of taxis (i.e., sudden peaks of trip volumes in a day) indicates a better vitality proxy. This result is reasonable because that suddenly increased demand for taxis reflects diverse types of travel purposes at a location, which relates to the figure of vibrant socioeconomic activities. We also found that CTaxi coefficients have a higher magnitude in week 2, suggesting that taxi trips are preferred over public transport during the Spring Festival week. There is similar evidence for VPub weighing more in weeks 1 and 3, indicating that public transport plays a more important role during nonfestival weeks.

The LUM and RNR coefficients do not follow the common view that diversity of land use relates to a better vitality proxy. In our results, the coefficients of built environment diversity have negative values in most places. This does not mean that built environment diversity harms nighttime vitality; instead, we argue that other indices such as ridership diversity are more positively related to nighttime vitality, indicating the importance of dynamic vitality indices. Because all selected indices explain the vitality proxy in a compound manner, the strong positive contribution of some variables may lead to negative coefficients of other variables. In our case, indices derived from transit data are found to be more positively related to vitality proxy. Similar evidence is also found in the smaller bandwidths of transit-based indices, as presented in the previous section. Based on our results, we argue that human mobility indices more effectively explain the vitality proxy. However, a nonhuman mobility index, i.e., DOC, is a strong positive predictor of vitality in all weeks. This result agrees with the research that proposed DOC as a sole indicator of vitality [24], [50]. Furthermore, DSI is positively related to the vitality proxy, which is surprisingly consistent throughout these weeks with varying temporal contexts.
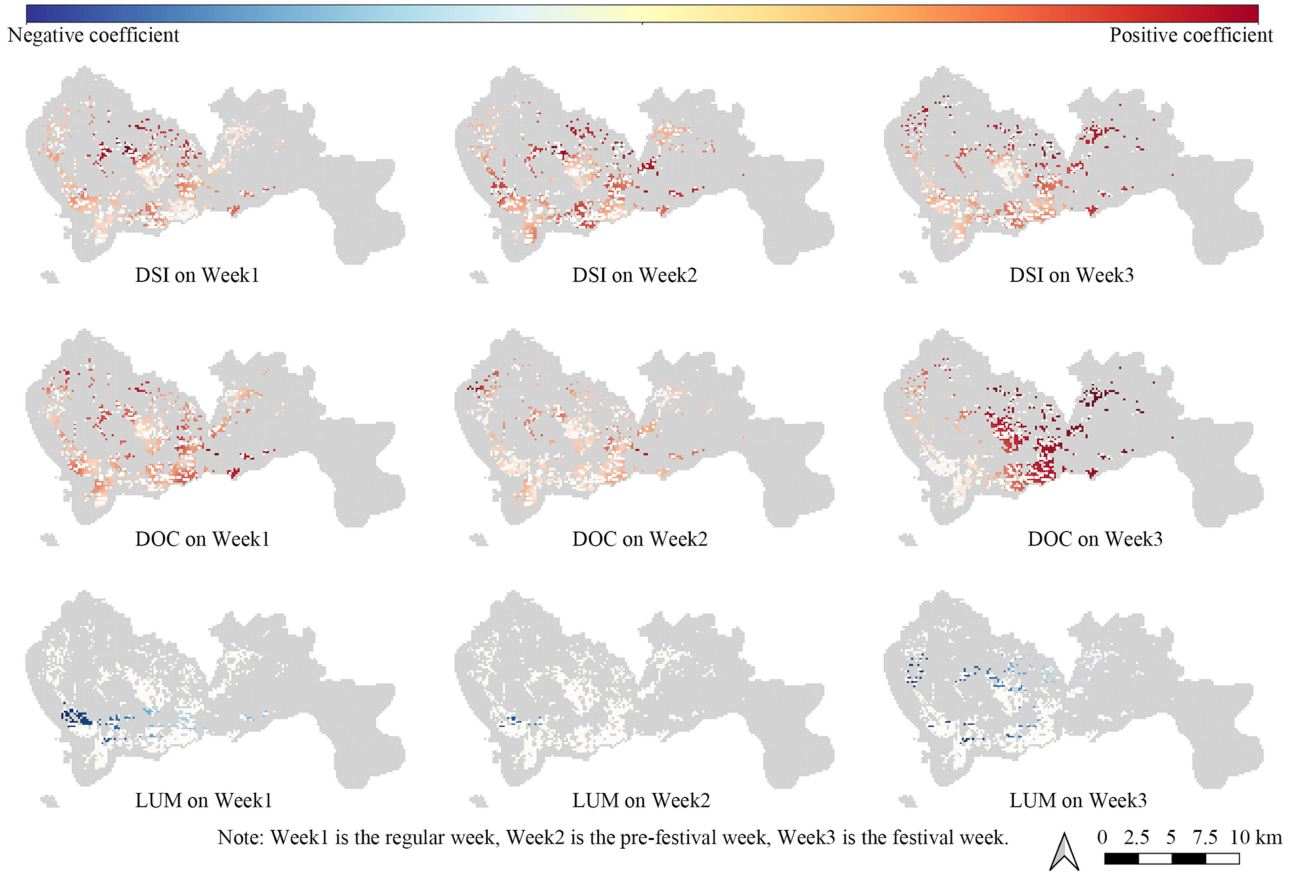
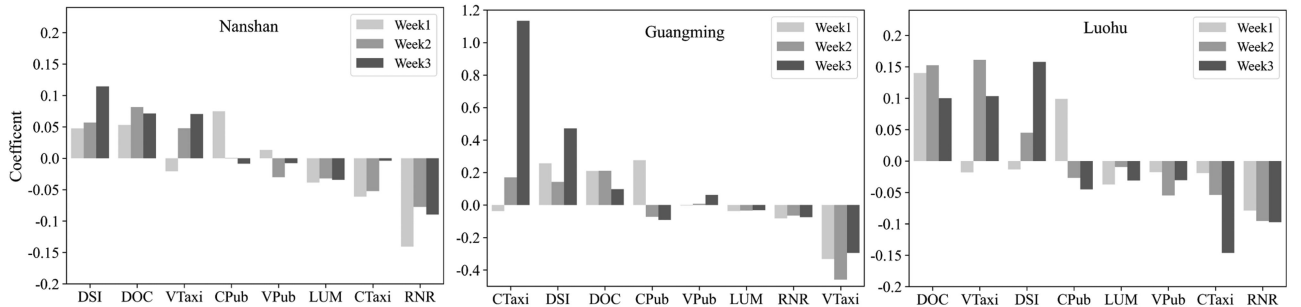Fig. 8.    Spatial distribution of DSI, DOC, and LUM coefficients.



Fig. 9.    Ranked barplot of coefficients in different districts and weeks.

*2) Spatial Variation of the Vitality–Diversity Relationship:* To augment the statistical features of the coefficients, the MGWR coefficients are mapped to improve the understanding of city-level patterns (see Fig. 8). The horizontal axis represents week 1 to week 3, and the vertical axis represents different indices. DSI, DOC, and LUM are selected for demonstration. Positive values are represented by red colors, negative values by blue colors, and nonsignificant estimated locations are identified using a *t*-value filter and are masked in white [32]. Although we have discussed the DSI, DOC, and LUM in the previous paragraphs, what cannot be fully interpreted from Table V is how these indices explain the vitality proxy differently across different locations. In Fig. 8, we find that DSI is positively linked to vitality in a wide range of locations, whereas DOC

only explains vitality in some clustered areas. Overall, the maps clearly show that the relationships between the vitality indices and vitality proxy are spatially variable with different patterns. Thus, previous studies that constructed global relationships to quantify urban vitality may have failed to consider the local, leading to under/overestimation in some places [48], [57].

*3) Temporal Variation of the Vitality–Diversity Relationship:* Temporal changes of coefficients can be observed using standard deviation (see Table V and Fig. 9). This metric is calculated across locations, which means that higher standard deviations indicate more substantial coefficient variation in space. We ranked the vitality indices according to the magnitude of the std value in Fig. 9. VTaxi, with the highest std(coef), and LUM are the most stable indices to explain vitality across the city. In
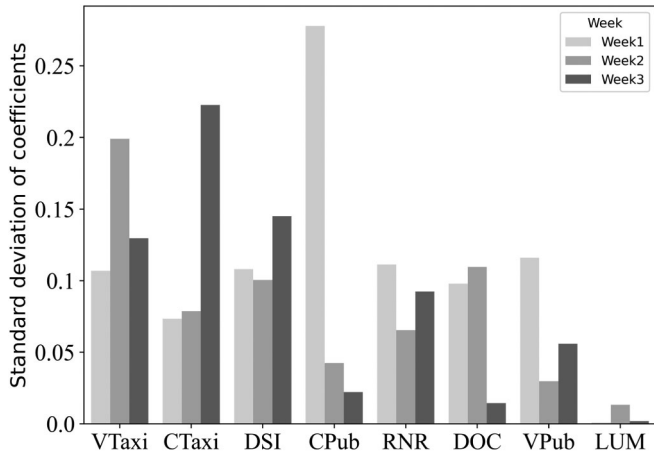
Fig. 10. Ranked barplot of coefficients standard deviation in different weeks.

Fig. 9, different shades of grey are used to represent the std(coef) in different weeks. Public transport and taxi trip indices present coefficient variations in both the space and time dimensions. DSI and DOC coefficients change across space but are relatively stable throughout time. The results based on std reveal characteristics of different vitality indices at a higher level, providing useful references for future research that may prioritize the uncertainty (variation) of vitality indices for representing the vitality proxy.

*4) Profile of the Unique District Context of Urban Vitality:* We conduct further analysis by plotting coefficients at the district level. Because administrative boundaries within the city are barriers that impact many important processes such as economic policy and resource allocation, analysis at the district level may lead to different vitality results. Heterogeneous development of districts is reported in Shenzhen [58]. The combination of dominant vitality indices and their changes throughout time can serve as an alternative angle to profile the vitality in the district context. In Fig. 8, we use three districts as a demonstration. Nanshan is a district known for the clustering of technology companies, such as Tencent, where DSI is found to most strongly explain vitality. The travelers to and from Nanshan may be associated with multiple types of business purposes. In contrast, Guangming is a more distant district known for its clustering of factories. The diversity of within-day taxi trips is the most positive variable related to vitality, followed by DSI. Luohu is a similarly central district as Nanshan, though its most dominant vitality index is DOC. Luohu has a different urban context due to its location next to Hong Kong, which is more urbanized by entertainment and residential entities. Overall, this analysis is an effective follow-up tool to utilize the MGWR-based vitality model coefficients to profile the district context.

## V. CONCLUSION

In this article, we proposed a framework to integrate NTL, public transport and taxi data, and POIs to evaluate the relationships between nighttime vitality and urban diversity (i.e., ridership diversity, built environment diversity, and spatial interaction diversity). By comparing the results in regular and holiday weeks, the major conclusions are as follows.

The MGWR-based vitality method outperforms OLS- and GWR-based vitality in several ways. MGWR model performance is satisfactory in all weeks in terms of goodness of fit and the significance of variables. Flexible scales inferred in MGWR provide additional information beyond varying coefficients.

NTL is an effective vitality proxy. During the study period, spatiotemporal patterns of NTL changed obviously, showing its feasibility for observing intracity dynamics. The significant relationship between the vitality proxy and diversity indices proves that NTL captures the vitality that suits the core concept in vitality theory very well. Diversity is important for urban vitality.

The inferred scales and coefficients of the MGWR vitality models provide several practical insights:

1) Although the relationships between nighttime vitality and urban diversity are strong in all weeks, the relationships varied with different patterns. Across space, ridership diversity and DSI explain the vitality proxy on a smaller spatial scale than built environment diversity. Across time, DOC and DSI are the most stable indices associated with the vitality proxy.

2) Regarding the overall relationship, diversity is not always positively related to vitality. For example, less diverse (i.e., stable) public transport trips relate to better vitality, while diverse (i.e., fluctuating) taxi trips relate to vitality in Shenzhen.

3) Spatial interaction diversity (DSI), a new index proposed in this study, is found to consistently play an important role in depicting the vitality proxy as a strong positive variable.

4) Aggregated coefficient values are useful for profiling the district-level socioeconomic context. For example, spatial interaction is found to be the most important vitality index in Nanshan, which happens to be a district known for its clustering of technology companies, such as Tencent. DSI refers to travelers from diverse locations. Taxi ridership diversity is most important in Guangming, which is a more distantly located district and is known for having clustered factories.

It should be noted that these practical insights are discussed specifically in the context of Shenzhen. In the future work, comparative studies can be conducted in multiple cities using our framework. Urban vitality is a comprehensive concept involving many aspects of the city. Although we follow a common rule (i.e., strength of economic activity) to choose NTL as a vitality proxy, the limitation of NTL data should be noted. NTL represents socioeconomic activity more instead of real-time and all kinds of activity. However, examining the integration of NTL in an urban vitality framework is still useful, considering that NTL is more accessible and provides wider coverage compared to private datasets such as mobile phone and social media data.

## REFERENCES

[1] Y. R. March, S. Wilkinson, and E. Firidin Özgür, "Measuring building adaptability and street vitality," *Plan. Pract. Res.*, vol. 27, no. 5, pp. 531–552, Oct. 2012, doi: 10.1080/02697459.2012.715813.

[2] H.-G. Sung, D.-H. Go, and C. G. Choi, "Evidence of Jacobs's street life in the Great Seoul City: Identifying the association of physical environment with walking activity on streets," *Cities*, vol. 35, pp. 164–173, Dec. 2013, doi: 10.1016/j.cities.2013.07.010.

[3] H. Sung and S. Lee, "Residential built environment and walking activity: Empirical evidence of Jane Jacobs' urban vitality," *Transp. Res. Part D: Transp. Environ.*, vol. 41, pp. 318–329, Dec. 2015, doi: 10.1016/j.trd.2015.09.009.

[4] J. Jacobs, *The Life and Death of Great American Cities*. New York, NY, USA: Random House, 1961, pp. 222–240.

[5] J. Gehl, *Life Between Buildings: Using Public Space*. Washington, DC, USA: Island Press, 2011, pp. 161–162.

[6] J. Montgomery, "Making a city: Urbanity, vitality and urban design," *J. Urban Des.*, vol. 3, no. 1, pp. 93–116, Feb. 1998, doi: 10.1080/13574809808724418.

[7] S. Williams, W. Xu, S. B. Tan, M. J. Foster, and C. Chen, "Ghost cities of China: Identifying urban vacancy through social media data," *Cities*, vol. 94, pp. 275–285, Nov. 2019, doi: 10.1016/j.cities.2019.05.006.

[8] A. Zhang, W. Li, J. Wu, J. Lin, J. Chu, and C. Xia, "How can the urban landscape affect urban vitality at the street block level? A case study of 15 metropolises in China," *Environ. Plan. B, Urban Anal. City Sci.*, vol. 48, no. 5, pp. 1245–1262, 2021, doi: 10.1177/2399808320924425.

[9] C. D. Elvidge, K. E. Baugh, E. A. Kihn, H. W. Kroehl, E. R. Davis, and C. W. Davis, "Relation between satellite observed visible-near infrared emissions, population, economic activity and electric power consumption," *Int. J. Remote Sens.*, vol. 18, no. 6, pp. 1373–1379, Apr. 1997, doi: 10.1080/014311697218485.

[10] X. Li, H. Xu, X. Chen, and C. Li, "Potential of NPP-VIIRS nighttime light imagery for modeling the regional economy of China," *Remote Sens.*, vol. 5, no. 6, pp. 3057–3081, Jun. 2013, doi: 10.3390/rs5063057.

[11] X. Li, N. Levin, J. Xie, and D. Li, "Monitoring hourly night-time light by an unmanned aerial vehicle and its implications to satellite remote sensing," *Remote Sens. Environ.*, vol. 247, Sep. 2020, Art. no. 111942, doi: 10.1016/j.rse.2020.111942.

[12] C. D. Elvidge, K. E. Baugh, S. J. Anderson, P. C. Sutton, and T. Ghosh, "The night light development index (NLDI): A spatially explicit measure of human development from satellite data," *Social Geogr.*, vol. 7, no. 1, pp. 23–35, Jul. 2012, doi: 10.5194/sg-7-23-2012.

[13] T. Ma, Y. Zhou, Y. Wang, C. Zhou, S. Haynie, and T. Xu, "Diverse relationships between Suomi-NPP VIIRS night-time light and multi-scale socioeconomic activity," *Remote Sens. Lett.*, vol. 5, no. 7, pp. 652–661, Jul. 2014, doi: 10.1080/2150704x.2014.953263.

[14] M. Zhao, W. Cheng, C. Zhou, M. Li, N. Wang, and Q. Liu, "GDP spatialization and economic differences in South China based on NPP-VIIRS nighttime light imagery," *Remote Sens.*, vol. 9, no. 7, Jul. 2017, Art. no. 673, doi: 10.3390/rs9070673.

[15] R. J. R. Elliott, E. Strobl, and P. Sun, "The local impact of typhoons on economic activity in China: A view from outer space," *J. Urban Econ.*, vol. 88, pp. 50–66, Jul. 2015, doi: 10.1016/j.jue.2015.05.001.

[16] K. Shi, Y. Chen, L. Li, and C. Huang, "Spatiotemporal variations of urban CO2 emissions in China: A multiscale perspective," *Appl. Energy*, vol. 211, pp. 218–229, Feb. 2018, doi: 10.1016/j.apenergy.2017.11.042.

[17] C. C. M. Kyba *et al.*, "Artificially lit surface of Earth at night increasing in radiance and extent," *Sci. Adv.*, vol. 3, no. 11, Nov. 2017, Art. no. e1701528, doi: 10.1126/sciadv.1701528.

[18] M. Zhao *et al.*, "Applications of satellite remote sensing of nighttime light observations: Advances, challenges, and perspectives," *Remote Sens.*, vol. 11, no. 17, Aug. 2019, Art. no. 1971, doi: 10.3390/rs11171971.

[19] T. Ma, "An estimate of the pixel-level connection between visible infrared imaging radiometer suite day/night band (VIIRS DNB) nighttime lights and land features across China," *Remote Sens.*, vol. 10, no. 5, May 2018, Art. no. 723, doi: 10.3390/rs10050723.

[20] T. Ma, "Multi-level relationships between satellite-derived nighttime lighting signals and social media–derived human population dynamics," *Remote Sens.*, vol. 10, no. 7, Jul. 2018, Art. no. 1128, doi: 10.3390/rs10071128.

[21] T. Ma, "Quantitative responses of satellite-derived night-time light signals to urban depopulation during Chinese new year," *Remote Sens. Lett.*, vol. 10, no. 2, pp. 139–148, Nov. 2018, doi: 10.1080/2150704x.2018.1530484.

[22] G. Xu, T. Xiu, X. Li, X. Liang, and L. Jiao, "Lockdown induced night-time light dynamics during the COVID-19 epidemic in global megacities," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 102, Oct. 2021, Art. no. 102421, doi: 10.1016/j.jag.2021.102421.

[23] C. Zeng, Y. Song, Q. He, and F. Shen, "Spatially explicit assessment on urban vitality: Case studies in Chicago and Wuhan," *Sustain. Cities Soc.*, vol. 40, pp. 296–306, Jul. 2018, doi: 10.1016/j.scs.2018.04.021.

[24] Y. Long and C. Huang, "Does block size matter? The impact of urban design on economic vitality for Chinese cities," *Environ. Plan. B, Urban Anal. City Sci.*, vol. 46, no. 3, pp. 406–422, Jun. 2017, doi: 10.1177/2399808317715640.

[25] Y. Yue, Y. Zhuang, A. G. O. Yeh, J.-Y. Xie, C.-L. Ma, and Q.-Q. Li, "Measurements of POI-based mixed use and their relationships with neighbourhood vibrancy," *Int. J. Geogr. Inf. Sci.*, vol. 31, no. 4, pp. 658–675, Aug. 2016, doi: 10.1080/13658816.2016.1220561.

[26] W. Wu and X. Niu, "Influence of built environment on urban vitality: Case study of Shanghai using mobile phone location data," *J. Urban Plann. Dev.*, vol. 145, no. 3, Sep. 2019, Art. no. 04019007, doi: 10.1061/(asce)up.1943-5444.0000513.

[27] C. Jia, Y. Du, S. Wang, T. Bai, and T. Fei, "Measuring the vibrancy of urban neighborhoods using mobile phone data with an improved pagerank algorithm," *Trans. GIS*, vol. 23, no. 2, pp. 241–258, Mar. 2019, doi: 10.1111/tgis.12515.

[28] W. Yue, Y. Chen, Q. Zhang, and Y. Liu, "Spatial explicit assessment of urban vitality using multi-source data: A case of shanghai, China," *Sustainability*, vol. 11, no. 3, Jan. 2019, Art. no. 638, doi: 10.3390/su11030638.

[29] C. Xia, A. G.-O. Yeh, and A. Zhang, "Analyzing spatial relationships between urban land use intensity and urban vitality at street block level: A case study of five Chinese megacities," *Landscape Urban Plan.*, vol. 193, Jan. 2020, Art. no. 103669, doi: 10.1016/j.landurbplan.2019.103669.

[30] Y.-L. Kim, "Seoul's wi-fi hotspots: Wi-Fi access points as an indicator of urban vitality," *Comput., Environ. Urban Syst.*, vol. 72, pp. 13–24, Nov. 2018, doi: 10.1016/j.compenvurbsys.2018.06.004.

[31] C. Wu, X. Ye, F. Ren, and Q. Du, "Check-in behaviour and spatio-temporal vibrancy: An exploratory analysis in Shenzhen, China," *Cities*, vol. 77, pp. 104–116, Jul. 2018, doi: 10.1016/j.cities.2018.01.017.

[32] A. S. Fotheringham, W. Yang, and W. Kang, "Multiscale geographically weighted regression (MGWR)," *Ann. Amer. Assoc. Geogr.*, vol. 107, no. 6, pp. 1247–1265, Aug. 2017, doi: 10.1080/24694452.2017.1352480.

[33] C. D. Elvidge, K. E. Baugh, M. Zhizhin, and F.-C. Hsu, "Why VIIRS data are superior to DMSP for mapping nighttime lights," *Proc. Asia-Pac. Adv. Netw.*, vol. 35, pp. 62–69, Jun. 2013, doi: 10.7125/apan.35.7.

[34] C. Small and C. D. Elvidge, "Night on earth: Mapping decadal changes of anthropogenic night light in Asia," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 22, pp. 40–52, Jun. 2013, doi: 10.1016/j.jag.2012.02.009.

[35] C. Elvidge, T. Ghosh, F.-C. Hsu, M. Zhizhin, and M. Bazilian, "The dimming of lights in China during the COVID-19 pandemic," *Remote Sens.*, vol. 12, no. 17, Sep. 2020, Art. no. 2851, doi: 10.3390/rs12172851.

[36] M. O. Román and E. C. Stokes, "Holidays in lights: Tracking cultural patterns in demand for energy services," *Earth's Future*, vol. 3, no. 6, pp. 182–205, Jun. 2015, doi: 10.1002/2014ef000285.

[37] X. Zhao *et al.*, "NPP-VIIRS DNB daily data in natural disaster assessment: Evidence from selected case studies," *Remote Sens.*, vol. 10, no. 10, Sep. 2018, Art. no. 1526, doi: 10.3390/rs10101526.

[38] A. Crooks *et al.*, "Crowdsourcing urban form and function," *Int. J. Geogr. Inf. Sci.*, vol. 29, no. 5, pp. 720–741, Jan. 2015, doi: 10.1080/13658816.2014.977905.

[39] T. Hu, J. Yang, X. Li, and P. Gong, "Mapping urban land use by using landsat images and open social data," *Remote Sens.*, vol. 8, no. 2, Feb. 2016, Art. no. 151, doi: 10.3390/rs8020151.

[40] R. Yang, H. Yan, W. Xiong, and T. Liu, "The study of pedestrian accessibility to rail transit stations based on KLP model," *Procedia - Social Behav. Sci.*, vol. 96, pp. 714–722, Nov. 2013, doi: 10.1016/j.sbspro.2013.08.082.

[41] D. Murakami *et al.*, "The importance of scale in spatially varying coefficient modeling," *Ann. Amer. Assoc. Geogr.*, vol. 109, no. 1, pp. 50–70, Dec. 2018, doi: 10.1080/24694452.2018.1462691.

[42] A. S. Fotheringham, H. Yue, and Z. Li, "Examining the influences of air quality in China's cities using multi-scale geographically weighted regression," *Trans. GIS*, vol. 23, no. 6, pp. 1444–1464, Sep. 2019, doi: 10.1111/tgis.12580.

[43] T. J. Hastie and R. J. Tibshirani, *Generalized Additive Models*. London, U.K.: Routledge, 2017.

[44] M. O. Román *et al.*, "NASA's Black Marble nighttime lights product suite," *Remote Sens. Environ.*, vol. 210, pp. 113–143, Jun. 2018, doi: 10.1016/j.rse.2018.03.017.

[45] Z. Wang, M. O. Román, V. L. Kalb, S. D. Miller, J. Zhang, and R. M. Shrestha, "Quantifying uncertainties in nighttime light retrievals from Suomi-NPP and NOAA-20 VIIRS day/night band data," *Remote Sens. Environ.*, vol. 263, Sep. 2021, Art. no. 112557, doi: 10.1016/j.rse.2021.112557.

[46] C. Kang, D. Fan, and H. Jiao, "Validating activity, time, and space diversity as essential components of urban vitality," *Environ. Plan. B, Urban Anal. City Sci.*, vol. 48, no. 5, pp. 1180–1197, May 2020, doi: 10.1177/2399808320919771.

[47] C. Zhong *et al.*, "Variability in regularity: Mining temporal mobility patterns in London, Singapore and Beijing using smart-card data," *PLoS One*, vol. 11, no. 2, Feb. 2016, Art. no. e0149222, doi: 10.1371/journal.pone.0149222.

[48] P. Sulis, E. Manley, C. Zhong, and M. Batty, "Using mobility data as proxy for measuring urban vitality," *J. Spatial Inf. Sci.*, no. 16, pp. 137–162, Jun. 2018, doi: 10.5311/josis.2018.16.384.

[49] Y. Liu, Z. Sui, C. Kang, and Y. Gao, "Uncovering patterns of inter-urban trip and spatial interaction from social media check-in data," *PLoS One*, vol. 9, no. 1, Jan. 2014, Art. no. e86026, doi: 10.1371/journal.pone.0086026.

[50] K. Stopar, D. Drobne, K. Eler, and T. Bartol, "Citation analysis and mapping of nanoscience and nanotechnology: Identifying the scope and interdisciplinarity of research," *Scientometrics*, vol. 106, no. 2, pp. 563–581, Nov. 2015, doi: 10.1007/s11192-015-1797-x.

[51] L. Leydesdorff, C. S. Wagner, and L. Bornmann, "Interdisciplinarity as diversity in citation patterns among journals: Rao-Stirling diversity, relative variety, and the Gini coefficient," *J. Informetrics*, vol. 13, no. 1, pp. 255–269, Feb. 2019, doi: 10.1016/j.joi.2018.12.006.

[52] L. D. Frank *et al.*, "The development of a walkability index: Application to the neighborhood quality of life study," *Brit. J. Sports Med.*, vol. 44, no. 13, pp. 924–933, Apr. 2009, doi: 10.1136/bjsm.2009.058701.

[53] Y. Ye, D. Li, and X. Liu, "How block density and typology affect urban vitality: An exploratory analysis in Shenzhen, China," *Urban Geogr.*, vol. 39, no. 4, pp. 631–652, Sep. 2017, doi: 10.1080/02723638.2017.1381536.

[54] H. Hofmann, H. Wickham, and K. Kafadar, "Letter-Value plots: Boxplots for large data," *J. Comput. Graphical Statist.*, vol. 26, no. 3, pp. 469–477, Jul. 2017, doi: 10.1080/10618600.2017.1305277.

[55] M. C. González, C. A. Hidalgo, and A.-L. Barabási, "Understanding individual human mobility patterns," *Nature*, vol. 458, no. 7235, pp. 238–238, Mar. 2009, doi: 10.1038/nature07850.

[56] X. Zhang, Y. Xu, W. Tu, and C. Ratti, "Do different datasets tell the same story about urban mobility — A comparative study of public transit and taxi usage," *J. Transport Geogr.*, vol. 70, pp. 78–90, Jun. 2018, doi: 10.1016/j.jtrangeo.2018.05.002.

[57] X. Delclòs-Alió and C. Miralles-Guasch, "Looking at Barcelona through Jane Jacobs's eyes: Mapping the basic conditions for urban vitality in a Mediterranean conurbation," *Land Use Policy*, vol. 75, pp. 505–517, Jun. 2018, doi: 10.1016/j.landusepol.2018.04.026.

[58] B. Güneralp and K. C. Seto, "Environmental impacts of urban growth from an integrated dynamic perspective: A case study of Shenzhen, South China," *Global Environ. Change*, vol. 18, no. 4, pp. 720–735, Oct. 2008, doi: 10.1016/j.gloenvcha.2008.07.004.

[59] C. Mellander, J. Lobo, K. Stolarick, and Z. Matheson, "Night-time light data: A good proxy measure for economic activity?," *PLoS One*, vol. 10, no. 10, Oct. 2015, Art. no. e0139779, doi: 10.1371/journal.pone.0139779.

[60] C. Côté-Lussier, A. Knudby, and T. A. Barnett, "A novel low-cost method for assessing intra-urban variation in night time light and applications to public health," *Social Sci. Med.*, vol. 248, Mar. 2020, Art. no. 112820, doi: 10.1016/j.socscimed.2020.112820.

[61] D. Ma, R. Guo, Y. Jing, Y. Zheng, Z. Zhao, and J. Yang, "Intra-urban scaling properties examined by automatically extracted city hotspots from street data and nighttime light imagery," *Remote Sens.*, vol. 13, no. 7, Mar. 2021, Art. no. 1322, doi: 10.3390/rs13071322.

[62] J. Huang, X. Liu, P. Zhao, J. Zhang, and M.-P. Kwan, "Interactions between bus, metro, and taxi use before and after the Chinese Spring festival," *ISPRS Int. J. Geo-Inf.*, vol. 8, no. 10, Oct. 2019, Art. no. 445, doi: 10.3390/ijgi8100445.

[63] J. Liu, W. Shi, and P. Chen, "Exploring travel patterns during the holiday season—A case study of Shenzhen Metro system during the Chinese Spring festival," *ISPRS Int. J. Geo-Inf.*, vol. 9, no. 11, Oct. 2020, Art. no. 651, doi: 10.3390/ijgi9110651.

**Junwei Zhang** received the B.Eng. degree in resource prospecting engineering from the China University of Geosciences, Beijing, China, in 2015, and the M.Sc. degree in geomatics from The Hong Kong Polytechnic University, Hong Kong, in 2017, where he is currently working toward the Ph.D. degree in geomatics with the Department of Land Surveying and Geo-Informatics.

His research interests include spatiotemporal data mining, complex network analysis, and GIScience.

**Xintao Liu** received the B.Eng. degree in survey from Hohai University, Nanjing, China, in 1998, the M.Sc. degree in cartography and GIS from Nanjing Normal University, Nanjing, China, in 2003, and the Ph.D. degree in geoinformatics from the Royal Institute of Technology, Stockholm, Sweden, in 2012.

In 2012, he joined the Department of Civil Engineering, Ryerson University, Toronto, ON, Canada, where he was a Postdoctoral Fellow in GIS and transportation until 2016. Since 2015, he has been a Sessional Lecturer with Ryerson University. He is currently an Assistant Professor with the Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University. He is also a PI and a Co-PI of several national projects funded by Sweden, Canada, and Hong Kong. His research interests include GI services and science, urban computing, and GIS in transportation. His research goal is to use the state-of-the-art technologies to advance smart city for a better urban life.

Dr. Liu is currently a Reviewer of a series of major international journals, such as International Journal of Geographical Information Science and Annals of the American Association of Geographers in his field. He was the recipient of the Ph.D. Scholarship from Lars Erik Lundbergs.

**Xiaoyue Tan** received the B.S. degree in geographical information system from China University of Mining and Technology, Xuzhou, China, in 2016, and the M.S. degree in geography from Beijing Normal University, Beijing, China, in 2019. She is currently working toward the Ph.D. degree in geomatics with the Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University.

Her research interests include nighttime light remote sensing and spatiotemporal data mining.

**Tao Jia** received the Ph.D. degree in geoinformatics from the KTH Royal Institute of Technology, Stockholm, Sweden, in 2012.

He is currently an Associate Professor with the School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China. His research interests include geographical information science in general, spatiotemporal data mining, complex network analysis and modeling, traffic analysis, and human dynamics in particular.

**Ahmad M. Senousi** received the B.Eng. and M.Sc. degrees in civil engineering from Cairo University, Cairo, Egypt, in 2013 and 2017, respectively, and the Ph.D. degree in geographical information science from the Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong, in 2021.

He is currently an Assistant Lecturer with the Civil Engineering Department, Cairo University. His research interests include geographic information science in general, spatiotemporal data mining, complex network analysis, multilayer network, urban structure, and transport geography.

**Ling Yin** received the B.S. and M.S. degrees in geographical information system (GIS) from Nanjing University, Nanjing, China, in 2003 and 2006, respectively, and the Ph.D. degree in geographical information science from the University of Tennessee at Knoxville, Knoxville, TN, USA, in 2011.

Since 2011, she has been with the Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Beijing, China, where she became a Full Professor in 2020. Her research interests include spatiotemporal data mining, spatially epidemic modeling, urban computing, and GIS for smart city.

**Jianwei Huang** received the B.Sc. degree in geographical information science from Yunnan University, Kunming, China, in 2014, the M.Sc. degree in geoinformatics from the Salzburg University, Salzburg, Austria, and the M.Sc. degree in cartography and geographical information system from Nanjing Normal University, Nanjing, China, in 2017. He is currently working toward the Ph.D. degree in earth system and geoinformation science with the Institute of Space and Earth Information Science, The Chinese University of Hong Kong, Hong Kong.

His research interests include human mobility, environmental health, and GIScience.

**Fan Zhang** received the Ph.D. degree in communication and information system from the Huazhong University of Science and Technology, Wuhan, China, in 2007.

He was a Postdoctoral Fellow with the University of New Mexico and the University of Nebraska-Lincoln from 2009 to 2011. He is currently a Professor with the Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China. He is also the Director of the Shenzhen Institute of Beidou Applied Technology. His research topics include intelligent transportation systems, urban computing, and big data and AI technology.