

# Chapter 4

## Population Characterization in Location Modeling: Alternatives, Impacts, and Insights



Daoqin Tong, Wangshu Mu, and Changfeng Li

**Abstract** Location analysis and modeling have been widely applied to support locational decisions for service provision. The general idea of such analysis has been for sited facilities to serve the demand of interest in an efficient and/or effective way. In many applications, service demand involves either general people or certain population groups in a region. Currently, population-based demand has been assessed mainly based on where people live, primarily using census population count data. This can be problematic given that people do not always stay at home or originate their trips from home. As a result, relying upon residential information may lead to an inaccurate evaluation of service demand in location modeling. This study investigates the impacts of alternative population characterizations on the classic  $p$ -median problem. A new model incorporating time-varying population distributions is introduced. An empirical study was conducted in three regions in Shanghai, China, where time-varying population distributions were derived using cell phone data. Analysis results show that solutions generated based on where people live can be far from the optimal that considers the temporal variability of population distributions. Discussion is provided on ways to remedy the issue.

**Keywords** Location modeling · Population · Demand

---

D. Tong (✉) · W. Mu  
School of Geographical Sciences and Urban Planning, Arizona State University, Tempe, AZ,  
USA  
e-mail: [Daoqin.Tong@asu.edu](mailto:Daoqin.Tong@asu.edu)

C. Li  
China Academy of Urban Planning and Design, Beijing, China

## 4.1 Introduction

Location analysis and modeling have been widely used to support locational decisions for various service provisions in both the public and private sectors. Applications include urban green space design (Zhang et al. 2017), emergency humanitarian logistics (Boonmee et al. 2017), health service development planning (Rahman and Smith 2000; Ahmadi-Javid et al. 2017), and farmers' market placement (Tong et al. 2012). The general idea of these models has been to locate facilities to serve the demand of interest in an efficient and/or effective way. While some models focus on ensuring a certain level of service with minimal resources, others try to achieve the maximal efficiency or equity given a limited budget.

In many location models, people often serve as the demand for the intended services. Such services include healthcare (Murawski and Church 2009; Meskarian et al. 2017), public transportation (Wu and Murray 2005), cell phone signal coverage (Akella et al. 2010), emergency responses (Marianov 2017), and disaster relief good planning (Widener and Horner 2011; Chen et al. 2013). In these studies, population has been mainly characterized based on where people live. Such information can be obtained from Census Bureau in certain aggregate form. For example, American Community Survey (ACS) provides population count data based on census data collection units. Using census data, Socioeconomic Data and Applications Center (SEDAC) provides population estimates at certain grid level (e.g., 1 km).

Characterizing demand based on where people live might be problematic in many real-world applications as people may need to be present at workplaces, move on roadways, and visit parks depending on the time of the day. Except for a few studies, relying on the static residential information for population representation may lead to an inaccurate evaluation of service demand in location modeling. It remains unknown to what extent such an inaccurate demand assessment affects the optimal solution and whether the impact varies with population distributions. Drawing on spatiotemporal cell phone data collected in Shanghai, China, this chapter aims to provide an investigation on these questions. The next section provides a literature review on population characterization in existing location analysis studies and the associated problems. This is followed by an introduction to a classic location model and a new model incorporating temporal variability of population. An empirical study is then conducted with the results presented. We conclude with some discussion and future research directions.

## 4.2 Background

In many location modeling studies, population-based demand is often approximated using census population count data given the data availability. The count data summarize the total population information in the associated data collection unit, such as census block group or tract. One common practice has been to aggregate

these areal units into points each assigned with the corresponding population information. For example, in a study on the sexual health service provision, Meskarian et al. (2017) aggregated postcode-level population into a set of points representing the service demand. In seeking the best sites for locating mobile food markets, Widener et al. (2012) used block group centroids to represent the demand site for fresh food. While most studies directly rely on census data for population demand, some applications may need finer population information. To better characterize the spatial variation, Wu and Murray (2005) made a population interpolation analysis at the  $30 \times 30$  m scale and applied the associated population estimates to determine the best modification strategy for transit service provision.

Using census population data as the proxy for service demand is equivalent to assuming people receive or originate their trips for the intended service at home. For certain services that people need to receive all the time such as cell phone signals and emergency responses, ensuring coverage of residential areas is insufficient. This is because in addition to home people frequently visit and stay at other important sites such as workplaces, schools, parks, etc. For example, the 2017 American Time Use Survey reported that workers on average spent about 8 h on an average weekday at work, and 83% of workers did some or all of their work at workplace. Also, people spend a significant amount of time traveling to their activity destinations. The 2017 National Household Survey found that on average, American drivers and passengers spent about an hour in a vehicle every day.

For services where people need to make their trips for, such as transit services, grocery stores, and medical care, current modeling results may only cover home-based trips by ensuring that the service provided is most convenient to homes. However, studies showed that people may also initiate their travel from other important locations such as school and workplace. Based on a survey, Mack and Tong (2015) reported that about 42% of the farmers' market trips originated from nonhome places. In general, nonhome-based trips have been found to account for over 30% of all daily trips (McGuckin et al. 2005). Recognizing that workplaces may serve as important sites where people originate their trips from, several studies expanded the demand representation to also include employment in their location modeling. For example, in a transit stop removal study, Wu and Murray (2005) considered the amount of employees in each census block along with the census population for potential transit service. Similarly, in determining the best farmers' markets sites, Tong et al. (2012) considered workers at their workplaces for potential demand for farmers' markets. While in some applications, it is important to locate services/facilities close to where people live or work, a more general framework has been to capture the commute flows by locating facilities close to commute routes. The approach to incorporating commute-based trip chain has been used to site children day care centers (Hodgson 1981) and determine the location and operation time of farmers' markets (Tong et al. 2012).

In addition to home and workplace, people may be at different places during different times of the day for various purposes, and these may include a large number of non-commute chained trips. Based on a 2008 National Household Travel Survey add-on dataset, Li and Tong (2017) observed 78% non-commute chained trips. They

then developed a facility location model to address a full spectrum of trip chaining by incorporating travelers' daily activity-travel program into service provision planning. According to their approach, people have the flexibility of visiting a facility from any activity site, including home, workplace, and other activity destinations, and along any trip made based on an individual's daily activity-travel program. Such an approach assumes the knowledge of an individual's daily activity-travel program, which can be very challenging due to the data availability issue.

To have a better understanding of the issue associated with population characterization in location modeling, this study analyzes the temporal variability of population distributions and examines how locational decisions may vary with alternative demand characterization. Different from detailed activity-travel data used in Li and Tong (2017), data involved in this study are relatively easier to obtain, especially considering the increasing availability of large geotagged data collected through cell phones and wearable devices. We also note the difference between people's daily movement and seasonal or long-term migration. For example, Ndiaye and Alfares (2008) provided a study on healthcare facility location where population groups migrated seasonally. In their study, during a season, population groups were fixed at home locations, and people's daily movement was not considered. We use the classic  $p$ -median problem (ReVelle et al. 2008) to demonstrate the nuances of alternative population characterization and investigate whether and how problem solutions may be impacted.

## 4.3 Methodology

### 4.3.1 PMP

The  $p$ -median problem (PMP) is one of the classic location problems that aims to site a number of facilities so that the overall demand-weighted travel distance/time to the closest facility is minimized. The problem was first introduced by Hakimi (1964, 1965) in a network context, where the optimal sites are called "medians" of the network. The PMP linear programming model was first provided by ReVelle and Swain (1970). Since then, the PMP has been widely studied in the literature. The problem has been applied to support cluster analysis (Klastorin 1985), transportation logistics (Pamučar et al. 2016), political redistricting (Hess et al. 1965), bike-sharing station planning (Park and Sohn 2017), and healthcare center siting (Jia et al. 2014).

Consider the following notation:

- $i$ : index of demand
- $j$ : index of candidate facility site
- $h_i$ : demand associated with  $i$
- $d_{ij}$ : distance between  $i$  and  $j$
- $p$ : the number of facilities to be sited

$$x_j = \begin{cases} 1 & \text{if candidate site } j \text{ is selected} \\ 0 & \text{otherwise} \end{cases}$$

$$y_{ij} = \begin{cases} 1 & \text{if demand } i \text{ is allocated to facility at } j \\ 0 & \text{otherwise} \end{cases}$$

The PMP can be formulated as

$$\text{Minimize } \sum_i \sum_j h_i d_{ij} y_{ij} \quad (4.1)$$

subject to

$$\sum_j y_{ij} = 1 \quad \forall i \quad (4.2)$$

$$y_{ij} \leq x_j \quad \forall i, j \quad (4.3)$$

$$\sum_j x_j = p \quad (4.4)$$

$$x_j, y_{ij} \in \{0, 1\} \quad \forall i, j \quad (4.5)$$

The PMP objective (4.1) minimizes the total demand-weighted travel distance. Constraints (4.2) require that each demand be assigned to one facility. Constraints (4.3) ensure that demand  $i$  is assigned to facility at  $j$  only when site  $j$  is selected for siting. Constraint (4.4) specifies  $p$  number of facilities to be sited. Constraints (4.5) impose binary conditions on decision variables. As specified in the original PMP, demand at  $i$  and  $h_i$  does not vary with time.

### 4.3.2 The PMP-Time-Varying Demand (PMP-TD)

As discussed previously, population distributions may vary with time ( $t$ ), resulting in a time-varying demand distribution  $h_{it}$ . The corresponding problem will then become identifying the spatial configuration of  $p$  facilities so that they are most accessible, considering the temporal variability of demand. Consider the additional notation,

$$y_{ijt} = \begin{cases} 1 & \text{if demand at } i \text{ is allocated to facility at } j \text{ during time } t \\ 0 & \text{otherwise} \end{cases}$$

$d_{ijt}$ : distance between  $i$  and  $j$  during time  $t$

The PMP-TD can be formulated as,

$$\text{Minimize } \sum_i \sum_j \sum_t h_{it} d_{ijt} y_{ijt} \quad (4.6)$$

subject to

$$\sum_j y_{ijt} = 1 \quad \forall i, t \quad (4.7)$$

$$y_{ijt} \leq x_j \quad \forall i, j, t \quad (4.8)$$

$$\sum_j x_j = p \quad (4.9)$$

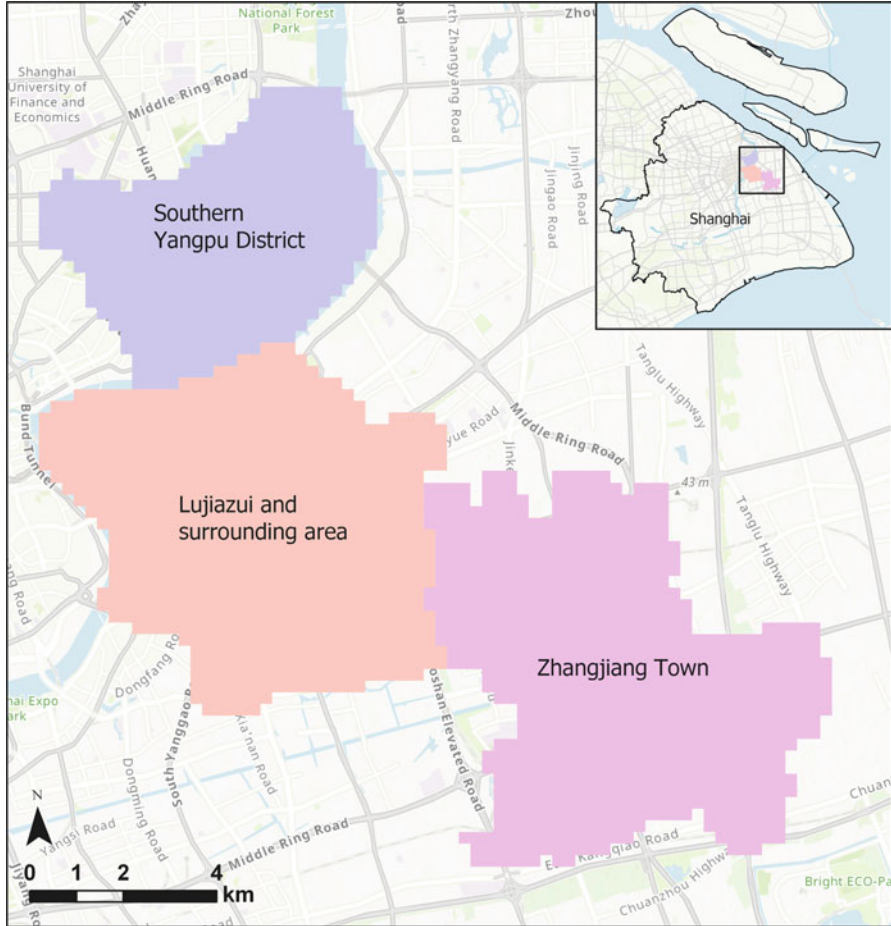
$$x_j, y_{ijt} \in \{0, 1\} \quad \forall i, j, t \quad (4.10)$$

Objective (4.6) minimizes the overall demand-weighted travel across all times. Constraints (4.7) specify that demand at  $i$  during time  $t$  can be assigned to only one facility. Constraints (4.8) states that demand at  $i$  during time  $t$  can be assigned to facility at  $j$  only when site  $j$  is selected for siting. Constraint (4.9) is the same as constraint (4.4). Constraints (4.10) impose binary integer conditions on the decision variables.

We note here that given a spatial configuration of facilities, the assignment of demand  $i$  does not change with time. This is because in the PMP demand ( $h_{it}$ ) at  $i$  is always assigned to the closest facility based on the minimal travel distance objective. That is, demand allocation only depends on the spatial distribution of  $i$  and  $j$ . Given that  $d_{ij}$  is fixed over time  $t$  ( $d_{ij} = d_{ijt}$ ),  $y_{ijt}$  is always the same for a given pair of  $i$  and  $j$ . The time-independent nature of demand allocation  $y_{ijt}$  also makes constraints (4.7) irrelevant to time: given  $i$  and  $j$ ,  $y_{ijt}=1$  for one time period  $t'$  ensures satisfaction of constraints (4.7) for all other time periods. As a result, in the PMP-TD,  $y_{ijt}$  collapses into  $y_{ij}$ , and constraints (4.7, 4.8, 4.9, and 4.10) can be replaced by constraints (4.2, 4.3, 4.4, and 4.5). Therefore, the new problem involves a new assessment of demand  $\sum_t h_{it}$  at demand site  $i$ . The new demand is a sum of the demand at  $i$  across all times. Objective (4.6) then becomes  $\sum_i \sum_j \sum_t h_{it} d_{ij} y_{ij}$  (11)

## 4.4 Empirical Study

We will use a case study to demonstrate how to incorporate time-varying population distributions into the PMP. We will also compare whether and how the solutions based on time-varying demand differ from those obtained based on where people live. The case study consisted of three regions in Shanghai, China (also see Fig. 4.1). The first region contains Lujiazui and its surrounding neighborhood communities with an overall area of 48 km<sup>2</sup>. This region serves as one of central business districts (CBDs) in Shanghai and is known as one of the most important financial districts in China. The second region is composed of the southern part of Yangpu district. This region is primarily residential area with a size of 30 km<sup>2</sup>. The third region is located in Zhangjiang Town with an area of 50.4 km<sup>2</sup>. This region contains a major



**Fig. 4.1.** The study area

technology park hosting many IT companies in the northwest and some residential areas in the southeast.

Spatiotemporal distributions of the population in the three regions were derived based on cell phone data provided by China Unicom. The data were collected for an entire week from Monday, November 20, 2017 to Sunday, November 26, 2017. Hourly population count was summarized using  $250 \times 250$  meter grids. In each of the three regions, we assumed five facilities to be sited ( $p = 5$ ). For each region, the PMP-TD was performed to obtain the optimal solution considering the temporal variation in the population distribution. Meanwhile, for each hour during weekdays and weekends, we obtained the PMP optimal solutions based on the population observed during that hour. For each of these solutions, we mapped the facility sites obtained and computed the overall travel using the time-varying demand. Such travel was then compared with the optimal travel obtained using the PMP-TD to compare the solution quality.

## 4.5 Results

### 4.5.1 Temporal Variability in the Population Distribution

Figure 4.2 shows the temporal variation of the population in the three regions during weekdays and weekends. The horizontal axis records the 24 h of a day; it starts from midnight (0) of the first day and ends before the midnight of the following day (23). On weekdays, compared to the midnight population, Lujiazui region gained a significant amount of population (58%) during the daytime, especially during the work hours (8 am–6 pm). This is not surprising given the CBD functionality of the region. In Zhangjiang Town, we also note significant population gain (56%) during weekday work hours. In contrast, Yangpu region had a stable population distribution with a daily population change of 8%. On weekends, while Yangpu and Zhangjiang had relatively consistent population throughout the day, Lujiazui attracted as much as 37% of people to this region as it also serves as an important tourist attraction site.

Figure 4.3 maps the spatial distribution of population gain/loss at 10 am compared with that at midnight for both weekdays and weekends. For each grid in a region, the population count at midnight is used as the baseline. Compared with the population at midnight, blue areas are places losing population at 10 am, whereas red areas correspond to locations gaining population. During weekday workday hours, Lujiazui has substantially more areas that gained population than areas that lost population (Fig. 4.3a). We notice that areas gaining population in Lujiazui were distributed extensively throughout the region. This is different from the pattern we observe in Zhangjiang (Fig. 4.3c). In Zhangjiang, areas gaining population are

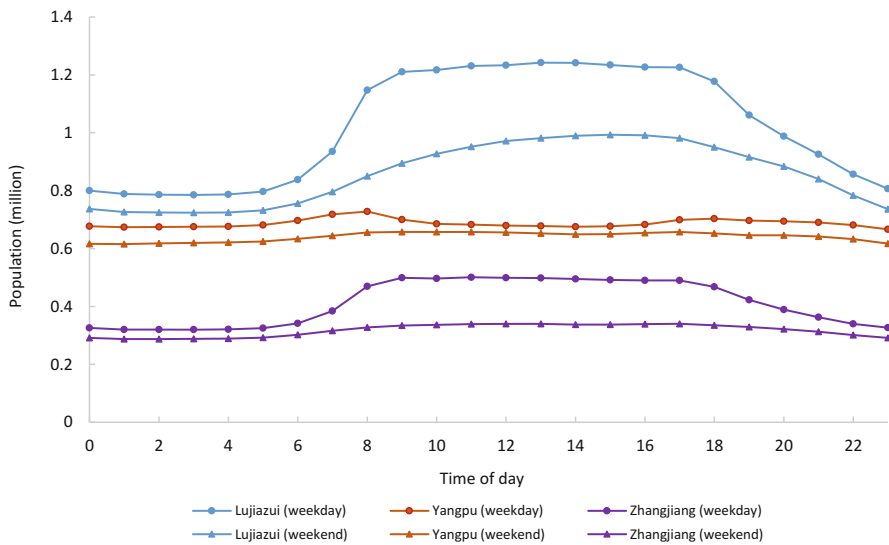
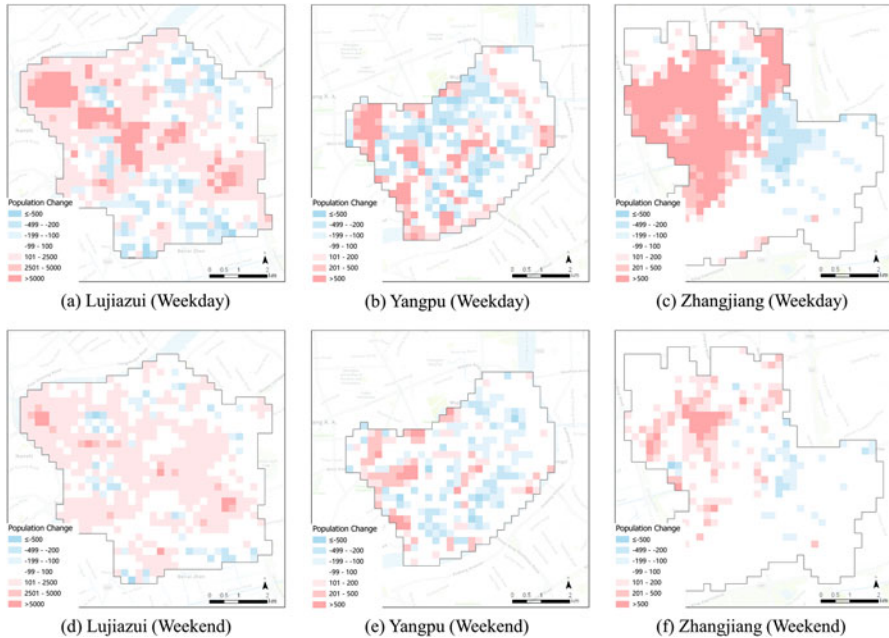


Fig. 4.2. Overall population change in the three regions

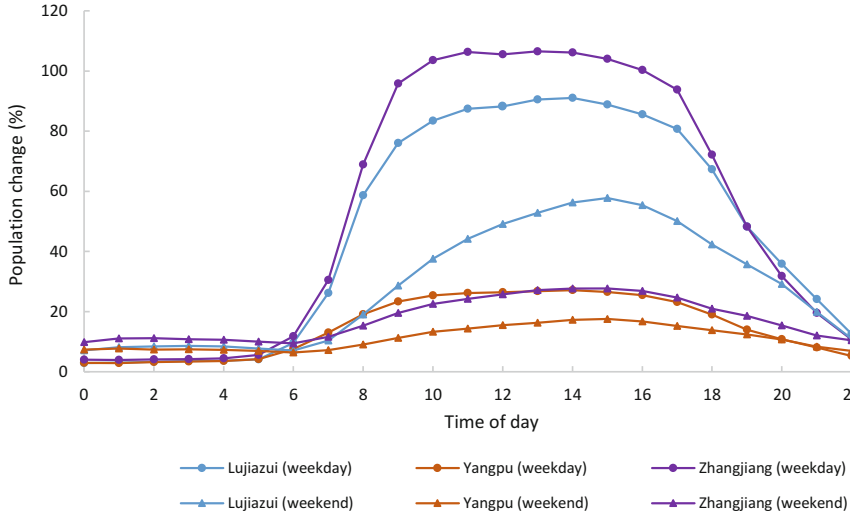




**Fig. 4.3.** Population gain/loss assessed at 10 am compared with midnight. (a) Lujiazui (weekday). (b) Yangpu (weekday). (c) Zhangjiang (weekday). (d) Lujiazui (weekend). (e) Yangpu (weekend). (f) Zhangjiang (weekend)

mainly clustered in the northwestern part of the town where the technology park is located, and areas losing population are concentrated in the residential areas east to the technology park. Different from Lujiazui and Zhangjiang, during weekday work hours, Yangpu has minimal areas gaining population, and these areas are highly dispersed in the region. During weekends, most of the areas gaining population in Lujiazui are similar to those during weekdays though with a smaller magnitude of gain. For Yangpu and Zhangjiang, much fewer areas gained population during weekends. For both weekdays and weekends, we notice that the magnitude of population gain in Lujiazui is much higher than the other two regions as reflected in the legend, indicating the ability of many zones in this region in attracting people during the daytime.

Figure 4.4 summarizes the average absolute population change (%) in a region for both weekdays and weekends. Similar to Fig. 4.3, we computed the change using the overall midnight population as the baseline. Population change in a grid at time  $t$  was computed using the percentage of population gain/loss at time  $t$  compared with the midnight population. Here, we did not differentiate population gain from loss as the focus is on the magnitude of change. Summarizing all grids in a region, the average population change was then calculated at time  $t$ . For all the three regions, population changes during weekdays were higher than those during weekends. Different from the overall population change shown in Fig. 4.2, Fig. 4.4 also captures people's

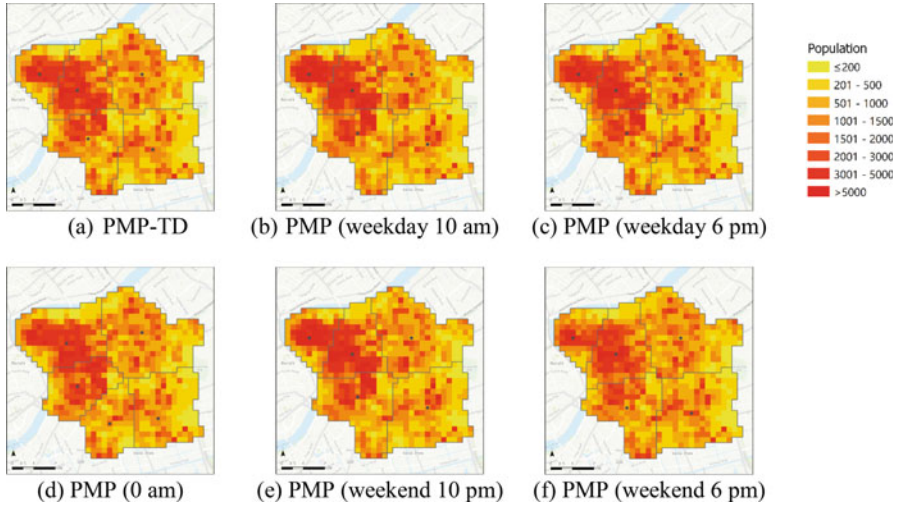


**Fig. 4.4** Average absolute population change in the three regions

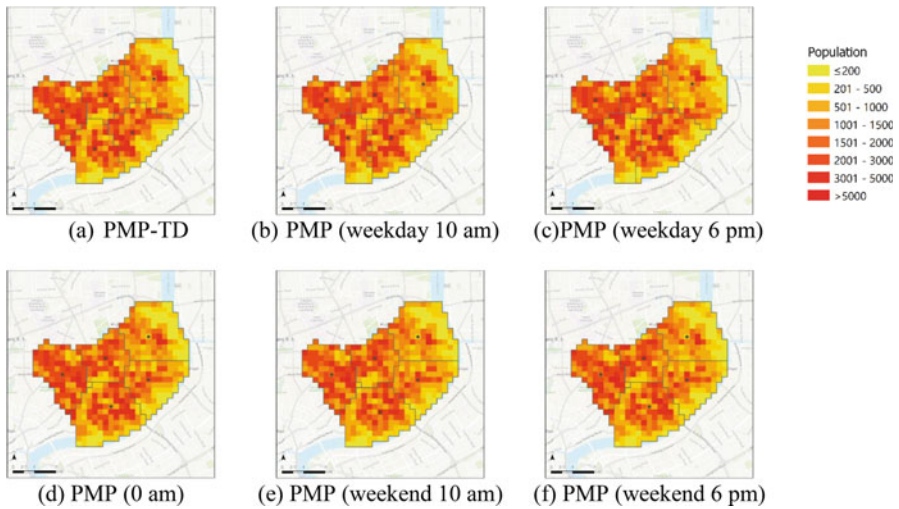
movement within a region. As shown in Fig. 4.4, in a general weekday, population changes were higher than the weekend population changes, which is also consistent with the overall population change in Fig. 4.2. However, when intraregional movement is considered, Fig. 4.4 gives different population change profiles. According to Fig. 4.4, on weekdays, Zhangjiang had the largest population change (106%), followed by Lujiazui (91%) and Yangpu (27%). This is different from the overall population change curve in Fig. 4.2, where Lujiazui had the highest overall population gain (58%), followed by Zhangjiang (56%) and Yangpu (8%). This suggests significant intraregional population exchange in Lujiazui and Zhangjiang. Unlike weekdays, Lujiazui had the highest average population change (58%) during weekends followed by Zhangjiang (28%) and Yangpu (18%).

## 4.5.2 Optimal Solution Comparison

We compared the PMP-TD solutions with the PMP solutions for three times ( $t$ ), 0 am, 10 am, and 6 pm (Figs. 4.5, 4.6 and 4.7). While we separated the weekday and weekend solutions for  $t = 10$  am and 6 pm, we used the average midnight population across the entire week to derive the solution for the PMP with  $t = 0$  am, given that the population distribution at midnight did not vary much across days. Figures 4.5, 4.6, and 4.7 plot the facilities selected (black stars) and the associated allocation (black lines) of population to its nearest facility. For the PMP solutions (Figs. 4.5b–f, 4.6b–f, and 4.7b–f), the population at the corresponding time  $t$  is shown as the



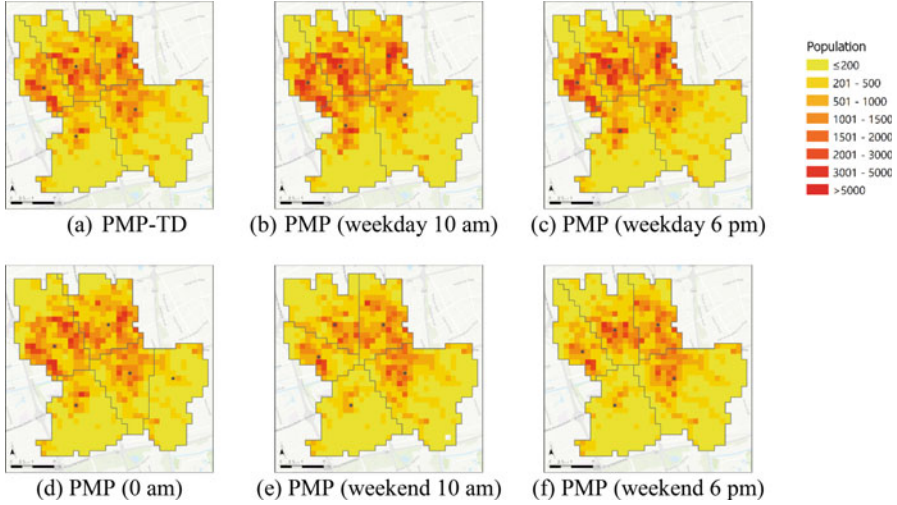
**Fig. 4.5.** Solution comparison in the Lujiazui region. (a) PMP-TD. (b) PMP (weekday 10 am). (c) PMP (weekday 6 pm). (d) PMP (0 am). (e) PMP (weekend 10 pm). (f) PMP (weekend 6 pm)



**Fig. 4.6.** Solution comparison in the Yangpu region. (a) PMP-TD. (b) PMP (weekday 10 am). (c) PMP (weekday 6 pm). (d) PMP (0 am). (e) PMP (weekend 10 am). (f) PMP (weekend 6 pm)

background. For the PMP-TD solutions, the overall average population incorporating the hourly variation across the entire week is mapped (Figs. 4.5a, 4.6a, and 4.7a).

For Lujiazui, the spatial configurations of sited facilities drawn based on the PMP during the daytime (e.g.,  $t = 10$  am and 8 pm) tend to resemble those given by the PMP-TD. This is as expected. As we show previously, the region gains a significant amount of population during the daytime (58%). The significantly higher demand in



**Fig. 4.7** Solution comparison in the Zhangjiang region. (a) PMP-TD. (b) PMP (weekday 10 am). (c) PMP (weekday 6 pm). (d) PMP (0 am). (e) PMP (weekend 10 am). (f) PMP (weekend 6 pm)

these hours lead to higher weights, given these hours in Objective (11), which eventually helps pull the PMP-TD optimal solution toward sites that best serve the population distribution during these hours. We also note that the population distribution during weekend daytime in this region is similar to that during weekday daytime with population concentrations in the northwestern part of the region. The PMP solutions during weekend daytime (Figs. 4.5e and f) are therefore similar to the PMP-TD solution. We notice significant difference in the spatial configuration of sited facilities when comparing the solution given by the PMP ( $t = 0$ ) with that given by the PMP-TD. While both models prescribe three facilities to serve the western part of the region, the PMP-TD sites two facilities in the northwest, whereas the PMP ( $t = 0$ ) locates only facility in that area.

In the Yangpu region, the solutions given by the PMP and PMP-TD are similar due to the overall small population change throughout the day. Only slight difference exists between the weekday and weekend solutions. We notice that weekend daytime PMP solutions are similar to the midnight PMP solutions. This is also to our anticipation, given that the weekend population distribution has the minimal temporal variation.

In Zhangjiang, we have a similar comparison observation to that in Lujiazui. The spatial configuration of the sited facilities using the PMP-TD (Fig. 4.7a) is very similar to that based on the PMP solutions during weekday daytime (Figs. 4.7b and c). This is because similar to Lujiazui, Zhangjiang gains substantial population (56%) during the weekday daytime. As we discussed previously, such an increase will result in higher weights in Objective (11) given to sites to better serve the weekday daytime population. The PMP solutions based on weekend and weekday daytime population are also similar except for the slight difference in the PMP

weekend morning solution (Fig. 4.7e). The spatial configuration of the PMP midnight solution is found to be drastically different from that given by the PMP-TD solution: while the PMP-TD prescribes three facilities to serve the technology park area, the PMP locates only two facilities in the area (Fig. 4.7d).

We use the population distribution at midnight ( $t = 0$ ) to approximate the census population. Assuming a spatial configuration of facilities sited using the PMP ( $t = 0$ ), we computed the travel involved using the time-varying demand and compared it with the travel based on the PMP-TD solutions. Figure 4.8 shows the comparison for both weekdays and weekends. Here, positive/negative additional travel means the PMP solutions need more/less travel when compared with the PMP-TD solutions. In general, the PMP solutions involve more travel during daytime (e.g., 7 am–8 pm) and less travel at night (e.g., 9 pm–6 am). As for daytime, the PMP solutions require significantly more travel during weekdays than weekends with the highest weekday additional travel of 15.3% for Lujiazui and 15.2% for Zhangjiang, respectively, compared to highest weekend additional travel of 7.5% for Lujiazui and 3.3% for Zhangjiang, respectively. Combining weekdays and weekends, the PMP solutions require an additional daytime travel of 9.2% for Lujiazui, 0.8% for Yangpu, and 8.6% for Zhangjiang. As for nighttime, the PMP solutions give a travel reduction of 5% for Lujiazui and 3% for Zhangjiang. Summarizing the entire week, the additional travel brought about by the PMP is 4.1% for Lujiazui, 0.3% for Yangpu, and 4.3% for Zhangjiang.

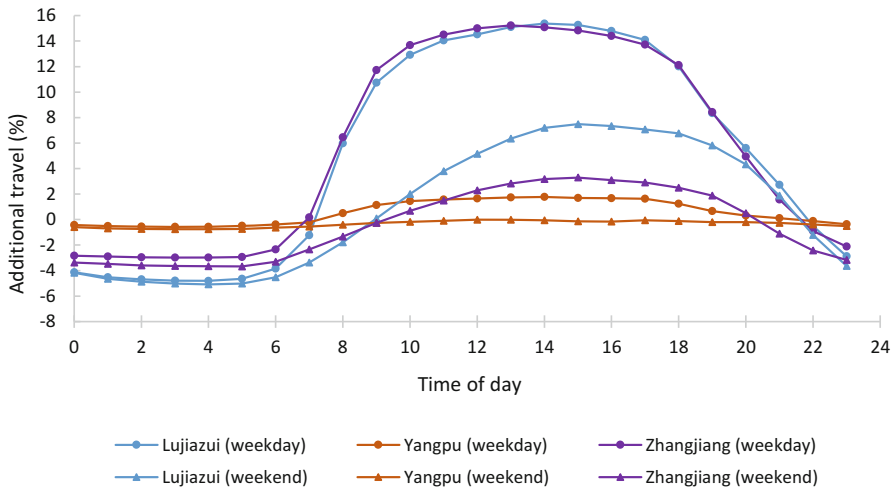


Fig. 4.8. Additional travel based on the PMP ( $t = 0$ ) solutions

## 4.6 Discussion and Conclusion

In this study, we observe distinctly different population distributions by time of day (daytime vs. nighttime) and across days (weekdays vs. weekends) especially in Lujiazui and Zhangjiang. While many of existing location models focus on siting facilities based on where people live, our empirical study indicates that the solutions obtained using this approach could be very different from the optimal one when the temporal variation of population is considered. Using a classic location model as an example, we find that the existing approach may result in much worse solutions when the temporal variation of population is large. We note that for many public services that are only available during daytime, such as postal offices and public libraries, the existing approach may give even worse solutions. Although the existing approach appears to be more applicable during weekends, our empirical study shows that the nontrivial temporal variation of population in two of the three regions has led to significantly more travel needed by the existing approach.

Numerous studies have been developed to seek the optimal solutions to the PMP, especially for medium- and large-sized problems. Many of these approaches try to close the optimal gap at the magnitude of less than 1%. For example, Mu and Tong (2018) introduced a spatial-knowledge-enhanced Teitz and Bart (STB) algorithm for solving the PMP with an improvement of less than 1% for most test cases. Irawan and Salhi (2015) developed a PMP solution heuristic based on a demand aggregation strategy and reported an improvement of less than 0.5%. If the temporal variability of population is not considered, solutions provided by the PMP could be farther from the optimal when compared with heuristic approaches. As we show previously, compared to the PMP-TD, the PMP solutions have an overall optimality gap of 4% for Lujiazui and Zhangjiang and 0.3% for Yangpu.

In real-world applications, it is therefore worthwhile to examine the temporal distribution of the targeted population before the implementation of a location model. For an area where the population distribution does not vary much with time, as in the case of Yangpu, a direct application of the associated location model based on census population information might yield a solution that is not very far from the optimal one. If an area involves significant population change throughout the day, solely relying upon census population data can be very problematic. In this case, a better characterization of the population will be needed. As for the PMP, we show that the average population throughout the day will be appropriate. In this case, the 24-h average population data (e.g., LandScan data) could be used. However, whether such data are suitable for other location models remains unknown.

In this research, we incorporate the temporal variability of population into location modeling by discretizing the population distribution into a finite number of time periods. The population in each time period has an equal probability of being served by the sited facility. This assumption is more appropriate for certain applications, such as cell phone signal coverage and emergency services. However, it can be less appropriate for some other applications, such as grocery stores and dining places



as people may be at work or have other constraints during a particular time period that may prevent them from using the service. In these applications, a better characterization of the population as potential demand will be needed.

Nowadays, big geospatial data have been widely collected through taxis, shared mobility applications such as Uber and Mobike, wearable devices, and social media platforms. The emergence of big data provides statistically sound samples with finer spatial and temporal resolutions. These data provides the opportunity to revisit some of the assumptions we make in many location models (Tong and Murray 2017). In this study, the temporal variation of population derived using cell phones allows us to examine the impact of the population assumption made in one classic location model. In addition, big data offer the opportunity for us to study individual-level mobility and travel activity, which will be helpful for a better characterization of service access. For example, the PMP assumes that people visit the closet facility, which might not always be true. Some people may chain a visit to a facility with other important trips even if the facility is far away. How to incorporate more complex travel behavior into location modeling points to another venue for future research.

## References

- Ahmadi-Javid A, Seyedi P, Syam SS (2017) A survey of healthcare facility location. *Comput Oper Res* 79:223–263
- Akella M, Delmelle E, Batta R, Rogerson P, Blatt A (2010) Optimal adaptive facility location using geostatistics. *Geogr Anal* 42:227–244
- Boonmee C, Arimura M, Asada T (2017) Facility location optimization model for emergency humanitarian logistics. *Int J Disaster Risk Reduct* 24:485–498
- Chen Z, Chen X, Li Q, Chen J (2013) The temporal hierarchy of shelters: a hierarchical location model for earthquake-shelter planning. *Int J Geogr Inf Sci* 27(8):1612–1630
- Hakimi SL (1964) Optimum locations of switching centers and the absolute centers and medians of a graph. *Oper Res* 12(3):450–459
- Hakimi SL (1965) Optimum distribution of switching centers in a communication network and some related graph theoretic problems. *Oper Res* 13(3):462–475
- Hess SW, Weaver JB, Siegfeldt HJ, Whelan JN, Zitlau PA (1965) Nonpartisan political redistricting by computer. *Oper Res* 13(6):998–1006
- Hodgson MJ (1981) The location of public facilities intermediate to the journey to work. *Eur J Oper Res* 6(2):199–204
- Irawan CA, Salhi S (2015) Solving large p-median problems by a multistage hybrid approach using demand points aggregation and variable neighborhood search. *J Glob Optim* 63:537–554
- Jia T, Tao H, Qin K, Wang Y, Liu C, Gao Q (2014) Selecting the optimal healthcare centers with a modified P-median model: a visual analytic perspective. *Int J Health Geogr* 13:42
- Klasterin TD (1985) The p-median problem for cluster analysis: a comparative test using the mixture model approach. *Manag Sci* 31(1):84
- Li R, Tong D (2017) Accessibility maximization for facility siting: an incorporation of activity space and trip chaining. *Socio Econ Plan Sci* 60:1–14
- Mack J, Tong D (2015) Characterizing the spatial and temporal patterns of farmers' market visits. *Appl Geogr* 63:43–54

- Marianov V (2017) Location models for emergency service applications. *INFORMS Tutorials Oper Res* 11:237–262
- Mcguckin N, Zmud J, Nakamoto Y (2005) Trip-chaining trends in the United States: understanding travel behavior for policy making. *Transp Res Rec J Transp Res Board* 1917:199–204
- Meskarian R, Penn ML, Williams S, Monks T (2017) A facility location model for analysis of current and future demand for sexual health services. *PLoS One* 12(8):e0183942. <https://doi.org/10.1371/journal.pone.0183942>
- Mu W, Tong D (2018) A spatial-knowledge-enhanced heuristic for solving the p-median problem. *Trans GIS* 22:477–493
- Murawski L, Church RL (2009) Improving accessibility to rural health services: the maximal covering network improvement problem. *Socio Econ Plan Sci* 43(2):102–110
- Ndiaye M, Alfares H (2008) Modeling health care facility location for moving population groups. *Comput Oper Res* 35:2154–2161
- Pamučar D, Vasin LJ, Atanasković P, Miličić M (2016) Planning the city logistics terminal location by applying the green p-median model and type-2 neuro-fuzzy network. *Comput Intell Neurosci* Article ID 6972818. <https://doi.org/10.1155/2016/6972818>
- Park C, Sohn SY (2017) An optimization approach for the placement of bicycle sharing stations to reduce short car trips: an application to the city of Seoul. *Transp Res A Policy Pract* 105:154–166
- Rahman S, Smith DK (2000) Use of location-allocation models in health service development planning in developing nations. *Eur J Oper Res* 123:437–452
- ReVelle CS, Swain RW (1970) Central facilities location. *Geogr Anal* 2(1):30–42
- ReVelle CS, Eiselt HA, Daskin MS (2008) A bibliography of some fundamental problem categories in discrete location science. *Eur J Oper Res* 184:817–848
- Tong D, Murray AT (2017) Location analysis: developments on the horizon. In: Jackson R, Schaeffer P (eds) *Regional research frontiers*, vol 2. Springer International Publishing, New York, pp 193–208
- Tong D, Ren F, Mack J (2012) Locating Farmers' Markets with an incorporation of spatio-temporal variation. *Socio Econ Plan Sci* 46(2):149–156
- Widener MJ, Horner MW (2011) A hierarchical approach to modeling hurricane disaster relief goods distribution. *J Transp Geogr* 19(4):821–828
- Widener MJ, Metcalf SS, Bar-Yam Y (2012) Developing a mobile produce distribution system for low-income urban residents in food deserts. *J Urban Health* 89(5):733–745
- Wu C, Murray AT (2005) Optimizing public transit quality and system access: the multiple-route, maximal covering/shortest-path problem. *Environ Plann B Plann Des* 32:163–178
- Zhang Y, Murray A, Turner BL (2017) Optimizing green space locations to reduce daytime and nighttime urban heat island effects in Phoenix, Arizona. *Landsc Urban Plan* 165:162–171