Utilizing Public Domain Data to Estimate Non-motorized Trips Monthly: A Case Study in Washington Metropolitan Area

Yixuan PAN[1] and Lei ZHANG[2*] and Bo PENG[3*] and Xiaowei XU[4*]

[1]Graduate Research Assistant, Department of Civil and Environmental Engineering, University of Maryland, 1173 Glenn Martin Hall, College Park, MD 20742; email: ypan1003@umd.edu
[2]Herbert Rabin Distinguished Professor, National Transportation Center, Department of Civil and Environmental Engineering, University of Maryland, 1173 Glenn Martin Hall, College Park, MD 20742; email: lei@umd.edu
[3]Ph.D. Candidate, Department of Civil and Environmental Engineering University of Maryland, 1173 Glenn Martin Hall, College Park, MD 20742; email: bp105@terpmail.umd.edu

ABSTRACT

Non-motorized travel modes—mainly referred to as biking and walking—have drawn research attention for years. However, people's knowledge of the non-motorized travel demand is still interrupted and limited. Existing studies either looked into the microscopic behaviors of bicyclists and pedestrians with high temporal resolution, or focused on the macroscopic behaviors with lower temporal resolution (e.g., annual). However, the combination of macroscopic spatial scale and relatively high temporal resolution have yet to be examined for the non-motorized modes. Hence, the paper proposes a two-module framework to estimate the number of biking and walking trips monthly at the metropolitan level. Various public domain data sources are utilized, including the American Community Survey (ACS), non-motorized count data, and regional household travel survey. They help to first estimate the number of annual non-motorized trips in the study area, which is later disaggregated by the monthly trend factors derived from the Poisson multilevel model (PMM). The application in the Washington–Arlington–Alexandria, DC–VA–MD–WV metropolitan statistical area (D.C. MSA) demonstrates the feasibility and reliability of the proposed method.

Keywords: Non-motorized, Bicycle, Pedestrian, Poisson Multilevel Model

1. INTRODUCTION

Non-motorized transportation, also known as active transportation and human powered transportation, has tremendous benefits to not only non-motorized travelers, but also to the whole network. Non-motorized modes, including walking, biking, and other variants, are considered as a green and sustainable way to travel, enhancing riders' health while reducing pollution and congestion [1]. The public's discovery of those advantages has stimulated the non-motorized travel demand, which at the same time prompts the need for monitoring the temporal trend of non-motorized trips in order to optimize the transportation planning and traffic management.

However, the data on non-motorized transportation is not as abundant or standardized as other modes, and thus is not intuitive for producing continuous statistics. The most popular type of non-motorized data may be count data, but they are usually limited to a few locations and time periods. As a result, current studies mostly work on the microscopic travel behavior of non-motorized modes and provide trip summary of certain facilities or at certain times. Another data source is travel surveys, which provides more travel information at individual level.

Not too many papers have worked on the spatiotemporal analysis of non-motorized modes in the literature. Phung and Rose made use of automatic count data in Melbourne to study bike path usage, and further developed a regression model for each trail at both hourly level and monthly level [2]. Lewin analyzed temporal patterns of available bicycle counts in weekly, monthly and seasonal levels using the five-year continuous detector data from two permanent bicycle count locations (representing four stations). Weather influence (high temperature) was later incorporated to develop a regression model for estimating bicycle volume [3]. Nordback tried to estimate the annual average daily bicyclists and pedestrians (AADBP) utilizing both the short-term and the continuous count data for bicycle trips from the count program of Colorado Department of Transportation (CDOT). They also explored multiple statistical methods concerning the hourly weather information [4]. Nordback and Sellinger outlined a sample-based method to calculate bicycle and pedestrian miles traveled (BMT and PMT) for the state of Washington. They derived the seasonal, daily and hourly adjustment factors from the continuous count data and further applied them to the short-term counts conducted in the sample locations, each from a roadway functional group [5]. Gosse and Clarens addressed the temporal factoring of sparse bicycle counts through Markov Chain Monte Carlo (MCMC) sampling [6]. Lindsey et al. utilized the American Community Survey (ACS) and travel survey to estimate the annual trip total for the non-motorized modes [7].

It can be concluded that the previous studies are usually limited in the study scope or temporal resolution. There lacks a comprehensive technique that is able to ascertain the macroscopic volume and temporal trend of non-motorized transportation. This paper will integrate traditional count data and public survey results to provide such a method to estimate the monthly non-motorized trips at the metropolitan level. In contrast to the existing strategy, the proposed method features an inverse process where the annual trip total is first estimated and later apportioned to the targeted resolution.

2. DATA AND METHODOLOGY

This paper targets at establishing a handy and adaptable method to monitor the regional trip total and travel trend for non-motorized modes (Figure 1). The data sources considered can be easily accessible in most areas of the United States, including the American Community Survey (ACS), regional household travel survey, and non-motorized count data. One of the data products (Table B08006: Sex of workers by means of transportation to work) from the ACS publishes the average daily number of workers using different travel modes to work every year at county level [8]. Based on the average weekday estimates of non-motorized commuters, the regional household travel survey helps to compute the specialized trip rates and further generate the weekday estimates of non-motorized trips. The underlying assumption is that the share of working trips does not change over time. Meanwhile, the count data can provide temporal adjustment factors to disaggregate the annual trip total. In addition, the trip number reported by the bike-sharing program can be used to validate the monthly trend if applicable.

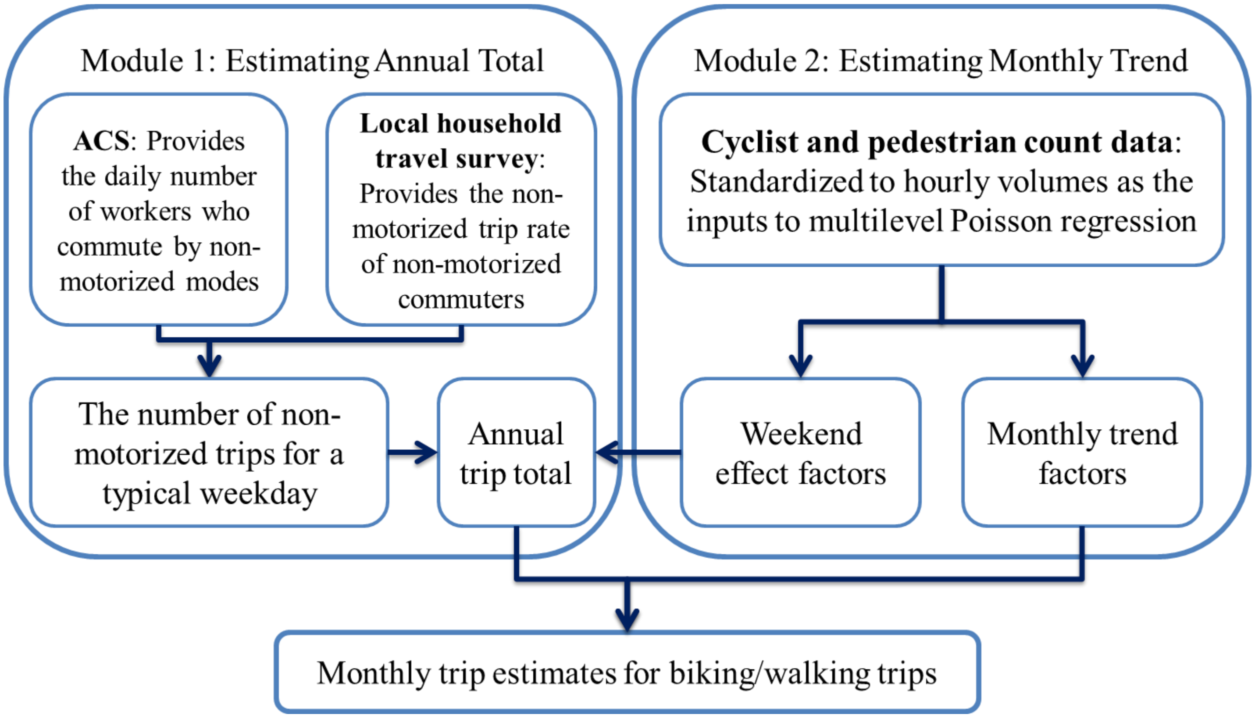

FIGURE 1. The framework for monthly non-motorized trip estimation

The aforementioned specialized trip rate for biking/walking commuters is derived from the raw data of the RHTS in Equation (1). After that, the specialized trip rate from RHTS performs as input to compute the average weekday biking/walking trips, together with the number of biking/walking commuters from ACS in Equation (2).

$$STR_{Biking/Walking} = N^{RHTS}_{Biking/WalkingTrips} \Big/ N^{RHTS}_{Biking/WalkingCommuters} \tag{1}$$

$$AWT_{\widehat{Biking/Walking}} = \sum_i N^{ACS}_{Biking/WalkingCommuters} \times STR_{Biking/Walking} \tag{2}$$

Where $N^{RHTS}_{Biking/WalkingTrips}$ is the weighted number of biking/walking trips reported to RHTS; $N^{RHTS}_{Biking/WalkingCommuters}$ is the weighted number of biking/walking commuters, who are identified by examining whether they bicycle or walk to work; $STR_{Biking/Walking}$ is the specialized trip rate for one single mode, either biking or walking; $N^{ACS}_{Biking/WalkingCommuters}$ is the daily number of biking/walking commuters estimated by ACS; $AWT_{\widehat{Biking/Walking}}$ is the average weekday biking/walking trip estimate.

To estimate the temporal trend, the Poisson Multilevel Model (PMM) is established with two parts: a Poisson regression model (3)–(4) and a multilevel model with random intercepts (5)–(6).

$$P(y|x) = \frac{\lambda^y e^{-\lambda}}{y!} \tag{3}$$

$$\log\big(E(Y|x)\big) = \log(\lambda) = \theta x \tag{4}$$

where y is the hourly bicycle/pedestrian volume, x is a set of the independent variables, and λ is the parameter of the predicted Poisson distribution. θ represents the coefficients to be estimated.

Level 1: $$\log(\lambda_{ij}) = \beta_{0j} + \beta_1 M_{ij} + \beta_2 Y_{ij} + \beta_3 WD_{ij} + \cdots \quad (5)$$
Level 2: $$\beta_{0j} = \beta_0 + u_{0j} \quad (6)$$

where i is the index of each data entry in the dataset and j is the identification of each count station; M_{ij} is a vector of indicators for months, where the indicator for the first available month is the reference category; Y_{ij} is a vector of indicators for different years, where the indicator for the first available year is the reference category; WD_{ij} is a binary variable indicating if the counting date is a weekend; β_{0j} is the random intercept composed of the fixed intercept β_0 and the random error term u_{0j}, whose distribution is assumed to be normal; β_1, β_2, ... are the fixed effects of the attributes to be estimated.

3. CASE STUDY

This section introduces the application of the proposed model in the Washington–Arlington–Alexandria, DC–VA–MD–WV metropolitan statistical area (shortened as D.C. MSA). The study year is 2016, when the latest ACS data was released. The regional household travel survey is the 2007/2008 TPB (Transportation Planning Board) Household Travel Survey [9]. Though there is a relatively large time gap, it is the best knowledge available about people's non-motorized travel behaviors in the D.C. MSA. The bicycle and pedestrian count data can be inquired via CommuterPage's web services [10]. It is observed that the counters are mainly located in the center of the MSA (Figure 2), which implies the assumption that the monthly trend and weekend effect are identical between the center urban areas and the outer rural areas.

The trip file of 2007/2008 TPB Household Travel Survey provides each trip record of every respondent. Following the steps in Section 3, 115 biking commuters are recognized out of 228 cyclists and 362 walking commuters out of 4,079 pedestrians. The weighted biking trip rate is 5.002 trips per biking commuter and the walking trip rate is 30.557 trips per walking commuter. Applying the specialized trip rate to the average weekday biking/walking commuter estimates from the ACS, the average weekday trip total can be achieved: 130,327 biking trips and 3,435,279 walking trips.

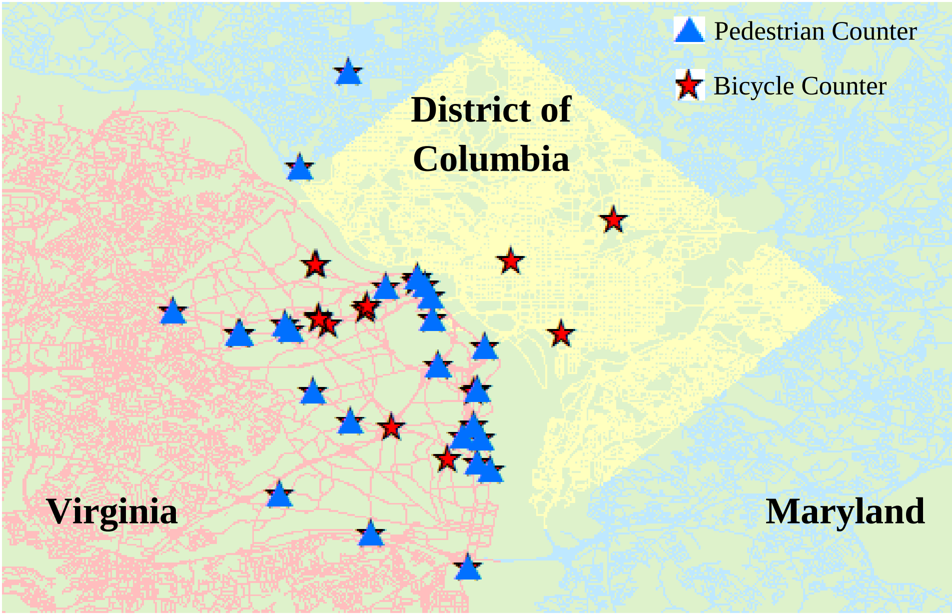

FIGURE 2. The spatial distribution of non-motorized automatic counters

The Poisson multilevel model then estimates the monthly trend and weekend effect factors based on the 2016 DC–MD–VA non-motorized automatic count data. The weekend effect factor is 0.55 for bicycle traffic and 1.07 for pedestrian traffic. In 2016, the annual linked biking trip total is 44,643,195 and the walking trip total is 1,259,150,276. Bike-sharing monthly trips from Capital Bikeshare are also utilized as a comparison to the monthly trend estimates (Figure 3). The trip comparison shows that the estimated monthly trend is very similar to that reported by the local bikeshare program, which to some degree demonstrates the reasonability of the PMM estimates.

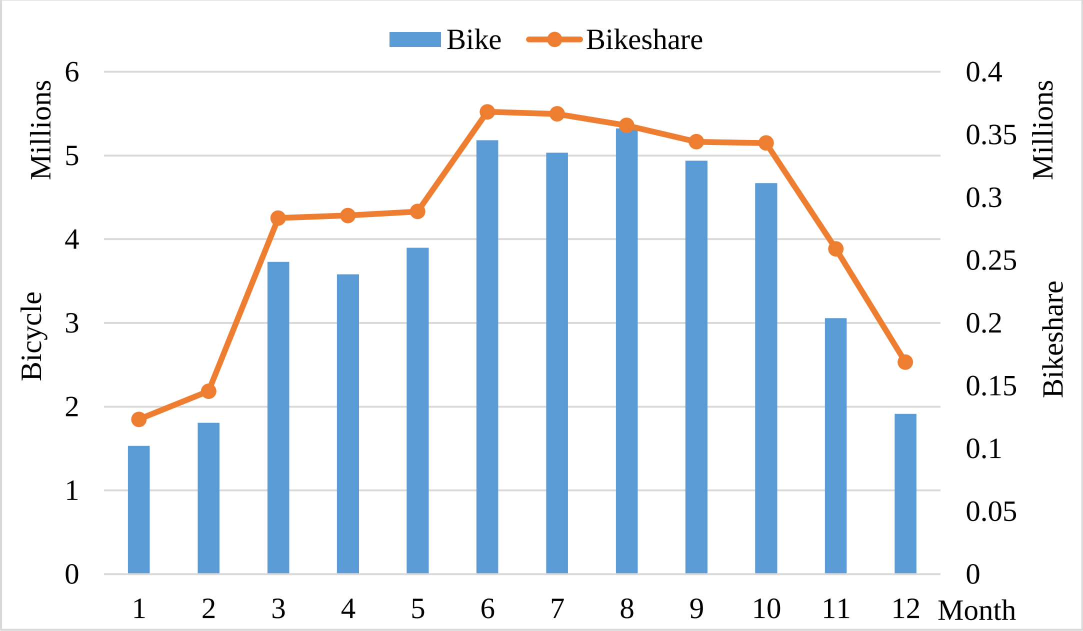

FIGURE 3. 2016 Monthly bicycle trip estimates and bike-sharing trips in D.C. MSA

The monthly trip esimtates for biking and walking mode are displayed in Figure 4. It can be noted that the monthly trend of walking trips is slightly smoother than that of biking trips except for the wintertime. Both modes indicate a dramatic demand increase in March and a minor decrease in July, which implies the weather influence on non-motorized travel.

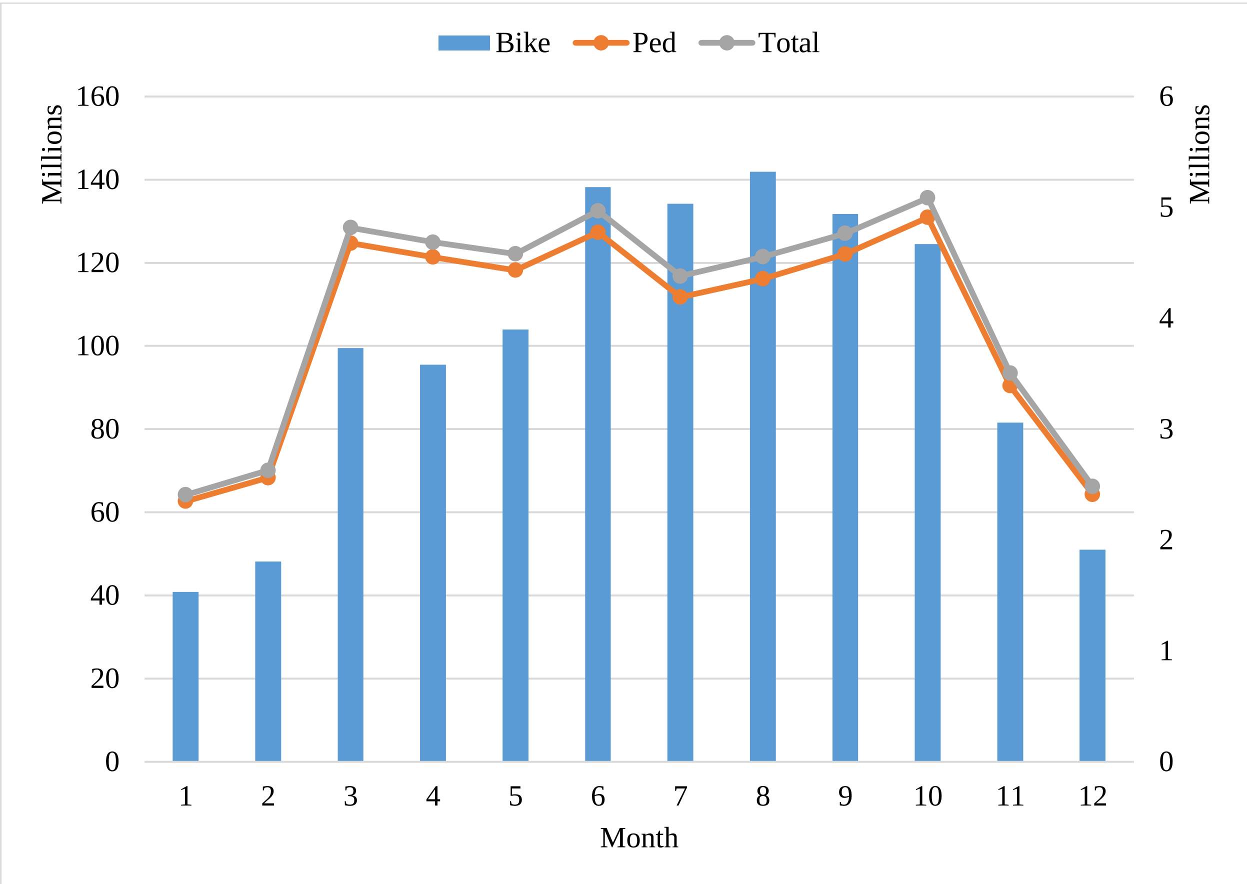

FIGURE 4. 2016 Monthly non-motorized trips in D.C. MSA

4. CONCLUSION

This paper has illustrated a novel methodology to estimate the monthly non-motorized trips at the metropolitan level. The case study has proved its feasibility as long as the three data sources are available. The contributions of this paper are two-fold: 1) the integration of multiple public domain data sources to enable the monthly demand estimation for a metropolitan area, and 2) a Poisson multilevel model to extract the average temporal trends considering the sparsity and scarcity of non-motorized count data.

The proposed static framework can be further extended to a dynamic and up-to-date data product even if the ACS estimates are lagged. Instead of using the month variables in a calendar year, count data of any consecutive months can serve as input to the multilevel model to generate the rolling monthly trend. Thus, the monthly trip total can be estimated based on the most recent number of trips in a known month and utilizing the rolling monthly trend.

5. ACKNOWLEDGEMENT

This research is financially supported by a Federal Highway Administration (FHWA) award, "Advancing Travel Behavior Information Gathering through Public Domain Data," FHWA BAA DTFH6115R00038. The opinions in this paper do not necessarily reflect the official views of FHWA. They assume no liability for the content or use of this paper. The authors are solely responsible for all statements in this paper.

REFERENCE

[1] Litman, T. (2017). Evaluating active transport benefits and costs: guide to valuing walking and cycling improvements and encouragement programs. Victoria Transport Policy Institute. Retrieved from http://www.vtpi.org/nmt-tdm.pdf.

[2] Phung, J., & Rose, G. (2007, September). Temporal variations in usage of Melbourne's bike paths. In Proceedings of *30th Australasian Transport Research Forum*, Melbourne.

[3] Lewin, A. (2011). Temporal and weather impacts on bicycle volumes. In *Transportation Research Board 90th Annual Meeting* (No. 11-2536).

[4] Nordback, K., Marshall, W. E., & Janson, B. N. (2013). Development of Estimation Methodology for Bicycle and Pedestrian Volumes Based on Existing Counts (No. CDOT-2013-18).

[5] Nordback, K., & Sellinger, M. (2014). Methods for estimating bicycling and walking in Washington state (No. WA-RD 828.1).

[6] Gosse, C., & Clarens, A. (2014). Estimating Spatially and Temporally Continuous Bicycle Volumes by Using Sparse Data. *Transportation Research Record: Journal of the Transportation Research Board*, (2443), 115-122.

[7] Lindsey, G., Quian, X., Linscheid, N., Tuck, B., Schoner, J., Pereira, M., & Berger, A. (2016). Assessing the Economic Impact and Health Benefits of Bicycling in Minnesota.

[8] U.S. Census Bureau. American Community Survey. 2016 American Community Survey 1-Year/5-Year Estimates, Table B08006; retrieved from https://factfinder.census.gov.

[9] Transportation Planning Board at the Metropolitan Washington Council of Governments. 2007/2008 Household Travel Survey: http://www1.mwcog.org/transportation/activities/hts.

[10] CommuterPage's web services: https://www.commuterpage.com/tools-resources/tools-for-developers.