Estimating Bicycle Demand of a Small Community

Seungkyu Ryu[1*], Jacqueline Su[2], Anthony Chen[3], Keechoo Choi[4]

[1*] TOD-based Sustainable Urban Transportation, Ajou University, Suwon, 442-749, Korea; (corresponding author); Email: ryuseungkyu@ajou.ac.kr

[2] Department of Urban Planning University of California, Los Angeles, Los Angeles, 90095-1656; E-mail: jacsu@ucla.edu

[3] Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, USA; Email: anthony.chen@polyu.edu.hk

[4] Department of Transportation Engineering, Ajou University, Suwon, 442-749, Korea; Email: keechoo@ajou.ac.kr

Abstract: Currently, there is a growing movement in the urban planning and transportation sectors advocating for the creation of sustainable and livable communities. Since these communities focus on the promotion of public health and the protection of environmental resources, it comes to no surprise that cycling is experiencing increasing popularity as an alternative mode of transportation. With anticipated increases in cycling mode share, there is a need to account for cycling in future transportation networks by estimating bicycle demand. Thus, the objective of this paper is to present a procedure for estimating bicycle trips in smaller communities with limited resources. A case study at the Utah State University campus in Logan, Utah is conducted to demonstrate the applicability of the bicycle demand estimation procedure. The case study involves data collection, initial bicycle origin-destination (O-D) estimation using a gravity model, and adjustment of the original bicycle O-D matrix using a path flow estimator with an in-depth analysis into the differences between observed and estimated data.

Keywords: Bicycle demand; bicycle counts; cyclist route choice; path flow estimator

1. Introduction

In recent years, cycling has gained popularity as an alternative mode of transportation that is healthier, more environmentally friendly, and more economical than motorized vehicles. This phenomenon can be partially explained by efforts to create livable communities, and it is especially prevalent in smaller, concentrated communities such as university campuses. Livable communities aim to promote public health and to protect environmental resources at the same time. Cycling fits well into these goals because it is a zero-emission mode of transport that supports active lifestyles and health improvement. Cycling is particularly suitable for university campuses because the general layout of a university campus encourages alternative and active transportation while discouraging motorized transport (Eom et al., 2009). Campuses are also likely to have high population density, relatively short distances between trip destinations, and more bicycle facilities (Krizek et al., 2009), all of which are factors that influence bicycle mode share in a community. Since small communities such as university campus environments are so favorable for biking, there is a need to estimate these communities' bicycle demand patterns to guide bicycle facility planning.

However, transportation planners and engineers often neglect bicycles in the travel forecasting modeling process. The traditional transportation planning model is called the four-step model, which is comprised of the following processes: trip generation, trip distribution, mode choice, and traffic assignment. The four-step modelling approach has become the standard practice, but it is flawed, notably for its incapability to model non-motorized modes.

Unlike the general motorized mode, only a few studies perform bicycle network analysis (e.g., Klobucar and Fricker, 2007; Broach et al., 2011; Mekuria et al., 2012;). The Wardrop principle, 1952, indicates that motorist route choice is governed by a single dominating factor, travel time. In contrast, cyclist route choice is governed by multiple influential factors, which may include, but are not limited to, road grade, trip distance, bicycle facility type, quantity of intersections, and other safety considerations. Research indicates that travel distance or travel time and safety are both important considerations in route choice decision-making (i.e. Hopkinson and Wardman, 1996; Stinson and Bhat, 2003; Dill and Carr, 2003; Hunt and Abraham, 2007; Akar and Clifton, 2009; Broach et al., 2011, and Winters et al., 2011). However, most research efforts regarding cyclist route decision-making rely on the simplistic all-or-nothing (AON) assignment method, which focuses the analysis on a single attractiveness measure such as distance, safety, or a composite measure of safety multiplied by distance. The current bicycle modeling tools and approaches do not account for the complexities and nuances of cyclist route choice. To address this problem, Ryu et al. (2015, 2017) developed network analysis tools for estimating bicycle demand in small communities. The purpose of this paper is to apply the procedure for estimating bicycle trips at the Utah State University (or USU) campus in Logan, Utah. The bicycle demand estimation procedure consists of three main parts: data collection, adoption of a gravity model (for estimating an initial bicycle origin-destination (O-D) matrix), and adjustment of the O-D matrix with a path flow estimator.

The remainder of this paper is organized as follows. After the introduction, the methodology of the estimation of general university bicycle demand is described, followed by an application of the analysis tools at the USU campus, and some concluding remarks.

2. Methodology

There are three main phases in the procedure for estimating bicycle demand at a university campus. The first phase involves collecting data related to bicycle facilities and bicycle usage at the particular campus. The second phase adopts a doubly constrained gravity model for estimating an initial bicycle O-D matrix. The third phase uses a path flow estimator to adjust the initial bicycle O-D matrix from the second phase such that the estimated bicycle volumes from the analysis match the observed bicycle counts.

2.1 Phase 1: Data collection

The data collection process for the campus bicycle demand analysis is threefold; the analysis will require data on the campus bicycle facilities, general travel patterns, and the campus bicycle network.

2.1.1 Campus bicycle facilities

The first step in the data collection process for estimating bicycle demand at a particular university campus is to assess the current setup of campus bicycle facilities. Since the term 'bicycle facility' generally refers to facilities related to bicycle travel, this portion of the data collection will focus on bikeways, bicycle parking lots, and other bicycle-related infrastructure. Although the term ‘bikeway’ has been used as an umbrella term for routes, lanes, or paths for bicycle travel, there are notable differences between the three types of pathways. Bike routes are streets signed for bicycle use, but they are not exclusive to bicycles because motorists share the roadway with cyclists. Bike lanes and bike paths are exclusive to cyclists, but they differ in terms of proximity to motorized vehicular traffic. Bike lanes are a marked section of the roadway itself, so cyclists and motorists are essentially travelling side by side on the same road. Bike paths are physically separated from the motorized vehicular traffic, so cyclists are isolated from motorists and travel on different pathways.

The availability and quality of bicycle facilities are also important factors in promoting bicycle travel. Bicycle travel may be encouraged by improving upon on the spatial capacity, location, and security of bicycle parking. The provision of bicycle facilities such as lockers and weatherproofing roofs on bicycle parking structures may likewise increase bicycle mode share. While it is important to note the campus features that promote cycling, the factors that deter bicycle travel should also be considered. Based on bicycle level of service (BLOS) measures in the Highway Capacity Manual (HCM, 2011), these factors can be classified with respect to network infrastructure and motorized flow. In terms of network infrastructure, influential bicycle travel factors include average effective width of outside through lane, total number of directional through lanes, pavement surface condition rating, crossing distance, and number of unsignalized intersections. In terms of motorized flow, important factors include proportion of heavy vehicles in motorized vehicle volumes, effective speed factor, and directional motorized vehicle volumes.

2.1.2 Campus travel patterns

The second step in the data collection process is to gather information about travel patterns on the campus. Since the data required for this step is qualitative in nature, we recommend collecting this data through a campus travel survey that asks campus travelers (students, staff, and faculty) about all campus-related trips: their trips to campus from beyond campus boundaries, their trips from campus to other destinations beyond campus boundaries, and their trips from one campus destination to another campus destination. The goal of the survey would be to obtain the following details about the travelers’ trips: trip mode choice, average trip distance, trip origin point (geographic location at which the trip began), and trip destination point (geographic location at which the trip ended). Since this data will be crucial in calculating zonal productions and attractions later on, the data collection process should be planned carefully to ensure accurate depiction of campus travel patterns.

2.1.3 Campus network and bicycle counts

Lastly, the third step of the data collection involves building the campus bicycle network and obtaining bicycle counts at key campus locations. Building the campus bicycle network will involve digitizing an aerial photo of the campus into a network consisting of nodes and links for network analysis in a geographical information system (GIS) platform. In terms of the bicycle counts, the data collection will likely involve manual counts. Conventional traffic count collection focuses on motorized modes, so there may not be regular collection of bicycle data (such as bicycle counts on streets and intersections) at the university campus of study or in the area in general. Even though bicycle data may not be collected regularly now, we expect the practice of collecting bicycle counts to increase in the future. Many major cities (Portland, Los Angeles, San Francisco, etc.) have started collecting bicycle counts for transportation planning purposes in anticipation of higher bicycle usage.

2.2 Phase 2: Initial bicycle O-D matrix estimation

After the first phase (data collection), the second phase (initial bicycle O-D matrix estimation) begins. The second phase of the bicycle demand calculation procedure involves adopting a doubly constrained gravity model to estimate an initial bicycle O-D matrix. Details about campus-bound trips from Phase 1 are required to begin the estimation process. In particular, data about traveler mode choice upon arrival to campus and data about cyclist trip geographic origin and destination is needed. This information would then translate into mode choice proportion data, origin point proportion data, and destination point proportion data. From these data, we can construct the zonal productions and attractions that are needed as inputs to the doubly constrained gravity model. Once the flows are generated, the doubly-constrained gravity model is followed to produce the initial bicycle O-D matrix. Figure 1 shows a flowchart of the overall procedure for estimating the initial bicycle O-D demand matrix.

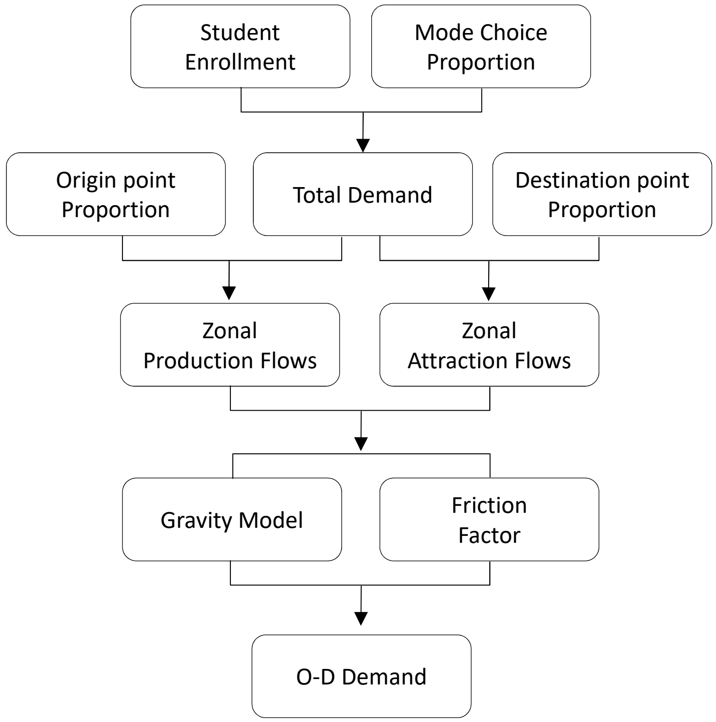


Fig. 1 Initial bicycle O-D matrix estimation flowchart

In finer detail, the initial bicycle O-D matrix estimation procedure involves the following steps: First, total demand is estimated by multiplying total enrollment and bicycle mode choice probability. Note that the O-D demand is approximated estimated by multiplying total enrollment in Phase 2 then the demand is adjusted with observed counts data in Phase 3. Then, zonal productions and attractions are estimated by multiplying the estimated total demand from the first step with origin and destination proportion data. The zonal productions and attractions trip data can then be used as inputs in a trip distributed model, which would generate a trip table using the trip ends produced from the trip generation models and network attributes (e.g., interzonal travel times). Traditionally, the gravity model has been used for the trip distribution model (e.g., the doubly-constrained gravity model) as follows.

$$T_{rs} = \frac{P_r}{\sum_{r \in R} P_r \cdot K_r \cdot F_{rs}} \cdot \frac{A_s}{\sum_{s \in S} A_s \cdot K_s \cdot F_{rs}} \cdot F_{rs}, \tag{1}$$

where T_{rs} = Number of trips produced in zone r and attracted to zone s;

P_r = Number of trips produced in zone r;

A_s = Number of trips attracted to zone s;

F_{rs} =An empirically derived "*friction factor*", which expresses the average area-wide effect of spatial separation on the trip interchanges between the two zones, *r* and *s*;

K_r , K_s = Empirically origin and destination adjustment factors or balancing factors, which are solved iteratively to account for the travel pattern.

In the absence of friction factors in the university area, we substitute the friction factors using the bicycle commute trip length introduced by Aultman-Hall et al. (1997). This bicycle commute trip length has a mean of 2.3 miles and a standard deviation of 1.24 miles. Finally, the doubly-constrained gravity model is performed using the zonal productions and attractions, and friction factors described above.

2.3 Phase 3: Refinement of the Initial Bicycle O-D Matrix

The third phase of the bicycle demand estimation process involves refining the initial bicycle O-D demand from Phase 2. The initial bicycle O-D demand matrix in Phase 2 is estimated based on a trip distribution model, (i.e., a doubly-constrained gravity model) that requires calibration of the friction factors for the study area. However, modelers often do not calibrate trip distribution models due to the unavailability of Trip Length Frequency Distribution (TLFD) data. To circumvent the need for TLFD data and still conduct calibration, modelers can arbitrarily alter friction factors and add k-factors to the trip distribution model such that the results of the traffic assignment would match the traffic counts on selective screenlines and critical links. Since this calibration procedure is hit-or-miss, it usually is a lengthy process. Moreover, the resultant models from the calibration process often contain factors that do not have the necessary behavioral foundations established from travel surveys.

For this study, we adopt the two-stage bicycle O-D demand adjustment procedure developed by Ryu et al. (2015). Stage 1 is responsible for the determination of efficient (or non-dominated) routes that represent the optimal tradeoffs between route distance and route bicycle level of service (BLOS) by generating a Pareto set of routes. It generates a set of efficient routes for each O-D pair by using a bi-objective shortest path algorithm (Ehrgott et al., 2012). Stage 2 is responsible for the adjustment of O-D demand based on observed bicycle counts and obtained zonal productions, zonal attractions and O-D demand estimated in Phase 2. It uses a path flow estimator (PFE) (Bell et al., 1997, Chootinan et al., 2005; Chen et al. 2005, 2009, 2010; Ryu et al.,

2014) to refine the initial bicycle O-D matrix from Phase 2 such that the readjusted, final bicycle O-D matrix can reproduce better matches with the observed bicycle counts when performing the bicycle traffic assignment procedure. The flexibility of aggregating path flows at different spatial levels, which allows the usage of various data (e.g., bicycle intersection counts, bicycle link counts, bicycle GPS data, bicycle miles traveled (BMT), etc.), makes the PFE a suitable approach for improving the accuracy of bicycle O-D estimation. In this study, we adopt the path-size logit (PSL) PFE model to consider the route overlapping issue. The PSL-based PFE formulation can be formulated with route utilities as a convex program with various side constraints as follows

$$Minimize: Z(\mathbf{f}) = \sum_{rs \in \mathrm{RS}} \sum_{k \in \mathrm{K}_{rs}} f_k^{rs} (\ln f_k^{rs} - 1) - \sum_{rs \in \mathrm{RS}} \sum_{k \in \mathrm{K}_{rs}} U_k^{rs} f_k^{rs} - \sum_{rs \in \mathrm{RS}} \sum_{k \in \mathrm{K}_{rs}} f_k^{rs} \ln PS_k^{rs} \tag{2}$$

subject to:

$$(1-\varepsilon_a) \cdot v_a \le x_a \le (1+\varepsilon_a) \cdot v_a, \qquad \forall\, a \in \bar{\mathrm{A}}, \tag{3}$$

$$(1-\varepsilon_r) \cdot O_r \le P_r \le (1+\varepsilon_r) \cdot O_r, \qquad \forall r \in \bar{\mathrm{R}}, \tag{4}$$

$$(1-\varepsilon_s) \cdot D_s \le A_s \le (1+\varepsilon_s) \cdot D_s, \qquad \forall s \in \bar{\mathrm{S}}, \tag{5}$$

$$(1-\varepsilon_{rs}) \cdot z_{rs} \le q_{rs} \le (1+\varepsilon_{rs}) \cdot z_{rs}, \qquad \forall\, rs \in \overline{\mathrm{RS}}, \tag{6}$$

$$f_k^{rs} \ge 0, \quad \forall k \in \mathrm{K}_{\mathrm{rs}}, rs \in \mathrm{RS}, \tag{7}$$

where

$$x_a = \sum_{rs \in \mathrm{RS}} \sum_{k \in \mathrm{K}_{\mathrm{rs}}} f_k^{rs} \delta_{ka}^{rs}, \qquad \forall\, a \in \mathrm{A}, \tag{8}$$

$$P_r = \sum_{s \in \mathrm{S}} \sum_{k \in \mathrm{K}_{\mathrm{rs}}} f_k^{rs}, \qquad \forall r \in \mathrm{R}, \tag{9}$$

$$A_s = \sum_{r \in \mathrm{R}} \sum_{k \in \mathrm{K}_{\mathrm{rs}}} f_k^{rs}, \qquad \forall s \in \mathrm{S}, \tag{10}$$

$$q_{rs} = \sum_{k \in \mathrm{K}_{\mathrm{rs}}} f_k^{rs}, \qquad \forall\, rs \in \mathrm{RS}, \tag{11}$$

where f_k^{rs} = Flow on route *k* between origin *r* and destination *s;*

U_k^{rs} = Utility route *k* between origin *r* and destination *s*;

PS_k^{rs} = Path size factor on route *k* between origin *r* and destination *s*;

v_a = Observed count on link *a*;

x_a = Estimated flow on link *a*;

ε_a = Percentage of measurement error allowed for the traffic count on link *a*;

$\bar{A}$ = Set of network links with measurements;

O_r and D_s = Generated trip production of origin *r* and generated trip attraction of destination *s* estimated from the survey data;

P_r and A_s = Estimated trip production of origin *r* and trip attraction of destination *s*;

ε_r and ε_s = Error bounds allowed for trip production of origin *r* and trip attraction of destination *s*;

$\bar{R}$ and $\bar{S}$ = Sets of zones with planning data;

z_{rs} = Target O-D flows (i.e., initial O-D demand matrix estimated from Stage 1) between origin *r* and destination *s*;

q_{rs} = Adjusted O-D flows between origin *r* and destination *s*;

ε_{rs} = Percentage measurement error allowed for the target O-D demands between origin *r* and destination *s*;

$\overline{RS}$ = Set of target (or initial) O-D pairs;

δ_{ka}^{rs} = Path-link indicator, 1 if link a is on path *k* between O-D pair *rs* and 0 otherwise.

The objective function (2) consist of three terms: an entropy term, a system optimal term, and a path size (PS) term. The entropy term seeks to spread trips onto multiple routes according to the dispersion parameter; the system optimal term clusters trips based on the minimum utilities; and the PS term penalizes routes with overlap.

To provide the analytical path flow solution, the Lagrangian function is shown in equations (12) - (17).

$$
\begin{aligned}
L = Z + & \sum_{a \in \bar{A}} u_a^- \cdot \left(v_a \left(1-\varepsilon_a\right) - \sum_{rs \in \mathrm{RS}} \sum_{k \in \mathrm{K}_{rs}} f_k^{rs} \delta_{ka}^{rs} \right) + \sum_{a \in \bar{A}} u_a^+ \cdot \left(v_a \left(1+\varepsilon_\mathrm{a}\right) - \sum_{rs \in \mathrm{RS}} \sum_{k \in \mathrm{K}_{rs}} f_k^{rs} \delta_{ka}^{rs} \right) \\
& + \sum_{r \in \bar{R}} \rho_r^- \cdot \left(O_r \left(1-\varepsilon_r\right) - \sum_{s \in \mathrm{S}} \sum_{k \in \mathrm{K}_{rs}} f_k^{rs} \right) + \sum_{r \in \bar{R}} \rho_r^+ \cdot \left(O_r \left(1+\varepsilon_r\right) - \sum_{s \in \mathrm{S}} \sum_{k \in \mathrm{K}_{rs}} f_k^{rs} \right) \\
& + \sum_{s \in \bar{S}} \alpha_s^- \cdot \left(D_s \left(1-\varepsilon_s\right) - \sum_{r \in \mathrm{R}} \sum_{k \in \mathrm{K}_{rs}} f_k^{rs} \right) + \sum_{s \in \bar{S}} \alpha_s^+ \cdot \left(D_s \left(1+\varepsilon_s\right) - \sum_{r \in \mathrm{R}} \sum_{k \in \mathrm{K}_{rs}} f_k^{rs} \right) \\
& + \sum_{rs \in \overline{RS}} o_{rs}^- \cdot \left(z_{rs} \left(1-\varepsilon_{rs}\right) - \sum_{k \in \mathrm{K}_{rs}} f_k^{rs} \right) + \sum_{rs \in \overline{RS}} o_{rs}^+ \cdot \left(z_{rs} \left(1+\varepsilon_{rs}\right) - \sum_{k \in \mathrm{K}_{rs}} f_k^{rs} \right)
\end{aligned} \tag{12}
$$

where $u_a^+, u_a^-, \rho_r^-, \rho_r^+, \alpha_s^-, \alpha_s^+, o_{rs}^-$ and o_{rs}^+ are the dual variables of constraints (3), (4), (5) and (6), respectively. The first-order partial derivative with respect to the path-flow variable can be expressed as follows:

$$
\begin{aligned}
\frac{\partial L}{\partial f_k^{rs}} = 0 \Rightarrow & \frac{1}{\theta} \ln f_k^{rs} - U_k^{rs} - \ln PS_k^{rs} - \sum_{a \in \bar{A}} u_a^- \delta_{ka}^{rs} - \sum_{a \in \bar{A}} u_a^+ \delta_{ka}^{rs} - \rho_r^- - \rho_r^+ - \alpha_s^- - \alpha_s^+ - \\
& o_{rs}^- - o_{rs}^+ = 0, \qquad \forall\, k \in \mathrm{K}_{rs}, rs \in \mathrm{RS}
\end{aligned} \tag{13}
$$

From Equation (13), we can obtain the path-flow solution:

$$
\begin{aligned}
f_k^{rs} = PS_k^{rs} \cdot \exp\left(\theta \left(U_k^{rs} + \sum_{a \in \bar{A}} \left(u_a^- + u_a^+ \right) \cdot \delta_{ka}^{rs} + o_{rs}^- + o_{rs}^+ \right) \right) \cdot \exp\left(\theta \left(\rho_r^- + \rho_r^+ \right) \right) \cdot \exp\left(\theta \left(\alpha_s^- + \alpha_s^+ \right) \right), \\
\forall\, k \in \mathrm{K}_{rs}, rs \in \mathrm{RS}
\end{aligned} \tag{14}
$$

Using Equation (11), the demand of O-D pair *rs* is

$$
q_{rs} = \sum_{k \in \mathrm{K}_{rs}} PS_k^{rs} \cdot \exp\left(\theta \left(U_k^{rs} + \sum_{a \in \bar{A}} \left(u_a^- + u_a^+ \right) \cdot \delta_{ka}^{rs} + o_{rs}^- + o_{rs}^+ \right) \right) \cdot \exp\left(\theta \left(\rho_r^- + \rho_r^+ \right) \right) \cdot \exp\left(\theta \left(\alpha_s^- + \alpha_s^+ \right) \right),
$$

$$
\forall\, rs \in \mathrm{RS}. \tag{15}
$$

Using Equation (14) and recognizing

$$
\begin{aligned}
& \frac{P_r}{\sum_{r \in \mathrm{R}} P_r \cdot K_r \cdot F_{rs}} = \exp\left(\theta \left(\rho_r^- + \rho_r^+ \right) \right), \quad \frac{A_s}{\sum_{s \in \mathrm{S}} A_s \cdot K_s \cdot F_{rs}} = \exp\left(\theta \left(\alpha_s^- + \alpha_s^+ \right) \right), \\
& \pi_{rs} = -\frac{1}{\theta} \ln \sum_{k \in \mathrm{K}_{rs}} PS_k^{rs} \cdot \exp\left(\theta \left(U_k^{rs} + \sum_{a \in \bar{A}} \left(u_a^- + u_a^+ \right) \cdot \delta_{ka}^{rs} + o_{rs}^- + o_{rs}^+ \right) \right)
\end{aligned} \tag{16}
$$

Hence, Equation (16) can be rewritten as the optimal solution of a doubly-constrained gravity model with a negative exponential deterrence function for the trip distribution step as follows:

$$q_{rs} = \frac{P_r}{\sum_{r \in \mathrm{R}} P_r \cdot K_r \cdot F_{rs}} \cdot \frac{A_s}{\sum_{s \in \mathrm{S}} A_s \cdot K_s \cdot F_{rs}} \cdot \exp\left(-\theta \pi_{rs}\right), \quad \forall\, rs \in \mathrm{RS} \tag{17}$$

The proposed PFE formulation with observed link counts constraints, zonal productions and attractions constraints and initial O-D demand constraints from Equations (3) to (6) can be considered as an extension of the combined distribution and assignment problem (Evans, 1976), in which the trip distribution follows a doubly constrained gravity model with a negative exponential impedance function and the traffic assignment follows the logit-based SUE model (Lundgren and Patriksson, 1998). Furthermore, the proposed PFE formulation can be interpreted as the most probable flow pattern under the efficient trip making behavior defined by the efficiency principle (Smith, 1978, 1983). Travelers generally prefer higher utility (or lower cost) than lower utility (or higher cost).

The use of error bounds in Equation (3)-(6) allows the inclusion of local knowledge and experience of the user to provide inputs to the PFE formulation to reflect the conditions within the study area (i.e., a smaller error bound indicates the data is more reliable, while a larger error bound indicates the opposite). Hence, the error bounds should be specified judiciously. If the error bounds are set too small, it may lead to infeasibility (i.e., a solution does not exist for such tight constraints). On the other hand, if the error bounds are set too large, it may lead to a biased solution with underestimated flows. If the user has difficulty in specifying reasonable error bounds, an alternative PFE with different norm approximations could be used to determine appropriate error bounds. See Chen et al. (2009) for the details of the different L_p-PFE models, where p could be 1, 2, or infinity (Chen et al., 2010).

3. Case study – USU campus

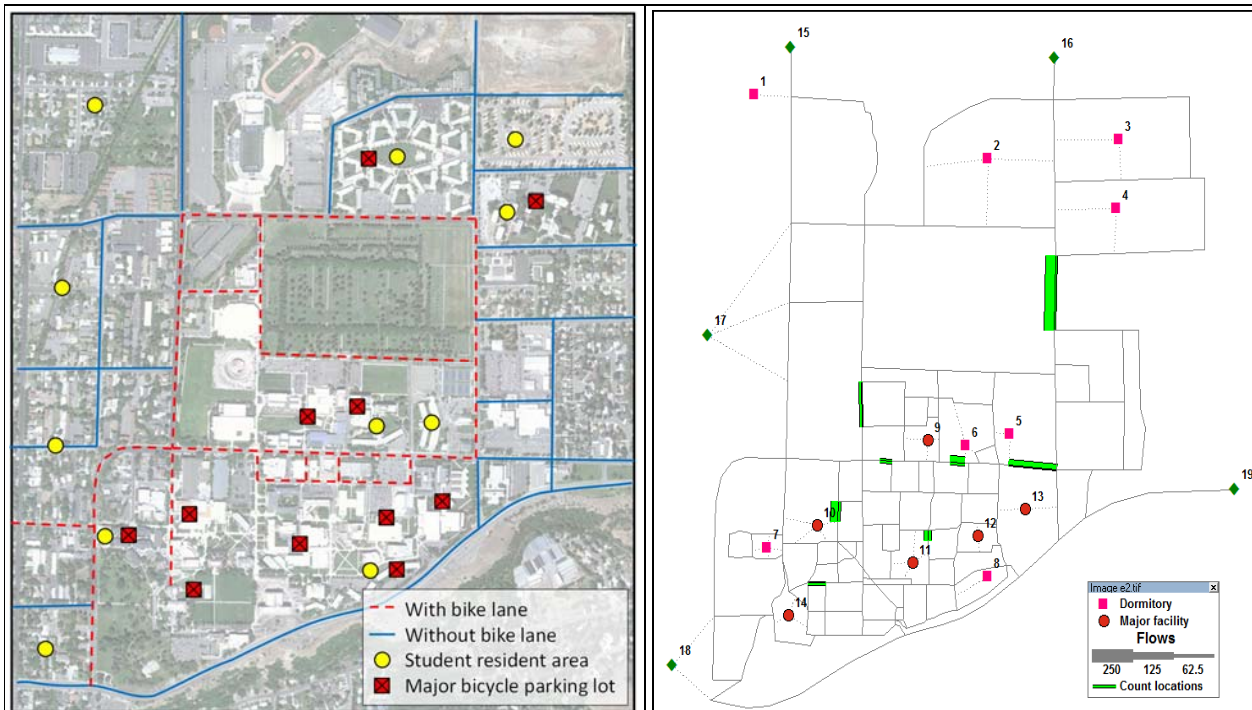
In this section, a case study of the Utah State University (USU) campus is conducted to demonstrate the applicability of the bicycle demand estimation process described above. The three-phase procedure of data collection, bicycle O-D demand generation, and adjustment of the bicycle O-D matrix is followed closely to analyze USU bicycle demand.

3.1 Results from Phase 1 (Data collection)

In the first phase of the bicycle demand analysis, we collected bicycle travel data for Utah State University (USU), focusing on student transportation mode choice and student geographic origin (or starting) points of campus-bound cycling trips.

Figure 2(a) shows an aerial overview map of USU and marks major student residential areas, bicycle parking lots, and bikeways. For scale, the university spans about 500 acres (2.0 km^2), in Logan, Utah and is home to more than 14,000 students who live on or near campus. In terms of bicycle parking, most buildings on campus feature bicycle parking structures that hold a range of 10 to 100 individual parking facilities. The residential area data in Figure 2 is based on student dormitory data for both campus and off-campus housing, and the bikeway data (e.g., bike lane) is gathered from the 2035 Regional Transportation Plan (CMPO, 2011).

Figure 2(b) summarizes our bicycle count results, which were conducted on a daily basis at specific locations on the campus. To obtain bicycle traffic information for the USU campus, we manually collected bicycle counts during September of 2014. However, since bicycle counts are not collected regularly at the USU campus, our bicycle count data may not be an accurate representation of actual cycling activity. If there are issues with data inconsistency between bicycle counts, it would affect the estimation of the bicycle O-D demand matrix.

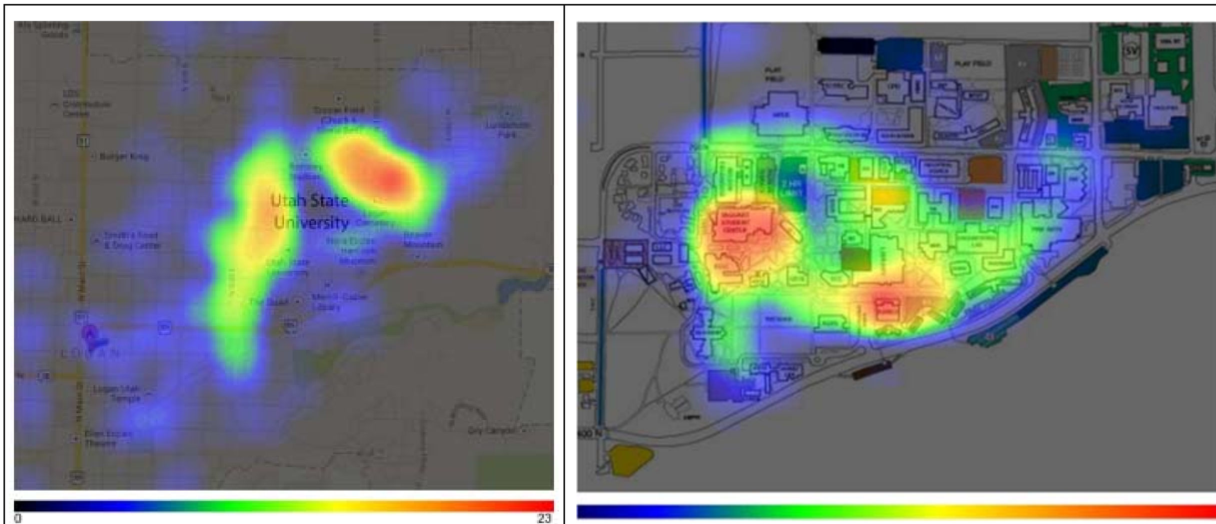


(a) Map of USU campus with bicycle facilities

(b) Campus network and bicycle counts

Fig. 2 Map of USU campus bicycle network, facilities, and counts

In addition to gathering data on campus bicycle infrastructure and facilities, we also used information from a 2014 USU transportation survey. 4,469 students (who make up approximately 26% of total USU enrollment) participated in the survey. Among other topics, the survey inquired about the students' mode choice upon arrival to campus and allowed for multiple mode choice selections ranging from walking, biking, riding motorcycles or scooters, driving, and using public transit. Survey results revealed that approximately 11% of the surveyed students traveled to campus by biking. Figure 3 gives a geographic overview of campus-bound bicycle trips. Overall results are referred in 2014 Transportation Survey Results (Utah State University, 2014).



(a) Map of students' starting point when biking to campus

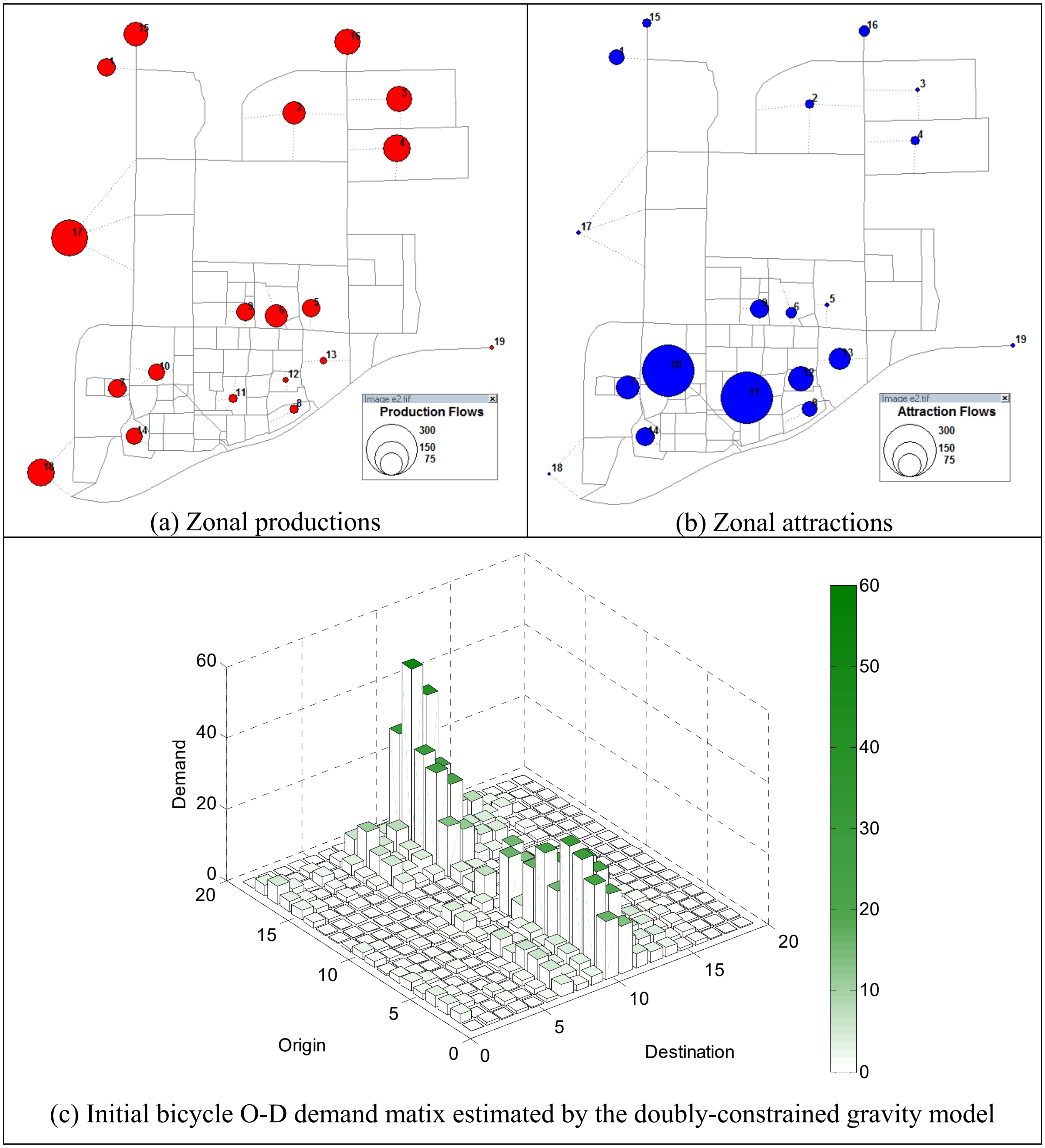
(b) Map of campus locations of high demand for more bicycle parking

Fig. 3 Map of starting points and arrival points for bicycle trips to campus

3.2 Results from Phase 2 (Initial bicycle O-D demand estimation)

In this section, we follow the procedures for Phase 2 to generate the initial bicycle O-D demand matrix for the USU network using data collected from Phase 1.

First, we estimate total demand by multiplying total enrollment (approximately 14,000 students) by bicycle mode choice probability (approximately 11%, according to survey results). This calculation generates 1,470 bicycle trips. Next, we estimate the zonal productions and the zonal attractions. To obtain the zonal productions, we multiply the total demand (1,470 bicycle trips) from the first step by the proportion of origin data from Figure 1. Similarly, to obtain the zonal attractions, we multiply the total demand (1,470 bicycle trips) from the first step by the proportion of destination data from Figure 1. Figures 4(a) and 4(b) present the estimated productions and attractions obtained in this step. Then, we use the estimated zonal productions and attractions with the bicycle commute trip length from Aultman-Hall et al. (1997) as inputs in the trip distribution model to obtain the initial bicycle O-D demand matrix. Figure 4(c) presents the initial bicycle O-D demand matrix estimated by the doubly-constrained gravity model.



(a) Zonal productions (b) Zonal attractions

(c) Initial bicycle O-D demand matix estimated by the doubly-constrained gravity model

Fig. 4 Estimated zonal productions, attractions, and initial bicycle O-D demand matix

3.3 Phase 3 Overview (Adjustment of Bicycle O-D Matrix)

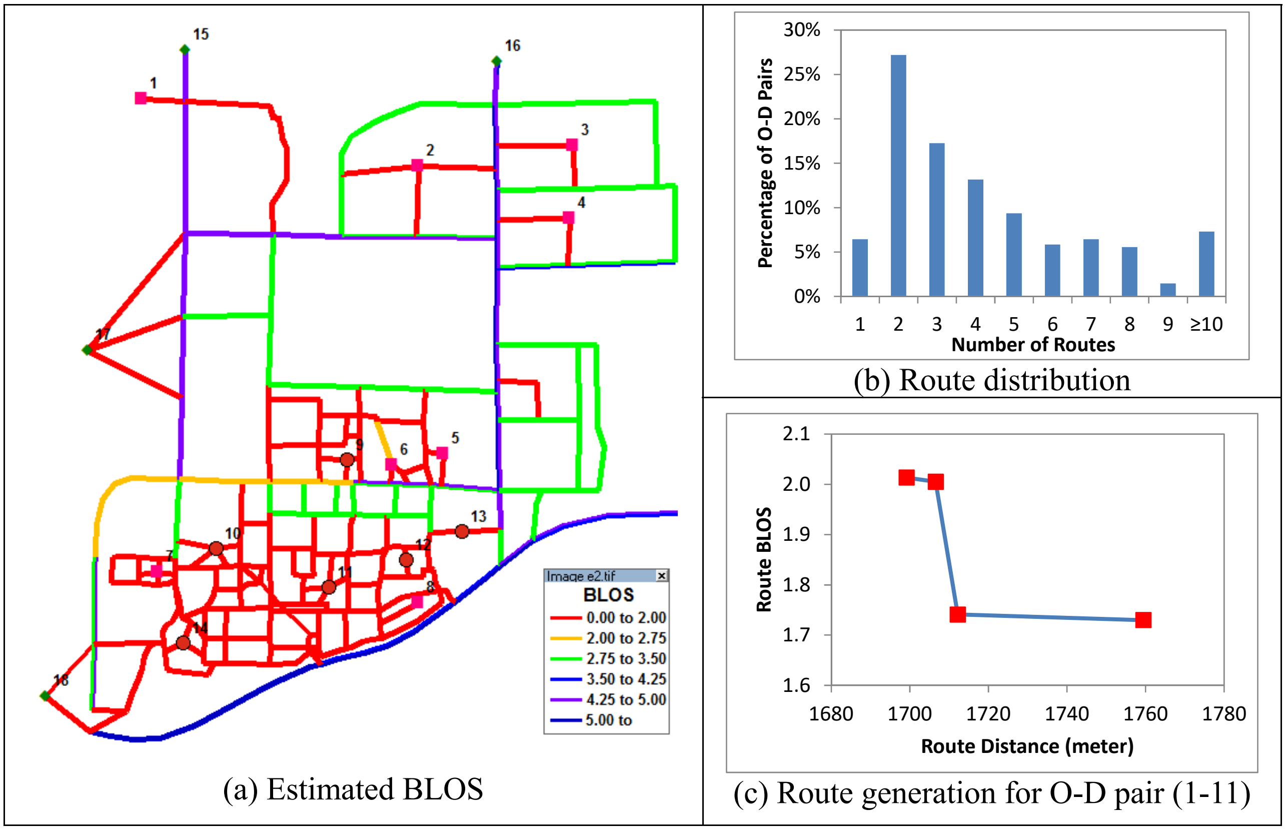
After generating the initial bicycle O-D matrix in Phase 2, we followed the procedure for Phase 3 to readjust the initial bicycle O-D matrix. The readjustment process goes through two stages: the first stage uses a bi-objective shortest path algorithm from Ehrgott et al. (2012) to generate the route set for each O-D pair, and the second stage uses a path flow estimator (see Ryu et al. (2014) for details) to adjust the initial bicycle O-D matrix from Phase 2 using the route set from the first stage.

Once the two-stage readjustment process is completed, we perform three key comparisons to show the benefits and the necessity of refining the initial bicycle demand matrix. In the first comparison, we compare the link flows observed from Phase 1 (the data collection process) to the link flows estimated by the readjustment process from Phase 3. In the second comparison, we compare the initial bicycle demand matrix from Phase 2 with the adjusted matrix from Phase 3. Lastly, in the third comparison, we compare the differences in bicycle traffic assignment resulting from the initial and adjusted bicycle demand matrices.

3.4 Results from Stage One: Bicycle BLOS analysis and route generation

As described above, the first stage of the bicycle O-D matrix adjustment process deals with BLOS analysis and route generation. To compute the BLOS measures, we used the route BLOS equation developed by the Highway Capacity Manual (HCM, 2011) using network topology data (e.g., road distance, road width, road slope, intersection, etc.), bicycle facility data (e.g., bike lane, bike path, bike parking, etc.), and traffic data (e.g., motorized vehicle volumes, speed limit, roadside parking, etc.) from the Cache Metropolitan Planning Organization (CMPO) planning model as inputs (CMPO, 2014). BLOS can serve as a surrogate measure for safety as it accounts for different attributes that are related to the safety of bicycle routes. Segments with higher motorized vehicle volumes typically generate higher BLOS values, while segments exclusive to non-motorized modes have lower BLOS values.

After the BLOS analysis, we followed the bi-objective route generation procedure to generate efficient routes for each O-D pair (efficiency is in terms of route distance and route BLOS). Figure 5 summarizes the results from stage one of Phase 3: Figure 5(a) shows the estimated BLOS results in the USU network; Figure 5(b) presents the distribution of the routes (e.g. about 6% of the O-D pairs have only one route, about 27% have two routes, about 18% have three routes, etc.); Figure 5(c) displays an example of route generation for O-D pair 1-11 (i.e., there are four efficient routes between this O-D pair).



(a) Estimated BLOS

(b) Route distribution

(c) Route generation for O-D pair (1-11)

Fig. 5 Network BLOS and generated routes

3.5 Results from Stage Two: Bicycle PFE

The second stage of the bicycle O-D matrix adjustment process deals with bicycle traffic assignment. Using the efficient routes generated from Stage 1, the bicycle PFE is performed by adjusting the initial bicycle O-D demands to match the observed data (bicycle counts from Figure 2(b), zonal productions and attractions from Figures 4(a) and 4(b), and target O-D demand from Figure 4(c) as much as possible. For the utility function, we adopted a composite utility function of route distance multiplied by route BLOS, $U_k^{rs} = -\left(\left(d_k^{rs}\right)^{\alpha} \cdot \left(BLOS_k^{rs}\right)^{\beta}\right)$, where U_k^{rs} is the utility of route k between O-D pair rs; d_k^{rs} is the distance of route k between O-D pair rs; and $BLOS_k^{rs}$ is the bicycle level of service of route k between O-D pair rs; α and β are parameters.

In this study, we adopted the parameter values (i.e., $\alpha = 0.862$ and $\beta = 0.117$) suggested by Kang and Fricker (2013). Since the path flow estimator (PFE) does not require the estimated flows to reproduce the observed values, we adopted different error bounds for different data types as shown in Table 1. As the reliability of a dataset increases, the size of the error bound decreases. For bicycle counts, a uniform error bound of 30% is used for the PFE estimation. For zonal productions, attractions, and target O-D demand (i.e., initial bicycle O-D demand matrix), a flow-dependent error bound is adopted (i.e., larger flow values are constrained to a smaller error bound as shown in Table 1).

Table 1. User-defined error bounds for different data types

Link Count	Zonal production and Attraction		Target O-D demand	
	Flows	Error bound	Demand	Error bound
30%	0≤ and <10	50%	0≤ and <5	Free
	10≤ and <30	40%	5≤ and <10	40%
	30≤ and <50	30%	10≤ and <30	30%
	50≤	20%	30≤	20%

Comparisons of Analysis Results

After completing the two-stage adjustment process of Phase 3, we analyze the results through the following three comparisons: (1) observed link flows from Phase 1 to estimated link flows from Phase 3, (2) initial bicycle matrix from Phase 2 to adjusted matrix from Phase 3, and (3) bicycle traffic assignment resulting and the initial and adjusted bicycle demand matrices.

In our first comparison, we compare the link flows observed from Phase 1 (the data collection process) to the link flows in Phase 3 (the readjustment process). Figure 6 shows the scatter plots of observed and estimated link flows in (a), zonal productions in (b), zonal attractions in (c), and target O-D demand in (d). To measure the accuracy, the root mean squared error (RMSE) value is adopted here.

Figure 6 (a), (b), (c) and (d) show that all estimated link flows, zonal productions, zonal attractions and O-D demand match the observed values with satisfactory error bounds. The RMSE values for estimated link flows, zonal productions, zonal attractions and O-D demand are 18.57, 15.47, 22.60 and 4.12, respectively. While most of the estimated flows match the observed flows within certain error bounds, Figure 6(b) shows that the university library is the only point that is out of the error bound. Since it is the nature of libraries to have many traversing visitors, there should be higher trips for both attractions and productions. However, the travel survey (i.e., see

Figure 3) only surveys the original starting point of a bicycle trip. This indicates that there is a data inconsistency problem between zonal productions and observed link counts.

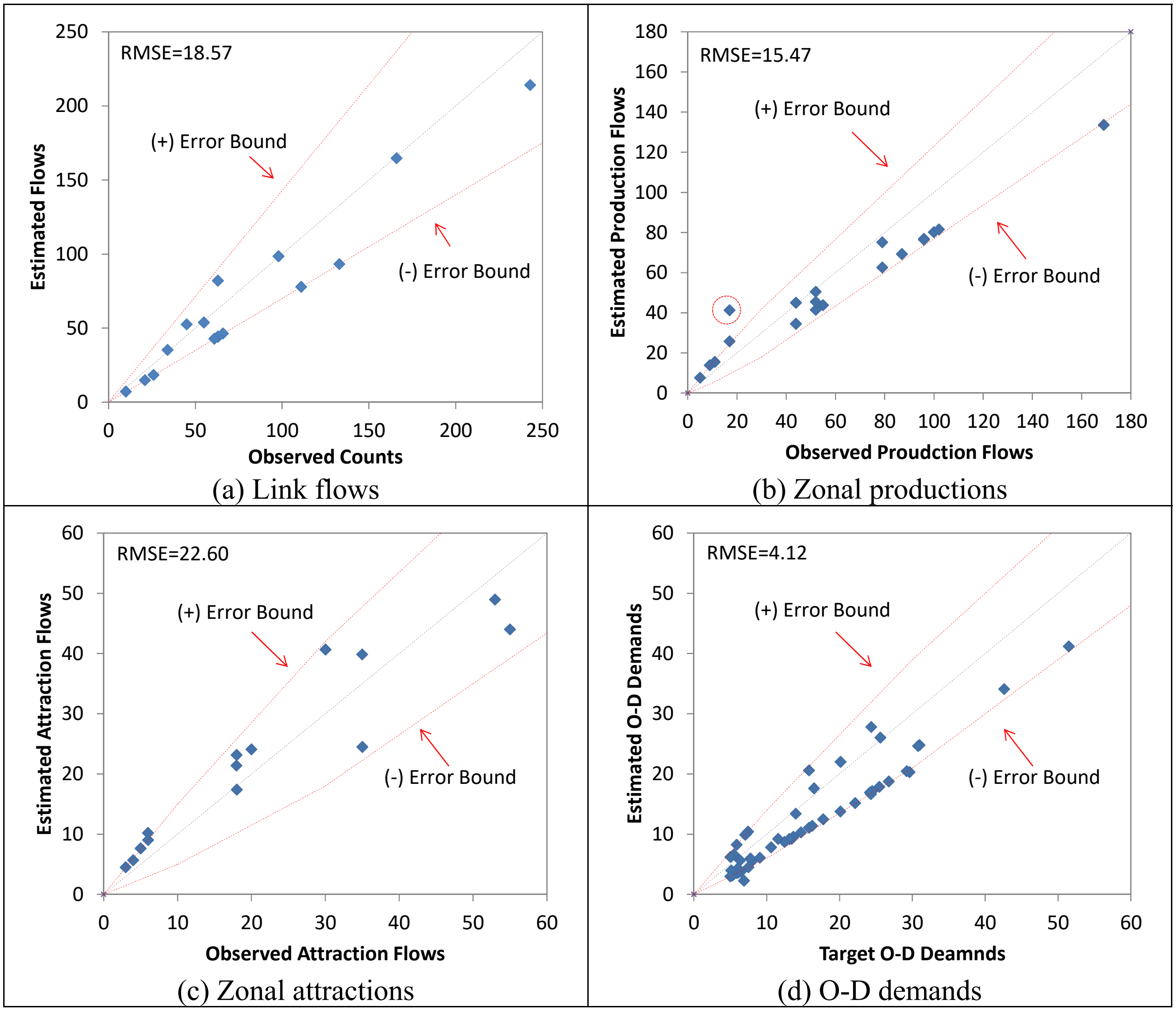
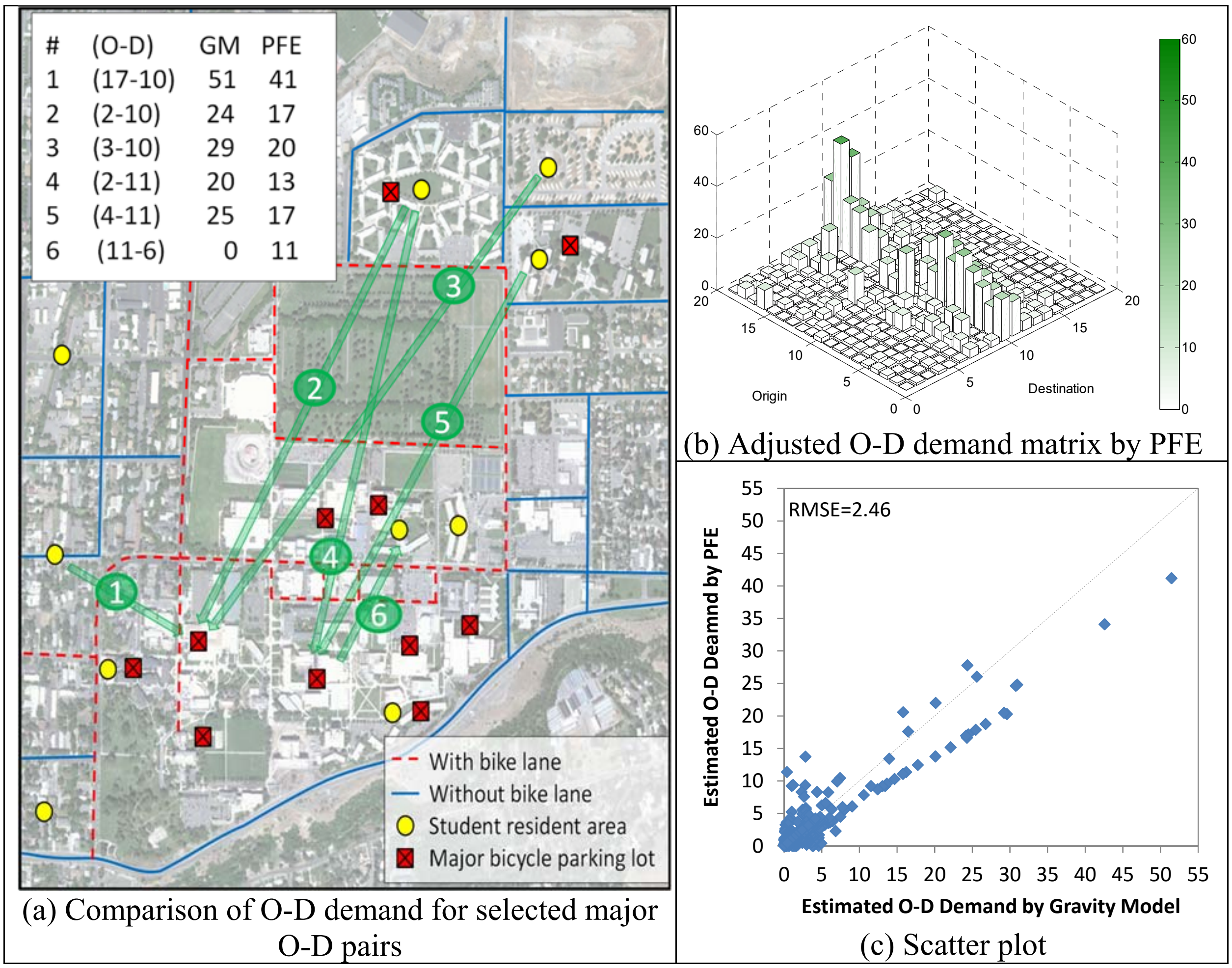


Fig. 6 Comparison between observed and estimated flows

For our second comparison, we compare the estimated bicycle O-D demand matrix from Stage 2 with the adjusted bicycle O-D demand matrix from Stage 3 and examine the bicycle traffic assignment results from both bicycle O-D demand matrices. It should be noted that the path-size logit (PSL) model was used for bicycle traffic assignment in the initial bicycle O-D demand matrix. The two-stage path flow estimator (PFE) model was used for bicycle traffic assignment and adjustment of the initial bicycle O-D demand matrix.

Figure 7 shows the adjusted O-D demand by PFE in (b) and compares it with the generated O-D by gravity model in (a) and (c). Based on Figure 7(b), the overall adjusted O-D demand is not significantly different compared to Figure 4(c) as there are many O-D pairs with very small demand (i.e., less than 5 trips).



(a) Comparison of O-D demand for selected major O-D pairs

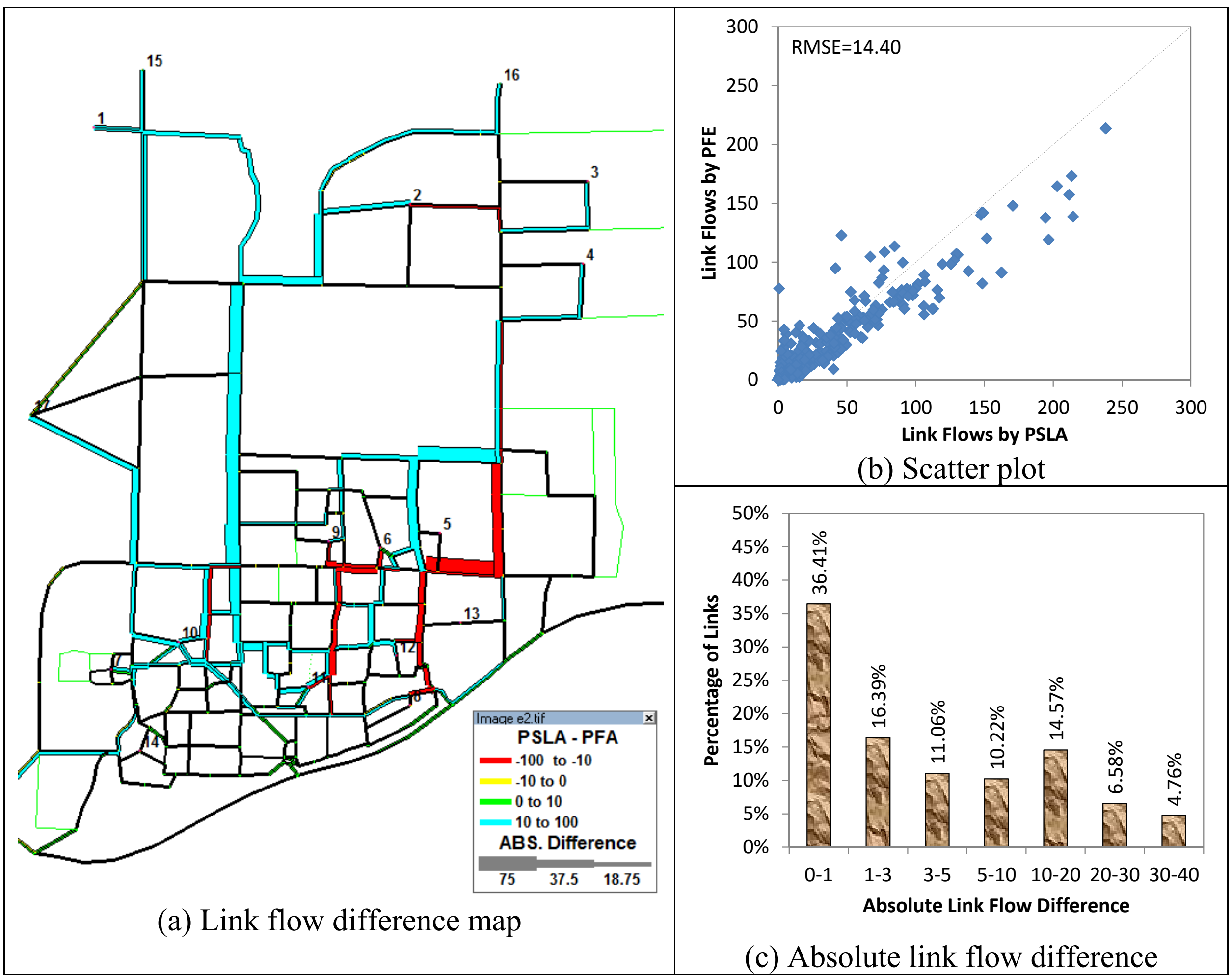
(b) Adjusted O-D demand matrix by PFE

(c) Scatter plot

Fig. 7 Comparison of O-D demand from Phase 1 (gravity model) and Phase 2 (PFE model)

The RMSE value is 2.46 in Figure 7(c). However, we can observe that some O-D pairs are significantly affected with about 10 flow differences in Figure 7(a). The O-D pairs related to the student center (zone number 10) and library (zone number 11) have significantly different demand patterns. This is because these two zones have relatively higher demand and these two observed count locations (four directional counts) are directly related to the zones.

In our third and final comparison, we look at the differences in traffic assignment results generated by the initial and the adjusted matrices. The initial matrix uses the Path Size Logit Assignment (PSLA) model, and the adjusted matrix uses the Path Flow Estimator (PFE) model. Figure 8 provides the flow allocation comparison between the PSLA and PFE models. Figure 8(a) depicts the flow difference on a color-coded GIS map. Red indicates that the link flows by the PFE model have higher flows than the PSLA model. The light blue (or cyan) color indicates the reverse; it indicates that the link flows by PSLA have higher flows than the PFE model.



(a) Link flow difference map

(b) Scatter plot

(c) Absolute link flow difference

Fig. 8 Link flow comparison between using PSLA and PFE

Figure 8(a) shows that the left-side links have more assigned flows by the PSLA model, while the right-side links have more assigned flows by the PFE model. Figure 8(b) shows the scatter plot of estimated flows by both models. The patterns may seem similar at first, but it is significantly different because bicycle flows are small. The RMSE value is 14.40. Figure 8(c) presents the absolute link flows difference. Only 36% (0-1) of links has similar flow patterns, while about 11% (i.e. greater than 20 flows) of links has significantly different flow patterns.

4. Concluding remarks and recommendations

There are many benefits to cycling. Cycling can support a healthy, active lifestyle and provide an alternate mode of transportation to the motorized vehicle. In addition, cycling can have positive impacts on public health, air quality (through the reduction of emissions), urban mobility, and livability. The 2009 National Households' Travel Survey (NHTS) (Kuzmyak et al., 2014) may state that the national mode share of bicycles is only slightly more than 1% of all trips nationwide, but the American Community Survey (ACS) shows that cycling is regaining popularity as a mode of transportation in recent years. Between 2008 to 2012, bicycle commuting increased by 61.6%, which is the largest percentage increase observed in any other commuting mode (McKenzie, 2014).

This paper presents a case study of estimating bicycle demand at the Utah State University (USU) campus in Logan, Utah using the methodology developed by Ryu et al. (2015). The estimation process involves three key phases: data collection, initial bicycle O-D estimation, and adjustment of the original bicycle O-D matrix. In the USU case study, we found that the incorporation of observed bicycle counts into the bicycle O-D matrix adjustment process yields some significantly different link flows. The differences in estimated and observed link flows may be partially explained by the absence of calibration for the adopted parameters in the trip distribution and traffic assignment models and by the data inconsistency problem between observed bicycle counts and estimated zonal flows.

This research and its findings are relevant to transportation planners and engineers who work for small communities such as university campuses; it may inform decision-making on a wide variety of bicycle capital and infrastructure topics, such as bicycle facility assessment, bicycle route guidance, bicycle counting location identification, and others. The three-stage process proposed in this research has been demonstrated to be useful for bicycle modeling and may have other applications, such as guiding the design of the larger transportation network to better accommodate bicycle travel. In future research concerning bicycle demand estimation, we recommend the following: (1) conducting parameter calibration; (2) including a greater variety of network topologies and bicycle facilities; (3) integrating analysis for motorized congestion as well as bicycle congestion; (4) accounting for travelers' characteristics and behavior; and (5) creating and implementing a more comprehensive data collection process to have a stronger observed count base and to avoid data inconsistency problems. Another consideration is observed counts data; compared to motorized modes, there may not be enough observed counts data for the bicycle mode for a large size network. This limitation could be addressed through the adoption of fuzzy theory as an alternative for consideration of data quality and user perception (Quattrone and Vitetta, 2011).

ACKNOWLEDGMENT

This research was supported by the Research Committee of the Hong Kong Polytechnic University (Project No. 1-ZE5T), the Research Grants Council of the Hong Kong Special Administrative Region (Project No. PolyU 115212217), the Transportation Research Center for Livable Communities (TRCLC), the Mountain-Plains Consortium (MPC) sponsored by the U.S. Department of Transportation, the Basic Science Research Program through the National Research Foundation (NRF) of Korea by the Ministry of Science (NRF-2016R1C1B2016254), and ICT & Future Planning (NRF-2010-0029443). These supports are gratefully acknowledged.

REFERENCES

1. *2014 Transportation survey result*, Utah State University Parking & Transportation, 2014 <http://parking.usu.edu/files/uploads/2014_Transportation_Report.pdf.> (February 20, 2014.)
2. Akar, G., and Clifton, K.J. (2009). "Influence of individual perceptions and bicycle infrastructure on decision to bike." *Transportation Research Record*, 2140, 165-171.
3. Aultman-Hall L., Hall, F., and Baetz, B. (1997). "Analysis of bicycle commuter routes using geographic information systems implications for bicycle planning." *Transportation Research Record 1578,* 102-110.
4. Bell, M.G.H., Shield, C.M., Busch, F., and Kruse, G. (1997). "A stochastic user equilibrium path flow estimator." *Transportation Research Part C*, 5(3-4), 197-210.
5. Broach, J., Gliebe, J., and J. Dill. (2011). "Bicycle route choice model developed using revealed preference GPS data." *Proc., Presented at the 90th Annual Meeting of the Transportation Research Board*, Washington D.C., USA.

6. *Cache Metropolitan Planning Organization (CMPO) Regional Transportation Plan 2035*, 2011 <http://cachempo.org/wordpress/wp-content/uploads/2012/04/CMPO-2035-RTP-adopted-June-20_2011-pub-format.pdf.> (March 20, 2014).
7. Chen, A., Chootinan, P., and Recker, W. (2005). "Examining the Quality of Synthetic Origin–Destination Trip Table Estimated by Path Flow Estimator." *Journal of Transportation Engineering* 131(7), 506-513.
8. Chen, A., Chootinan, P., and Recker, W. (2009). "Norm approximation method for handling traffic count inconsistencies in path flow estimator." *Transportation Research Part B*, 43(8), 852-872.
9. Chen, A., Ryu, S., and Chootinan, P. (2010). " L∞ -Norm Path Flow Estimator for Handling Traffic Count Inconsistencies: Formulation and Solution Algorithm." *Journal of Transportation Engineering*, 136(6), 565-575.
10. Chootinan, P., Chen, A., and Recke, W. (2005). "Improved path flow estimator for estimating origin-destination trip tables." *Transportation Research Record 1923*, 9-17.
11. Dill, J., and Carr, T. (2003). "Bicycle commuting and facilities in major US cities: if you build them, commuters will use them." *Transportation Research Record,* 1828, 311-316.
12. Ehrgott, M., Wang, J., Raith, A., and van Houtte, A. (2012). "A bi-objective cyclist route choice model." *Transportation Research Part A,* 46(4), 652-663.
13. Eom, J., Stone, J., and Ghosh, S. (2009). "Daily activity patterns of university students.", *Journal of Urban Planning and Development*, *135*(4), 141-149.
14. Evans, S. P. (1976). "Derivation and analysis of some models for combining trip distribution and assignment." *Transportation Research*, 10(1), 37-57.
15. *Highway Capacity Manual*. (2011). Transportation Research Board, Washington D.C., 2011.
16. Hopkinson, P. and Wardman, M. (1996). "Evaluating the demand for cycle facilities." *Transportation Policy*, 3(4), 241-249.
17. Hunt, J. D. and Abraham, J.E. (2007). "Influences on bicycle use." *Transportation*, 34(4), 453-470.
18. Kang, L., and Fricker, J. (2013). "A bicycle route choice model that incorporate distance and perceived risk. *Proc., Presented at the 92th Annual Meeting of the Transportation Research Board*, Washington D.C., USA.

19. Klobucar, M., and Fricker, J. (2007). "Network evaluation tool to improve real and perceived bicycle Safety." *Transportation Research Record 2031*, 25-33.
20. Krizel, G., and Thompson, K. (2009). "Analyzing the effect of bicycle facilities on commute mode share over time." *J. Urban Plann. Dev.*, 10.1061/(ASCE)0733-9488(2009)135:2(66), 66-73.
21. Kuzmyak, J., Walters, J., Bradley, M., and Kockelman, K. (2014) *Estimating Bicycling and Walking for Planning and Project Development: A Guidebook. National Cooperative Highway Research Program*, Report 770
22. Lundgren, J. and Patriksson, M (1998). "The combined distribution and stochastic assignment problem." *Annals of Operations Research*, 82, 309-330.
23. McKenzie, B., (2012) *Modes Less Traveled Bicycling and Walking to Work in the United States: 2008–2012. American Community Survey Reports*, ACS-25, 2014.
24. Mekuria, M., Furth, P., and Nixon, H. (2012). *Low-stress bicycling and network connectivity*. Mineta Transportation Institute, San José State University.
25. Quattrone, A., Vitetta, A. (2011). "Random and fuzzy utility models for road route choice." *Transportation Research Part E*, 47(6), 1126-1139.
26. Ryu, S., Chen, A., Zhang, H.M., and Recker, W. (2014). "Path flow estimator for planning application in small communities." *Transportation Research Part A*, 69, 212-242.
27. Ryu, S., Chen, A., Su, J., and Choi, K. (2017). Two-stage bicycle traffic assignment model. Journal of Transportation Engineering, Part A: Systems, 144(2), DOI: 10.1061/JTEPBS.0000108.
28. Ryu, S., Su, J., and Chen, A. (2015). A bicycle network analysis tool for planning applications in small communities. Utah Transportation Center Report UTC-1501.
29. Smith, T.E. (1978). "A cost-efficiency principle of spatial interaction behavior." *Regional Science and Urban Economics*, 8(4), 313-337.
30. Smith, T.E. (1983). "A cost-efficiency approach to the analysis of congested spatial-interaction behavior." *Environment and Planning A*, 15(4), 435-464.
31. Stinson, M. A. and Bhat, C.R. (2003). "Commuter bicyclist route choice: Analysis using a stated preference survey." *Transportation Research Record*, 1828, 107-115.
32. Wardrop, J.G. (1952). "Some theoretical aspects of road traffic research." *Proceedings of the Institution of Civil Engineers*, 2(1), 325-378.