Use of Direct-Demand Modeling in Estimating Nonmotorized Activity: A Meta-analysis

Project Number: UTC Safe-D 01-003
Project Title: Data Mining to Improve Planning for Pedestrian and Bicyclist Safety

Sirajum Munira, Texas A&M Transportation Institute
Ipek N. Sener, Ph.D., Texas A&M Transportation Institute

October 2017

Disclaimer

Acknowledgments

This research was funded by a grant from the Safety through Disruption (Safe-D) UTC supported by the U.S. Department of Transportation. The research described in this report was developed as part of a collaborative project between San Diego State University, the Texas A&M Transportation Institute, and the Virginia Tech Transportation Institute. The authors would like to thank Arash Jahangiri from San Diego State University, Justin Owens from the Virginia Tech Transportation Institute, and Shawn Turner from the Texas A&M Transportation Institute for their comments and feedback on a draft version of this report. The authors would also like to acknowledge Michelle Benoit from the Texas A&M Transportation Institute for her editorial review of the report.

Abstract

Growing interest in accurately estimating nonmotorized activities has generated a large number of studies over the past several years. The use of the direct-demand modeling approach to predict pedestrian and bicyclist volume is gaining attention due to its simplicity in application and the recent availability of large-scale datasets. While the objectives of the studies are similar, researchers and transport planners have used a myriad of approaches based on the magnitude of the available data and the characteristics of the study area. Preference for explanatory variables often varies by location and time period. Also, different modeling approaches have different strengths and weaknesses. This report summarizes the challenges and opportunities reflected in the literature associated with the use of direct-demand models, and compares the commonalities and site-specific approach differences. This report also discusses the transferability of the models under certain conditions. Significant explanatory variables and their interaction with nonmotorized volume across studies are summarized. The relevant findings may be important to decision makers, and observations about the variations in the different approaches may inform future studies.

Table of Contents

List of Tables

Executive Summary

Walking and bicycling make up a relatively small portion of transportation in the United States and yet account for a disproportionate share of the total fatal and serious injury crashes. Unfortunately, the lack of availability of nonmotorized exposure data often makes it difficult to discern a trend in crash rates and to identify high-risk locations. Transportation researchers and several federal agencies have highlighted the need for good exposure data to measure pedestrian and bicyclist risks, acknowledging that nonmotorized volume/exposure is one of the least understood areas for transportation planners. The counts for pedestrian and bicycle traffic are the key input for exposure analysis. Counts can either directly measure exposure on specific facilities or help agencies develop and calibrate network and regional models that evaluate exposure at various geographic scales. Among various modeling-based approaches for exposure analysis, the direct (facility) demand model is the most frequently used approach in the area of pedestrian/bicyclist safety.

The direct-demand model relates walking and bicycling demand directly to various associating factors such as sociodemographic factors and land use characteristics. Studies dating back 50 years have used this modeling approach to forecast nonmotorized traffic, and it has been widely used in different areas of transportation. Recently, this approach has attracted attention because it benefits from the availability of a large volume of good-quality data and spatial database management software such as geographic information systems, and results in comparatively simple tools that enable transport planners to predict nonmotorized traffic at relevant locations where count data are not available.

This report reviewed various studies that used the direct-demand modeling approach to estimate pedestrian, bicycle, and trail traffic volume at different locations. While the objectives of the studies are similar, researchers and transport planners have used a myriad of approaches based on the magnitude of the available data and the characteristics of the study area. This report summarizes the challenges and opportunities reflected in the literature associated with the use of the direct-demand models and compares the commonalities and differences in the site-specific approaches.

The generalized approach to develop a direct-demand model includes selection of a wide array of independent variables, often at various spatial scales, and choice of a suitable analysis method to estimate pedestrian, bicycle, or trail traffic in an area or location. Typically, the dependent variables of direct-demand models are pedestrian, bicycle, or trail traffic volumes for various time periods, such as during the peak period, hourly, daily, or annually. While some research has directly used data for the specific collection period, other studies have expanded short-period data to longer periods by using a scaling factor to be integrated and used in models. Studies have also explored a wide array of independent variables. Preference for explanatory variables often varies by location and time period. To review the independent variables used in these studies, this report categorizes the variables into nine groups: demographic, socioeconomic, network/interaction with vehicle traffic, pedestrian- or bicycle-specific infrastructure, transit facilities, major generators, weather and environmental, temporal or time related, and land use factors. Studies have highlighted that walking and bicycling trip behaviors differ substantially and need to be investigated separately. To identify the impact of land-use and built-environment characteristics on nonmotorized volume, a number of studies have considered a range of buffer widths. Investigating the influence of various independent variables by different buffer widths, studies have suggested that the best model may be

obtained using different scales of buffer zones for different variables because the variables are unlikely to be significant at the same buffer scale. To explore whether the variables have a consistent impact on nonmotorized volume across the studies, this report identifies the positive or negative impact of each of the significant variables on pedestrian, bicycle, and trail traffic across studies. It was observed that often the impact of the same independent variable varies across studies, most likely because of the land use characteristics of the study location. A wide variety of approaches and methods (e.g. ordinary least squares, negative binomial, and Poisson models) have been used in predicting nonmotorized activity using direct-demand models. Through the exploration of a wide array of modeling techniques, the majority of the studies have acknowledged the suitability of negative binomial models in predicting pedestrian and bicyclist volume.

The direct-demand modeling approach has some advantages over some other modeling approaches. The major advantage of the direct-demand modeling approach is that it can be developed largely using existing data and common software packages. Because the model explains the impact of different factors that influence people's travel choice, it can provide important contributions to the decision-making process. For example, a positive association between major streets and nonmotorized volume may promote initiatives such as Complete Streets. The significance of independent variables such as bus stops and off-street trails, at a small spatial scale, indicates that targeted improvements in infrastructure could encourage nonmotorized activities. However, the model has some limitations, especially when transferred far into the future and for large areas, and researchers and practitioners need to be judicious in developing and applying these models.

Although significant progress can be observed in the use of direct-demand modeling to estimate nonmotorized activity, challenges in data collection and model interpretation are yet to be fully resolved. The insights and findings in this report are intended to inform future research and transport-related policy making.

Introduction

Background

Recognizing the benefits of nonmotorized activity in physical fitness and sustainable transportation development, health advocates and transport planners have been steadfast in promoting walking and biking over the past several years. With the increasing development of pedestrian- and bicyclist-friendly infrastructures and design improvements, and perhaps the help of campaigns encouraging walking and bicycling, several U.S. cities are seeing a rise in nonmotorized activity. The American Community Survey reported that from 2000 to 2012, bicycle trips had the largest increase of any commuting mode (McKenzie, 2014). Although it varies widely by location, walking also appears to be a popular mode among young workers and students in some cities (McKenzie, 2014). However, walking and bicycling make up a relatively small portion of commuting activity in the United States (Centers for Disease Control and Prevention, 2016).

Despite their relatively low share of overall traffic, pedestrians and bicyclists account for a disproportionate share of the total fatal and serious injury crashes in the United States. In 2015, 5,376 pedestrians were killed and around 70,000 pedestrians were injured (National Highway Traffic Safety Administration, 2016). That same year, over 1,000 bicyclist deaths and almost 467,000 bicycle-related injuries were reported (Centers for Disease Control and Prevention, 2017). These two modes accounted for around 18 percent of the total U.S. traffic fatalities that year (National Highway Traffic Safety Administration, 2016). Unfortunately, the lack of exposure data for nonmotorized vehicles often makes it difficult to discern a trend in crash rates and identify high-risk locations (Turner et al., 2017). A National Highway Traffic Safety Administration and Federal Highway Administration (FHWA) report describes the need for good exposure data to measure pedestrian and bicyclist risks, stating that exposure for nonmotorized travel is one of the least understood areas for transportation planners (Hedlund, 2000). In addition to safety analysis, planners and decision makers need a reliable estimate of the current and future nonmotorized activity demand to plan and manage resources. Due to a lack of reliable volume estimates, bicycle development projects are often at a disadvantage when competing with projects for motorized vehicles despite their enormous potential to develop sustainable communities (Gosse and Clarens, 2014).

The question of how many people actually use current pedestrian or bicycle facilities, or how many will use a new or improved facility, can be answered by demand forecasting models that predict future levels of pedestrian and bicycle travel. The compelling benefits of nonmotorized demand forecasting modeling include better planning of infrastructure requirements, prioritization of projects based on benefits, and assessment of the safety of nonmotorized modes by developing exposure information for crash/safety models (Schwartz et al., 1999; Molino et al., 2009; Schneider et al., 2009a; Schepers, 2012; Aoun et al., 2015). Demand forecasting approaches can answer policy makers' most frequently asked questions (Porter et al., 1999), such as:

- How many people will use a new facility?
- Will a new facility increase demand?
- How will the facility affect overall mobility, traffic congestion, or air quality conditions?

Furthermore, some modeling approaches are particularly useful for isolating and quantifying the influence of specific factors on nonmotorized travel behavior and for assessing the interaction among the factors (Bhat et al., 2005). A comprehensive understanding of the influencing factors is also beneficial for long-term decision making. For example, the influence of precipitation on nonmotorized activity reveals people's reluctance to walk or bicycle in the rain, thus showing the need for measures to ensure pedestrian and bicyclist safety during adverse weather conditions (Lindsey, 2011).

The demand forecasting approaches vary widely in the quantity, specificity, and scale of data used. Data used for demand forecasting models can range from readily available census data to continuous cell phone or global positioning system (GPS) tracking data (Aoun et al., 2015). Although the methods of collecting motorized traffic data are well established and practiced by several transportation agencies, only a few states and municipalities have a formal approach for counting nonmotorized traffic (Fagnant and Kockelman, 2016). Often, transportation agencies quantify nonmotorized activity through manual short counts performed by staff and volunteers at sites. The amount of data collected by this method is limited and often not representative of the actual demand pattern over the course of days, months, or seasons (City of Greensboro, North Carolina, 2015). Moreover, systematic effort is required to gather and analyze data to understand the prevailing nonmotorized traffic patterns (Handy and McCann, 2010). Hunter and Huang (1995) looked at pedestrian and bicyclist counts in different locations and concluded that even within the same facility, pedestrian and bicycle activity varies widely with location and time. The authors also concluded that the short-term counts in many studies were not representative enough to be generalized and used in modeling and decision making. While a number of early studies have used short counts to estimate average daily or annual volume, continuous count data are necessary to account for daily, weekly, and seasonal variations in nonmotorized activity (Schneider et al., 2009b). Recently, several automated technologies such as pneumatic tubes, inductive loops, thermal cameras, infrared sensors, magnetometers, piezoelectric devices, radar sensors, and video imaging have been used to increase the accuracy of walking and bicycling data (Levinson et al., 2016; Nordback et al., 2016). A number of cities have established a robust program of continuous, automated pedestrian and bicycle counters to observe nonmotorized activity data. However, the count observations cannot be considered immediately policy relevant until they are scaled to a long-term representative value (Gosse and Clarens, 2014). Moreover, it is not cost effective to install automatic sensors at every location in a city to monitor continuous pedestrian and bicycle volume data (El Esawey et al., 2015).

The Current Study in Context

The counts for pedestrian and bicycle traffic are the key input for exposure analysis at a specific scale. As part of a recent FHWA project, Turner et al. (2017) conducted a comprehensive review of the methods for estimating exposure for nonmotorized traffic and indicated that the counts can either directly measure exposure on specific facilities, or help agencies develop and calibrate network and regional models that evaluate exposure at various geographic scales. Acknowledging the limitations of exposure analysis based solely on direct measurement, the authors identified several modeling-based methodologies for exposure analysis. Among these, Turner et al. (2017) highlighted that the direct (facility) demand model is the most frequently used modeling approach in the area of pedestrian/bicyclist safety. The other approaches include regional travel demand models (e.g., trip-based models and activity-based models), trip generation and flow models (e.g., pedestrian trip generation and flow models, and network simulation models), geographical information system (GIS)–based models (e.g., walk accessibility models), simulation-based traffic models (e.g., dynamic traffic assignment models), and other special focused models such as bicycle route choice models. While some of these methodologies are widely used by practitioners and researchers, some have been used infrequently or not at all; however, Turner et al. (2017) discussed different methods by explaining their potential in estimating exposure for nonmotorized travel. In addition, while some of these models provide direct volume exposure estimate (e.g., direct-demand models or regional travel demand models), the others (e.g., bicycle route choice models) need to be integrated into various tools to obtain the exposure estimate.

Emerging technologies are also paving the way for refining data collection methods to accurately assess demand in an easy and cost-effective way (Anda et al., 2017). GPS-enabled smartphones coupled with applications can track road users and their travel activities. App-based counting software such as Counterpoint; GPS-enabled route trackers such as Strava Metro, Cycletracks, and Cycle Atlanta; and GPS-enabled route trackers with hardware (magnetometers and infrared cameras) such as Ride Report are being used to collect pedestrian and bicyclist trip data, which public agencies then purchase and analyze for planning purposes (O'Toole and Piper, 2017). The availability of large-scale comprehensive datasets could provide advantages to several demand forecasting approaches. Studies indicate that the availability of a large volume of good-quality data is expected to be beneficial for two general classes of models: choice-based regional transportation models and direct- (or facility-) demand models (Kuzmyak et al., 2014), the latter being the focus of this report.

This report provides an extensive review of direct-demand modeling to estimate nonmotorized activity. Direct-demand models have been acknowledged as a useful tool for generating spatial estimates of nonmotorized activity and producing generalized results about the influence of specific built-environment features on pedestrian and bicycle traffic. The models are comparatively simple tools that enable transport planners to predict nonmotorized traffic at relevant locations where count data are not available.

The review presented here is based on both published and unpublished articles, papers, and reports. The majority of the studies were conducted in the United States, but examples are also included from Canada and Europe.

The remainder of this report is organized into the following sections:

- "Direct-Demand Model Studies"—an overview of direct-demand models and an in-depth comparative analysis of studies and noteworthy findings.
- "Data Analysis"—a discussion of the data analysis procedures and the dependent and independent variables.
- "Relationship with Independent Variables"—an overview of the nature of interaction between the volume and explanatory factors.
- "Model Benefits and Limitations"—a summary of the advantages and limitations of direct-demand models.
- "Conclusion"—a summary of findings and conclusions.

Direct-Demand Model Studies

An Overview of Direct-Demand Models

The increasing interest in forecasting nonmotorized travel demand has given rise to prolific research activity over the past several years. The traditional regional demand forecasting approach is a four-step process:

1. Trip generation.
2. Trip distribution.
3. Mode share.
4. Traffic assignment (Schwartz et al., 1999).

Other modeling approaches include direct-demand models, GIS-based models, trip generation and flow models, network analysis models, discrete-choice models, simulation-based traffic models, and more (Turner et al., 2017). The available modeling approaches have widely varying functions, strengths, and limitations. One of the most frequently used modeling approaches to estimate nonmotorized volume is the direct-demand model, which predicts volume or flow by combining all the elements of trip generation, attraction, distribution between zones, and modal choice in a single model (Domencich and McFadden, 1974).

Developed in the 1960s, direct-demand models have been widely used in estimating demand in different areas of transportation, such as nonmotorized demand (Pushkarev and Zupan, 1971), intercity passenger travel demand (Kraft, 1963), inter-urban rail travel demand (Wardman, 1997), air travel demand (Wardman et al., 1994), inter-regional commodity flow (Ranaiefar et al., 2014), and regional road freight movement (Sjafruddin et al., 1999). The model is comparable to no-constraint gravity models (Ranaiefar et al., 2014) and relates travel demand directly to mode, trip, and traveler attributes using different forms of regression analysis (Ortuzar and Willumsen, 2011). The resulting models can be used to predict travel activity at similar locations without counts. The models are widely used due to their simplicity in understanding and application. The approach is convenient, especially when it is impractical to collect continuous data at all locations in a large community (Schneider et al., 2012).

The concept of using a direct-demand model to estimate nonmotorized activity is not new. Studies dating back 50 years have forecast nonmotorized traffic using count and spatial data. Pushkarev and Zupan (1971) collected pedestrian count data using aerial photography in Manhattan, New York, and carried out multiple-correlation analysis models to forecast volume based on surrounding land-use characteristics such as walkway space and building floor space type. Behnam and Patel (1977) developed stepwise regression models where the pedestrian volume per hour per block was forecast based on land-use variables that included commercial space, office space, cultural and entertainment space, manufacturing space, residential space, parking space, vacant space, and storage and maintenance space.

Despite these early efforts, practitioners in the field did not widely recognize and adopt the approach, in part because of the difficulties in assembling a large amount of spatial data. This problem can now be overcome because of the availability of data and spatial database management software such as GIS (Lindsey, 2011).

Review of Nonmotorized Direct-Demand Models

This section of the report reviews the studies that have used the direct-demand modeling approach to estimate pedestrian, bicycle, and trail traffic activity. The studies show several design differences. Data collection approaches range from short-term manual observation to continuous automated counting. The counts were collected on signalized and unsignalized intersections or at the midblock locations along the street segments. Choice of explanatory variables and statistical method varies widely across the studies. Few of the studies applied their models to predicting traffic in similar locations without counts. Although all of the researchers acknowledged the need for model validation, few of them validated their models. Compared to earlier studies, the more recent studies generally used more extensive datasets and exhibited improvements in model building and validation. Table 1 summarizes the commonalities and differences of several studies and discusses their coverage scale, application scale, methods, validation, explanatory variables, and limitations. The table is followed by a brief discussion of some unique approaches and noteworthy observations of some of the recent research.

A number of studies have developed multiple statistical models for comparison. Table 1 lists the significant dependent variables ($P < 0.05$) but only discusses models with the best prediction accuracy (according to the studies). Moreover, some studies have also generated separate models for specific times and locations in addition to an overall model, which is aggregate in nature. Due to space constraints, the significant explanatory variable column lists only the findings of the aggregated models. The comment section briefly discusses the limitations and future scope outlined by the studies.

Table 1. Direct-Demand Models to Estimate Nonmotorized Traffic Volumes.

Author (Date)	Coverage and Data Collection Scale	Count Type and Time	Application Scale	Analysis Methods	Model Validation	Initial Number of Variables and Buffer Sizes	Significant Explanatory Variable (Buffer Size)		Comment (Limitations and Future Scope If Outlined by the Study)
							Pedestrian	Bicycle	
Hankey et al. (2017)	Location: Blacksburg, VA Count type and coverage: pedestrian and bicyclist counts at 101 locations on different street and trail segments	Data collection time and type: continuous counts for 1 year (2015) at four sites and a 1-week count at 97 sites Count method: pneumatic tubes, passive infrared, and radio beam	All street and trail segments in Blacksburg	Stepwise linear regression	Validated by goodness of fit, internal validation, and a Monte Carlo–based 20% holdout analysis	18 variables at 11 buffer sizes	Sidewalk length; off-street trail length; household income; residential addresses count in buffer; population density; bus stop count in buffer	Household income; centrality[a]; population density; on-street facility length; major roads length	Represents a unique small college town population. Factor groups to develop scaling factors were not developed, which may result in bias in average annual daily traffic estimation.
Hankey and Lindsey (2016)	Location: Minneapolis, MN Count type and coverage: pedestrian and bicyclist counts at 471 locations on different street and trail segments	Data collection time and type: 2-hour (4 to 6 p.m.) counts from 2007 to 2014, weekdays in mid-September Count method: trained volunteer-based counts	Midpoint of each block for the entire transportation network (both streets and trails) in Minneapolis	Stepwise linear regression model	Internal validation and Monte Carlo–based 10% holdout analysis	16 variables at 12 buffer sizes	Major roads (200 m); off-street trails (3,000 m); transit stops (400 m); retail areas (100 m); industrial areas (1,250 m); open space areas (100 m); job accessibility[b]; population density (750 m)	Off-street trails (200 m); on-street facilities (100 m); retail areas (100 m); industrial areas (1,250 m); open space areas (200 m); job accessibility; population density (1,250 m); precipitation; temperature	Proposed future research to investigate how count programs developed for spatial modeling can decrease the potential for spatial autocorrelation and improve the performance of the facility-demand models.

Author (Date)	Coverage and Data Collection Scale	Count Type and Time	Application Scale	Analysis Methods	Model Validation	Initial Number of Variables and Buffer Sizes	Significant Explanatory Variable (Buffer Size)		Comment (Limitations and Future Scope If Outlined by the Study)
							Pedestrian	Bicycle	
Fagnant and Kockelman (2016)	Location: Seattle, WA Count type and coverage: bicycle counts at 251 intersections	Data collection time and type: Tuesdays through Thursdays, 6 a.m. to 9 a.m. or 3 p.m. to 6 p.m. Count method: manual observation	Community of Shoreline, WA	Negative binomial and Poisson models	Not reported	23 variables	Not reported	Employment density; bicycle-trail access; bridges; number of lanes; curb-lane width; bike-lane width; separated paths; speed limit; residential areas; morning period count; League of American Bicyclists (LAB) gold[c]	The model may fall short of a comprehensive count program.
Tabeshian and Kattan (2014)	Location: Calgary, Canada Count type and coverage: pedestrian and bicycle counts at 34 intersections located on major arterials (excluding downtown)	Data collection time and type: 7 to 9 a.m., 11 a.m. to 1 p.m., 4 p.m. to 6 p.m. in different months from 2007 to 2012 Count method: manual observation	Not reported	Multiple linear and Poisson models	Validation based on prediction models of 18 intersections in southwest Calgary	26 variables at four buffer sizes	Number of bus stops (0.1 mi); street length (0.5 mi); total bus-km of bus routes (0.75 mi); total number of dwell counts (0.5 mi); hectares of commercial space (0.25 mi); number of schools (0.5 mi); pathway length (0.25 mi)	Hectares of commercial space (0.10 mi); hectares of low-density residential space (0.10 mi); number of bus stops (0.25 mi); hectares of institutional space (0.50 mi); number of street lanes reaching an intersection	Small sample size. The model excludes the downtown area.

Author (Date)	Coverage and Data Collection Scale	Count Type and Time	Application Scale	Analysis Methods	Model Validation	Initial Number of Variables and Buffer Sizes	Significant Explanatory Variable (Buffer Size)		Comment (Limitations and Future Scope If Outlined by the Study)
							Pedestrian	Bicycle	
Strauss et al. (2013)	Location: the island of Montreal, Quebec, Canada Location type and size: bicycle activity counts at 647 signalized intersections	Data collection time and type: 8-hour counts between April and November 2009 (when seasonal bicycle facilities are open) Count method: manual observation and loop detector	Not reported	Bayesian modeling	Not reported	14 variables at three buffer sizes	Not reported	Number of employment (400 m); presence of schools (400 m); presence of subway stations (800 m); land-use mix (800 m); length of bicycle facilities (800 m); commercial land-use area (50 m); presence of three approaches	A larger sample of intersections is needed to validate the results and draw more solid conclusions.
Strauss and Miranda-Moreno (2013)	Location: the island of Montreal, Quebec, Canada Location type and size: bicycle activity counts at 758 intersections	Data collection time and type: 8-hour weekday counts during 2008 and 2009 Count method: manual observation and loop detector	Not reported	Log-linear and negative binomial models	Not reported	27 variables at four buffer sizes	Not reported	Number of employment (400 m); number of schools (400 m); presence of subway stations (150 m); number of bus stops; land-use mix (800 m); mean income (50 m); presence of a bicycle lane; presence of a cycle track; length of bicycle facilities (800 m); presence of parking entrance	Future study may use additional variables such as the location and proximity to public bike stations, and elevation and slope of the intersection and its surrounding area.

Author (Date)	Coverage and Data Collection Scale	Count Type and Time	Application Scale	Analysis Methods	Model Validation	Initial Number of Variables and Buffer Sizes	Significant Explanatory Variable (Buffer Size)		Comment (Limitations and Future Scope If Outlined by the Study)
							Pedestrian	Bicycle	
Wang et al. (2013)	Location: Minneapolis, MN Location type and size: trail traffic counts at six locations on multiuse trails	Data collection time and type: varying number of daily counts from June 2010 to September 2011 Count method: active infrared monitors	Not reported	Negative binomial model	Validation based on predicted trail traffic for each location for 1 week, not included in the dataset, with the actual count during that time	10 variables	Percent of African-American residents; percent with a college education; percent of population over 64 or below 6; median household income; population density; recorded high temperature; deviation from the 30-year normal temperature; precipitation; average wind speed; weekends		Counters cannot distinguish between pedestrians and bicyclists. Data are available for only six locations for unequal time periods.
Schneider et al. (2012)	Location: San Francisco, CA Location type and size: counts of pedestrians who crossed each leg of the 50 intersections	Data collection time, type, and count method: manual 2-hour counts at 28 intersections in September 2009 and July and August 2010; automated counters at 25 intersections in March and September 2010 (hourly counts for 3 to 4 weeks at each intersection)	Not reported	Log-linear model	Validated against 2002 pedestrian volume at other 49 four-way intersections	16 variables	Number of households (0.25 mi); total employment (0.25 mi); intersection is in a high-activity zone; maximum slope on any intersection approach leg; intersection is within 0.25 mi of a university campus; intersection is controlled by a traffic signal	Not reported	The models may not perform well for special attractors. Future studies may test additional variables.

Author (Date)	Coverage and Data Collection Scale	Count Type and Time	Application Scale	Analysis Methods	Model Validation	Initial Number of Variables and Buffer Sizes	Significant Explanatory Variable (Buffer Size)		Comment (Limitations and Future Scope If Outlined by the Study)
							Pedestrian	**Bicycle**	
Hankey et al. (2012)	Location: Minneapolis, MN Location type and size: pedestrian and bicyclist counts at 259 locations, midblock portion of each street or sidewalk segment	Data collection time and type: total 330 2-hour (4 p.m. to 6 p.m.) counts and 43 12-hour (6:30 a.m. to 6:30 p.m.) counts on weekdays in September during 2007–2010 Count method: manual observation	Estimate 12-hour non-motorized traffic counts for nearly all street segments in Minneapolis (12,481 streets)	Ordinary least squares (OLS) and negative binomial models	Validates models based on predicted non-motorized traffic at 85 locations (46 new and 39 previously sampled locations)	14 variables	Percent of non-White residents; percent of residents with a college education; distance from the central business district (CBD); distance from the nearest body of water; recorded precipitation; principal arterial street (of count location); arterial street (of count location); collector street (of count location)	Percent of non-White residents; percent of residents with a college education; median household income; measure of mixing of land uses; distance from the CBD; recorded precipitation; off-street trail (of count location); arterial street (of count location); year	Since counts are from September, results should be interpreted as typical traffic volumes during September only. More data are required to better explain between-location and within-location variability.

Author (Date)	Coverage and Data Collection Scale	Count Type and Time	Application Scale	Analysis Methods	Model Validation	Initial Number of Variables and Buffer Sizes	Significant Explanatory Variable (Buffer Size)		Comment (Limitations and Future Scope If Outlined by the Study)
							Pedestrian	Bicycle	
Lindsey et al. (2012)	Location: Minneapolis, MN Location type and size: pedestrian and bicyclist counts at 240 different roads and trail locations	Data collection time and type: mostly 2-hour, peak-hour counts and some 12-hour counts from 2007 to 2009, in September or October, and occasionally at different times throughout the year Count method: manual observation	Not reported	OLS and negative binomial models	Model based on 2007–2009 data was used to predict 2010 traffic to compare with the actual 2010 count value	20 variables	Percent of non-White residents; percent of residents with a college education; recorded precipitation; principal arterial street (of count location); arterial street (of count location); and collector street (of count location); distance from the CBD; distance from the nearest body of water	Percent of non-White residents; percent of residents with a college education; median household income; measure of mixing of land uses; distance from the CBD; recorded precipitation; off-street trail (of count location); arterial street (of count location); on-street bicycle facility (of count location); year	Pedestrian and bicyclist manual counts are not available for all months of the year, and the data count locations are not representative of the whole city.

Author (Date)	Coverage and Data Collection Scale	Count Type and Time	Application Scale	Analysis Methods	Model Validation	Initial Number of Variables and Buffer Sizes	Significant Explanatory Variable (Buffer Size)		Comment (Limitations and Future Scope If Outlined by the Study)
							Pedestrian	Bicycle	
Lindsey (2011)	Location: Minneapolis, MN Location type and size: pedestrian and bicyclist counts at 240 locations on various street types	Data collection time and type: total 352 2-hour and 43 12-hour observations between 2007 to 2010 during fall of each year Count method: manual observation	Models were used to predict 12-hour non-motorized traffic counts for all street segments (n = 12,481) in Minneapolis	OLS model	Not reported	25 variables	Land-use mix[d]; percent of Black residents; percent of college students; median household income; recorded high temperature; minor arterial road facility of count location and major collector road facility of count location	Land use mix; percent of Black residents; percent of college students; percent of residents younger than 5 and older than 65 years; median household income; recorded high temperature; deviation from daily high temperature; minor arterial with a bicycle facility (of count location); minor arterial without a bicycle facility; off-street trail facility (of count location)	The model is not representative of the full range of weather conditions throughout the year. Additional measures of income, education, population age distribution, access to employment, access to recreation and other destinations, and even street classification could be tested.

Author (Date)	Coverage and Data Collection Scale	Count Type and Time	Application Scale	Analysis Methods	Model Validation	Initial Number of Variables and Buffer Sizes	Significant Explanatory Variable (Buffer Size)		Comment (Limitations and Future Scope If Outlined by the Study)
							Pedestrian	Bicycle	
Miranda-Moreno and Fernandes (2011)	Location: Montreal, Canada Location type and size: pedestrian counts at 1,018 signalized intersections	Data collection time and type: 8-hour observation during the a.m. peak, noon, and p.m. peak periods, mostly taken in 2009 and some in 2008 Count method: manual observation	Not reported	Log-linear model	Validated based on 20% of the sample holdout analysis	25 variables at three buffer sizes	Population (400 m); commercial space area (50 m); open space area (150 m); presence of a subway station (150 m); number of bus stops (150 m); number of schools (400 m); percent of major roads (400 m); average street length (400 m); three- or four-way intersection; distance to downtown; very warm temperature (max. temperature > 32°C)	Not reported	The sample of intersections was not randomly selected. The parameter estimates may be biased because the models did not account for the potential spatial correlation across intersections.

Author (Date)	Coverage and Data Collection Scale	Count Type and Time	Application Scale	Analysis Methods	Model Validation	Initial Number of Variables and Buffer Sizes	Significant Explanatory Variable (Buffer Size)		Comment (Limitations and Future Scope If Outlined by the Study)
							Pedestrian	Bicycle	
Griswold et al. (2011)	Location: Alameda County, CA Location type and size: bicycle counts at 81 intersections along arterial and collector streets	Data collection time and type: two 2-hour period counts at each intersection, one taken on a weekday and one taken on a Saturday during different times of the day in spring 2008 and 2009 Count method: manual observation	Not reported	Log-linear OLS model	Not reported	15 variables at three buffer sizes	Not reported	Number of commercial properties (1/10 mi); presence of bicycle markings on any approach; network distance to University of California, Berkeley campus edge; average slope (degrees) of terrain (1/2 mi); year	The model can be refined further by collecting counts at additional locations and conducting model validation. Additional land-use, transportation system, and socioeconomic variables can be added.
Jones et al. (2010)	Location: San Diego, CA Location type and size: bicycle and pedestrian counts at 80 intersections	Data collection time and type: two (2007 and 2008) peak-period counts at 80 locations and 1 year (August 2007 to July 2008) of automated 24-hour counts at five locations Count method: manual and combination of passive infrared counters and active infrared counters	San Diego	Stepwise regression model	Not reported	34 variables at three buffer sizes	Employment density (0.25 mi); population density (0.25 mi); presence of retail (0.25 mi)	Footage of Class I (Multi-Use Path) bicycle path (0.25 mi); employment density (0.25 mi)	Additional variables could be considered, such as presence of parks, retail establishments, choke points, and other factors that may affect walking and bicycling.

Author (Date)	Coverage and Data Collection Scale	Count Type and Time	Application Scale	Analysis Methods	Model Validation	Initial Number of Variables and Buffer Sizes	Significant Explanatory Variable (Buffer Size)		Comment (Limitations and Future Scope If Outlined by the Study)
							Pedestrian	Bicycle	
Liu and Griswold (2009)	Location: San Francisco, CA Location type and size: pedestrian counts at 63 intersections	Data collection time and type: 12 days in May, June, August, and September 2002 from 2:30 to 6:30 p.m., weekdays Count method: manual observation	Not reported	Regression analysis	Not reported	18 variables at five buffer sizes	Presence of bike lane; job density (1/4 mi); percent of residential land use (1/16 mi); number of municipal transportation agency bus or light-rail stops (3/8 mi); population density (1/2 mi); mean slope (1/16 mi); patch[e] richness density (1/16 mi)	Not reported	Perception of crime safety may be a better measure for modeling pedestrian flow. Due to data limitations, the model cannot be generalizable to some neighborhoods with steep terrain and curvilinear street patterns.
Schneider et al. (2009b)	Location: Alameda County, CA Location type and size: pedestrian counts at 50 intersections along arterial and collector roadways	Data collection time and type: two counts of each intersection between April and June 2008, one a weekday count and the other a Saturday count during different times of the day Count method: manual observation	Not reported	OLS model	Validated based on models developed for 46 new intersections in different parts of Alameda County	50 variables at three buffer sizes	Total population (0.5 mi); number of jobs (0.25 mi); number of commercial retail properties (0.25 mi); presence of a regional transit station (0.1 mi)	Not reported	The model did not include pedestrian facility quality variables. It may not be representative of locations near special attractors. Different weighting factors may be assigned to different pedestrian attractors.

Author (Date)	Coverage and Data Collection Scale	Count Type and Time	Application Scale	Analysis Methods	Model Validation	Initial Number of Variables and Buffer Sizes	Significant Explanatory Variable (Buffer Size)		Comment (Limitations and Future Scope If Outlined by the Study)
							Pedestrian	Bicycle	
Pulugurtha and Repaka (2008)	<u>Location</u>: Charlotte, NC <u>Location type and size</u>: pedestrian counts at 176 signalized intersections	<u>Data collection time and type</u>: 12-hour (7:00 a.m. to 7:00 p.m.) counts at each site on a normal day in 2005 <u>Count method</u>: manual observation	Not reported	Multiple regression analysis through backward elimination of indepen-dent variables	Not reported	32 variables at three buffer sizes	Population (0.5 mi); single-family residential areas (0.5 mi); number of transit stops (0.5 mi)	Not reported	The model did not include variables such as population density, population by age group, and automobile ownership, which may yield better models.
Lindsey et al. (2007)	<u>Location</u>: Indianapolis, IN <u>Location type and size</u>: 30 locations on five multiuse greenway trails	<u>Data collection time and type</u>: continuous counts for different durations from 2001 through 2005 <u>Count method</u>: infrared monitors	Not reported	OLS model	The model was validated comparing the actual and predicted traffic on two of the trails of the study	29 variables	Weekend; month of the count; deviation of daily average temperature; deviation of daily precipitation accumulation; deviation of daily snow accumulation; deviation of daily percentage sunshine; percent population less than 5 and greater than 64; percent African American; percent other ethnicities, excluding White and African American; percent population 25+ with college degree; median household incomes; mean Normalized Difference Vegetation Index (NDVI)[f] value in census blocks (0.5 mi); population density (0.5 mi); percent of commercial land use in the trail neighborhood; parking lot area in the trail neighborhood; average length of network street segments (0.5 mi)		The errors associated with extrapolation of a small sample of peak-hour counts to annual counts ranged from 6% to 36%.

Author (Date)	Coverage and Data Collection Scale	Count Type and Time	Application Scale	Analysis Methods	Model Validation	Initial Number of Variables and Buffer Sizes	Significant Explanatory Variable (Buffer Size)		Comment (Limitations and Future Scope If Outlined by the Study)
							Pedestrian	Bicycle	
Desyllas et al. (2003)	Location: central London Location type and size: pedestrian counts at the midpoint of 231 street segments	Data collection time and type: counts were made between 8 a.m. and 7 p.m. in March 2000, July 2001, and August 1999 Count method: manual observation	Predicted flow per hour values for all 7,526 street segments in the 25 km^2 area of central London	Stepwise regression model	Not reported	15 variables	Average visibility within the street network; accessibility to a London Underground station; pavement width; percentage of frontage that is retail	Not reported	Additional explanatory variables such as urban density should be included in future studies.
Qin and Ivan (2001)	Location: rural areas, CT Location type and size: total number of crossing pedestrians, pedestrian counts at 32 sites	Data collection time and type: weekday and weekend counts at each site, generally from 8:00 a.m. to 5:30 p.m. in May, June, October, and November 1999 Count method: manual observation	Not reported	Generalized linear model	Not reported	Four variables	Presence of sidewalk; number of lanes of road; campus area; tourist and downtown area	Not reported	Only the total number of crossing pedestrian exposures was of interest. Walking on the sidewalk was not included in the study.

Author (Date)	Coverage and Data Collection Scale	Count Type and Time	Application Scale	Analysis Methods	Model Validation	Initial Number of Variables and Buffer Sizes	Significant Explanatory Variable (Buffer Size)		Comment (Limitations and Future Scope If Outlined by the Study)
							Pedestrian	Bicycle	
Behnam and Patel (1977)	Location: Milwaukee, WI Location type and size: pedestrian counts from a midblock of the streets within the study area (CBD)	Data collection time and type: 6-minute counts at each station between 6 a.m. and 6 p.m. during summer 1971 through 1973 Count method: manual observation	Not reported	Stepwise regression model	Not reported	Eight variables	Commercial space area; office space area; cultural and entertainment space area; residential space area; vacant space area; storage and maintenance space area	Not reported	The model only considered land-use variables. Applications are limited to very dense CBDs during specific times of day.
Pushkarev and Zupan (1971)	Location: midtown Manhattan, NY Location type and size: pedestrian counts at block sectors	Data collection time and type: instantaneous counts during midday and evening rush hour Count method: aerial photographs	Not reported	Multiple-correlation analysis	Not reported	11 variables	Available walkway space area; office floor space in each of 10 buildings; retail floor space in each of 10 buildings; restaurant floor space in each of 10 buildings	Not reported	The model has data limitations because the observed values are based on instantaneous photographs.

[a] *Centrality* is the number of times a link in a network is used along the path of all shortest paths between all nodes.
[b] Accessibility measures were for the 60-minute time-based buffer.
[c] Gold-level communities are included in LAB bicycle-friendly community listings.
[d] *Land-use mix* is a measure of acres of retail, office, and commercial space per housing unit.
[e] *Patch density* is the areal density of patches (a contiguous area with the same land use).
[f] The NDVI determines the density of green on a patch of land.

As seen in Table 1, the generalized approach to developing a direct-demand model includes selection of independent variables, often at various spatial scales, and choice of a suitable analysis method to estimate pedestrian, bicycle, or trail traffic in an area or location. A few of the studies have used their developed model to predict nonmotorized traffic in locations where count data were not available. Additionally, to shed light on some unique approaches and noteworthy observations made by the studies, the following briefly discusses some recent research.

One recent study (Hankey et al., 2017) developed a direct-demand model for pedestrians and bicyclists in a small rural college town. For this study, the data collection sites were selected specifically to capture the spatial variability in pedestrian and bicycle activity. Since the study monitored nonmotorized activity in a small community, the sample size was deemed representative. While underscoring the need for developing separate models for pedestrians and bicyclists, the study demonstrates that nonmotorized count monitoring programs can be designed with multiple goals, including estimating performance measures and developing spatial models.

Hankey and Lindsey (2016) developed direct-demand models and explored trade-offs between fully specified (i.e., exploratory) models and reduced-form models that can be easily used in the field. The study also assessed the change in model performance when a location has multiple counts. Three types of facility-demand models were compared for both pedestrians and bicyclists. While the statistically optimal models for bicyclists included almost all the variables, the same model for pedestrians included only 11 variables, and most of those variables followed the expected directions for the variables. The researchers then developed a reduced-form core model that ensured that the direction of the effects of the variables conformed to the theories. In addition to the variables used in bicycle models, the models for pedestrians included a few other variables, including transit stops and major roads. The pedestrian model did not include weather variables because it was assumed that weather has a greater impact on bicyclists than on pedestrians. The research highlighted the need for further research to investigate how count programs developed for spatial modeling can decrease the potential for spatial autocorrelation.

Fagnant and Kockelman (2016) developed a suitability model to identify the relationship between a number of variables and the expected share of bicyclist counts. The study applied the developed model to Shoreline, Washington’s road network and mapped the *Highway Capacity Manual* (HCM) bicycle level of service (BLOS) values of the intersecting streets; the study found little correlation between bicyclist volume and HCM BLOS.

Wang et al. (2013) investigated the suitability of different regression models to predict nonmotorized activity and developed eight models (which encompassed general, location-specific, and trail-specific models) to estimate trail traffic in Minnesota. The general model is intended to be used at similar locations without traffic counts. The location-specific models can be used in the specific location when counts from monitors are not available, and the trail-specific models can explain the variation in trail activity in response to variations in the weather and day of the week. Comparing the prediction error across the models, the study indicated that the ranges of errors for the trail-specific models were smaller than for the other models.

Hankey et al. (2012) converted hourly counts to 12-hour daily counts using a scaling factor and developed models for estimating pedestrian and bicycle traffic. After analyzing the data, researchers concluded that 1-hour counts (peak hour) were highly correlated with 12-hour daily counts; thus, the peak-hour count can be scaled to a long-term value to inform decisions and planning when needed.

Schneider et al. (2012) used manual and automated pedestrian counts at 50 intersections in San Francisco, California. The manual 2-hour count was adjusted with automated-counter, temporal, and weather adjustment factors to estimate the annual pedestrian crossing volumes. The study also mentioned the limitations and applicability of the models, which may not perform well in areas where pedestrian volume is highly variable, such as major parks, waterfronts, and sports arenas. Researchers proposed testing other variables in future studies, including the overall mix of land use, special pedestrian generators (such as schools), sidewalk width and buffer between the roadway and sidewalk, roadway width and number of motor vehicle lanes, percentage of households with no vehicles available, types of transit facilities (e.g., light-rail stop versus bus stop), presence of bicycle parking, and components of the San Francisco pedestrian environmental quality index.

Griswold et al. (2011) developed pilot models of bicycle intersection volumes in Alameda County. Researchers developed four models using log-linear OLS regression and revealed differences in intersection volume between weekdays and weekends. One of the interesting findings of the study was the variation of influence of the explanatory variables across weekdays and weekends. While the proximity to commercial retail properties or a large university had greater impact on volumes on weekdays, marked bicycle facilities had greater association with volumes on weekends. Researchers concluded that the model can be used to estimate bicycle volume during specific time periods.

Miranda-Moreno and Fernandes (2011) developed four models for the a.m. peak period, noon period, p.m. peak period, and entire day to explain pedestrian activity in Montreal, Canada. The study was one of the early attempts to generate a spatiotemporal model of pedestrian activity for a large city. The study had three limitations:

- The potential spatial correlation of the intersections was not accounted for.
- Intersections were not randomly selected.
- Volumes were not standardized to account for the temporal and weather variations.

The California Department of Transportation sponsored the Seamless Travel Study (Jones et al., 2010), which was performed by Alta Planning and Design and the University of California, Berkeley Traffic Safety Center. At that time, the study was the largest and longest combined effort of manual and automated counts of nonmotorized activity in the United States. Researchers developed models to predict pedestrian and bicycle volumes at intersections during the morning peak period (7 to 9 a.m.) on weekdays in San Diego, California. The study went through a comprehensive, systematic procedure to screen initially selected independent variables. One of the interesting observations of the study suggested that separate bike lanes are not an indicator of bicycle use since often they are built based on location feasibility rather than high bicycle traffic. Although bike lanes are an attractive feature for bicyclists, all things being equal, bicyclists tend to follow the most direct route with the best combination of other features such as topography, lane width, and traffic speeds. The study indicated that a model with refinement factors—with variables triggered by specific thresholds of volumes—may improve prediction accuracy.

Using collected data in Alameda County, California, Schneider et al. (2009b) recommended a simple tool that can roughly estimate pedestrian crossing volumes at intersections. To avoid bias in location choice, the study used a structured process to select sites with a wide range of pedestrian volumes and other characteristics. A unique approach of the study was that it considered the variation in the pedestrian volume trend at different locations (e.g., employment centers, residential areas, neighborhood commercial areas, and multiuse trails) to make necessary adjustments to the model. The study also accounted for the differences in time (e.g., time of day and day of week) and weather (e.g., cloud cover and temperature) to adjust and extrapolate 2-hour counts to weekly volume. The study recommended future studies to assign differential weighting factors to various pedestrian trip attractors, such as commercial properties, regional transit stations, and schools.

A study in London (Desyllas et al. 2003) developed a citywide pedestrian-demand model using the direct-demand modeling approach. In addition to capacity, accessibility, and land-use variables, the study used street grid configuration (visibility within the street network) variables using visibility graph analysis, which was a noteworthy approach. The study found visibility the most significant variable and suggested that pedestrians may select routes based on complexity and not based on distance. Further, pedestrians with limited knowledge about street configuration may choose routes that are the most visible.

Qin and Ivan (2001) estimated pedestrian crossing volume as a measure of exposure based on sociodemographic, traffic signal, and land-use characteristics in rural areas in Connecticut. The count in the study did not include pedestrians using sidewalks. The number of lanes, area type, and sidewalk system were significantly associated with the weekly pedestrian exposure in that area. The study showed that factors influencing pedestrian activity in rural settings may differ from that of urban, suburb, and rural mixture conditions.

Data Analysis

This section outlines the data analysis approaches, including the use of statistical models and independent variables of different studies. Since the choice of models and independent variables varies widely across the research, the next section discusses the data analysis procedure in an aggregated manner.

Choice of Model

A wide variety of approaches and methods have been used in predicting nonmotorized activity using direct-demand models. However, only a few studies have discussed the implications of model choices to predict nonmotorized traffic activity (Kim and Susilo, 2011). A number of studies have used OLS regression to develop models (Lindsey et al., 2007; Schneider et al., 2009b). Studies have discussed the limitations of OLS models, where the output could be a negative pedestrian and bicycle traffic volume when in actuality the result must be zero or positive (Hankey et al., 2012). Hankey et al. (2012) compared the performance of negative binomial and OLS regression by comparing predicted and estimated pedestrian and bicyclist counts at 85 locations in Minneapolis, Minnesota, and concluded that generally the negative binomial model performs better. Wang et al. (2013) indicated that Poisson and negative binomial models tend to be a better fit for urban nonmotorized traffic characteristics than OLS regression models.

Kim and Susilo (2011) compared the performance of the Poisson regression model and negative binomial regression model in predicting pedestrian demand at a regional level. The authors concluded that the negative binomial regression is more appropriate to explain the overdispersed dependent variables than the Poisson regression. Other studies have also indicated that the negative binomial models tend to be statistically superior to Poisson regression models in explaining nonmotorized behavior (Cao et al., 2006; Baran et al., 2008; Fagnant and Kockelman, 2016).

Nordback (2012) investigated several models, including multivariate linear regression models, general linear models with log-linear regression, generalized additive models, Poisson models, Gaussian models, and negative binomial generalized linear models with log-link. The aim of the study was to develop a predictive model to estimate bicycle counts as a function of time and weather variables. The study concluded that the negative binomial generalized linear model with log-link was the best model to fit the positively skewed hourly count data.

Dependent Variables

Typically, the dependent variables of direct-demand models are pedestrian, bicycle, or trail traffic volumes for various time periods, such as during the peak period, hourly, daily, or annually. While some research has directly used data for the specific collection period, other studies have expanded short-period data to longer periods by using a scaling factor to be integrated and used in models (Lindsey, 2011; Hankey et al., 2012; Wang et al., 2016).

Several examples are present in the field. For instance, one study scaled hourly count data to 12-hour daily volume to model pedestrians and bicyclists on streets in Minneapolis, Minnesota (Hankey et al., 2012). Another study translated instantaneous counts to hourly flow rates (Pushkarev and Zupan, 1971). Another study conducted short pedestrian counts at 50 locations and adjusted the volume to account for the type of area and the difference in time and weather; the study then extrapolated to weekly counts (Schneider et al., 2009b). Nordback (2012) developed a method to estimate annual average daily bicycles in Boulder, Colorado, based on short-term counts (between 1 hour and 1 month). Lindsey et al. (2007) extrapolated hourly counts to annual estimates and used a natural logarithm to normalize the distribution in order to use it in the OLS regression.

Independent Variables

Although traditional belief is that the important dimension for bicycling is the proximity of cycling-specific infrastructure and the important dimension for walking is the proximity of neighborhood retail (Krizek, 2006), the volume of pedestrians and bicyclists tends to be governed by a broad array of factors. Often, the literature aggregates the two nonmotorized modes, walking and bicycling, when identifying the determinant factors (Pikora et al., 2003). However, walking and bicycling trip behaviors differ substantially and need to be investigated separately (Krizek, 2006; Hankey and Lindsey, 2016). Hankey and Lindsey (2016) indicated that pedestrian traffic is mostly influenced by activity centers with high job accessibility and public transport infrastructure, whereas bicycle traffic is mostly influenced by job accessibility and bicycle facilities. Moreover, Jones et al. (2010) revealed that explanatory variables such as the built environment and socioeconomic characteristics were different for high- and low-volume pedestrian intersections, while the same variables were not found significantly different for low- and high-volume bicycle locations.

To identify an initial set of independent variables, studies have taken into consideration the previously established relationships between nonmotorized activity and various sociodemographic, land-use, and surrounding built-environment factors (Kitamura et al., 1997; Moudon et al., 1997; Landis et al., 2001; Dill and Carr, 2003; Dill and Voros, 2007). Often, different measuring units are used to compute the same explanatory variable. For example, to compute weather variables (precipitation, snow, etc.), Lindsey et al. (2007) used deviations from the long-term daily average, while Nordback (2012) used the daily/hourly average.

To identify the impact of land-use and built-environment characteristics on nonmotorized volume, a number of studies have considered a range of buffer widths. Miranda-Moreno and Fernandes (2011) used buffers of 50 m, 150 m, and 400 m around intersections to investigate the impact of immediate surroundings, within-close-proximity urban form features, and within-walking-distance features, respectively, on pedestrian activity. Liu and Griswold (2009) concluded that the area around a one-block radius of an intersection has the strongest influence on pedestrian volume. Pulugurtha and Repaka (2008) investigated the influence of independent variables by different buffer widths and concluded that demographic and socioeconomic variables are more predominant within a 0.25-mi buffer width than a 0.50-mi or 1-mi buffer width, whereas speed limit and traffic volume variables are more predominant within a 0.50-mi or 1-mi buffer width than a 0.25-mi buffer width. This study also found that the 0.50-mi buffer width model tended to be a better fit than the 0.25-mi and 1-mi buffer models. However, Liu and Griswold (2009), Miranda-Moreno and Fernandes (2011), and Schneider et al. (2009b) suggested that the best model may be obtained using different scales of buffer zones for variables because the variables are unlikely to be significant at the same buffer scale. Hankey and Lindsey (2016) confirmed the assertion, stating that some variables (e.g., industrial area and population density) have the greatest influence on a large spatial area (more than 1 km), and some variables (e.g., bicycle facilities, retail areas, and open spaces) may influence a smaller spatial area (100 to 400 m).

The majority of the studies began with a large number of independent variables that were subsequently reduced through a systematic statistical procedure such as correlation and skewness testing. For example, Miranda-Moreno and Fernandes (2011) performed a multicorrelation analysis to identify variables with high correlations. The study generated a correlation matrix for the 50-m, 150-m, and 400-m buffer widths. Other than the criteria of 95th percentile significance and correlations less than 0.4, researchers also used intuition when selecting variables for the preferred models.

To review the independent variables used in several studies, this report categorizes the variables into the following nine groups:

- Demographic.
- Socioeconomic.
- Network/interaction with vehicle traffic.
- Pedestrian- or bicycle-specific infrastructure.
- Transit facilities.
- Major generators.
- Weather and environmental.
- Temporal or time related.
- Land use.

Table 2 depicts all the independent variables considered in bicycle, pedestrian, and trail traffic direct-demand model studies reviewed in this study. As mentioned previously, the same variable can have different measuring units. The table mainly lists a generalized form of the independent variables; for example, household income may represent either median household income or average household income.

Table 2. Explanatory Variables Used in Pedestrian, Bicycle, and Trail Traffic Model Studies.

Variable Category	Variable	References
Demographic	Population density Percentage of population younger than 5 and older than 65 years Percentage of African-American population Percentage of Hispanic population Percentage of other ethnicity population (excluding White and African American) Percentage of population 25+ with a college degree Number of children Household density Total housing units Vacant housing units Rented housing units Percentage of male, single, and multifamily housing Commuting population Walking and biking commuters	Hankey et al. (2012, 2017); Fagnant and Kockelman (2016); Hankey and Lindsey (2016); Tabeshian and Kattan (2014); Strauss and Miranda-Moreno (2013); Wang et al. (2013); Jones et al. (2010); Lindsey et al. (2006, 2007, 2012); Lindsey (2011); Liu and Griswold (2009); Miranda-Moreno and Fernandes (2011); Pulugurtha and Repaka (2008); Schneider et al. (2009b, 2012)
Socioeconomic	Household income Employment density Unemployment rate Households with no automobile Households below the poverty line Number of workers	Hankey et al. (2012, 2017); Fagnant and Kockelman (2016); Hankey and Lindsey (2016); Jones et al. (2010); Lindsey et al. (2006, 2007 2012); Lindsey (2011); Liu and Griswold (2009); Miranda-Moreno and Fernandes (2011); Schneider et al. (2009b, 2012); Tabeshian and Kattan (2014); Wang et al. (2013); Strauss et al. (2013); Strauss and Miranda-Moreno (2013)
Network/ interaction with vehicle traffic	Number of street segments Average length of network street segments[a] Steeper slopes Presence of traffic signals Number of intersections Percentage/length of major roads Length of local roads Presence of three-way or four-way intersections Mean block length Presence of arterial streets/freeways Maximum average daily traffic volume Average curb-to-curb length Average number of lanes Speed limit Bridges Intersection density Connected node ratio Road classification Lane visibility Network accessibility Maximum radial line of sight	Hankey et al. (2012, 2017); Fagnant and Kockelman (2016); Griswold et al. (2011); Hankey and Lindsey (2016); Jones et al. (2010); Lindsey et al. (2006, 2007, 2012); Lindsey (2011); Liu and Griswold (2009); Miranda-Moreno and Fernandes (2011); Pulugurtha and Repaka (2008); Schneider et al. (2009b, 2012); Qin and Ivan (2001); Strauss and Miranda-Moreno (2013); Desyllas et al. (2003)

Variable Category	Variable	References
Pedestrian- or bicycle-specific infrastructure	Presence of bike lanes Length of off-street trail Length/presence of bike paths Area of sidewalk coverage Sidewalks with buffer[b] Number of marked crosswalks Median refuge areas Bike-lane width Curb-lane width Sharrows, crosswalks, and pedestrian heads Bicycle facility characteristics on the road	Hankey et al. (2012, 2017); Fagnant and Kockelman (2016); Hankey and Lindsey (2016); Jones et al. (2010); Lindsey (2011); Liu and Griswold (2009); Pushkarev and Zupan (1971); Schneider et al. (2009b); Tabeshian and Kattan (2014); Qin and Ivan (2001); Strauss et al. (2013); Strauss and Miranda-Moreno (2013); Desyllas et al. (2003)
Transit facilities	Presence of subway stations Number of bus/light-rail stops Mileage of bus route Percentage of commuters who walk or take transit Distance to stations Number of jobs accessible by transit Transit ridership Total bus-km of bus routes	Hankey et al. (2012, 2017); Desyllas et al. (2003); Griswold et al. (2011); Hankey and Lindsey (2016); Jones et al. (2010); Lindsey (2011); Liu and Griswold (2009); Miranda-Moreno and Fernandes (2011); Schneider et al. (2009b); Tabeshian and Kattan (2014); Strauss et al. (2013); Strauss and Miranda-Moreno (2013)
Major generators	Distance to downtown Distance to ocean or a water body Distance to a university Number of schools (elementary, middle, and high) Number of college campuses	Griswold et al. (2011); Hankey et al. (2012); Lindsey et al. (2012); Miranda-Moreno and Fernandes (2011); Schneider et al. (2009b, 2012); Qin and Ivan (2001); Strauss et al. (2013); Strauss and Miranda-Moreno (2013)
Weather and environmental	Temperature Precipitation Snow accumulation Sunshine Solar radiation Wind Humidity	Fagnant and Kockelman (2016); Hankey and Lindsey (2016); Lindsey et al. (2006, 2007, 2012); Lindsey (2011); Nordback (2012); Wang et al. (2013); Strauss and Miranda-Moreno (2013)
Temporal or time related	Month, hour, or day Weekend Holiday School day Season or year	Fagnant and Kockelman (2016); Hankey and Lindsey (2016); Hankey et al. (2012); Lindsey et al. (2006, 2007, 2012); Nordback (2012); Wang et al. (2013)
Land use	Housing units, all households, residential addresses, non-residential addresses House density, low density residential space, medium density residential space, high density residential space, dwell, single family housing, multi-family housing Number of vacant housing, proportion of vacant housing, Number of rented housing, proportion of rented housing Urban residential area, urban residential commercial area, residential—mobile, resort residential, urban residential commercial area Neighborhood business Job accessibility Historic district, community service Park recreation education Manufactured house, public buildings Open space area, vacant space	Hankey et al. (2012, 2017); Behnam and Patel (1977); Desyllas et al. (2003); Fagnant and Kockelman (2016); Hankey and Lindsey (2016); Jones et al. (2010); Lindsey et al. (2006, 2007, 2012); Lindsey (2011); Liu and Griswold (2009); Miranda-Moreno and Fernandes (2011); Pulugurtha and Repaka (2008); Pushkarev and Zupan (1971); Schneider et al. (2012); Tabeshian and Kattan (2014); Qin and Ivan (2001); Strauss et al. (2013); Strauss and Miranda-Moreno (2013)

Variable Category	Variable	References
	Tree canopy, non-tree vegetation, patch richness density, shannon's diversity index, impervious surface Paved parking Slope Institutional, research district, neighborhood service district, cultural and entertainment space Business, office space, retail area, industrial area commercial space, storage and maintenance space Government Hotel, restaurant, commercial center Airport Direct control space Hazardous waste district High-activity zone, gigh crime Land use mix, land use charecteritics, land use type, mixed land use Visibility, maximum radial line of sight. Accessibility Planned unit development Innovative	

[a] *The average length of network street segments* is the total road length divided by the number of street segments.
[b] Surface streets are separated from the edge of the roadway by grass, trees, shrubs, or other types of buffers.
[c] Shannon's diversity index is characterized by species diversity in a community.
[d] *Innovative land use* is a nontraditional and new type of land use.

Relationship with Independent Variables

This section provides a brief discussion of the independent variables and their impact on nonmotorized activity, which are explained by several studies. Table 3, Table 4, and Table 5 identify the positive or negative impact of each of the significant variables on pedestrian, bicycle, and trail traffic, respectively. The tables provide clear insight into whether the variables have consistent impact on nonmotorized volume across the studies. The following discussion briefly explains the characteristics and influence of the independent variables depicted in the tables.

Table 3. Influence of Explanatory Variables on Pedestrian Volumes.

Variables	Behnam and Patel (1977)	Desyllas et al. (2003)	Hankey et al. (2017)	Hankey and Lindsey (2016)	Hankey et al. (2012)	Jones et al. (2010)	Lindsey (2011)	Liu and Griswold (2009)	Miranda-Moreno and Fernandes (2011)	Pulugurtha and Repaka (2008)	Qin and Ivan (2001)	Schneider et al. (2012)	Tabeshian and Kattan (2014)
Demographic													
Population density			+	+		+		+					
Population									+	+			
Percent of non-White residents					+								
Percent of residents with a college education					+		+						
Percent of Black residents							–						
Socioeconomic													
Household income			–				–						
Total employment												+	
Employment density						+		–					
Network/interaction with vehicle traffic													
Major roads length				+									
Percent of major arterials									–				
Number of street segment									+				
Street lengths													+
Principal arterial street (of count location)					–								
Arterial street (of count location)					+		+						
Collector street (of count location)					+		+						
Presence of four-way intersection									+				
Number of lanes											+		
Pedestrian- or bicycle-specific infrastructure													
Sidewalk length			+										
Off-street trail length			–	+									
Pathway length													+
Presence of bike lane								+					
Presence of sidewalk											+		
Footway pavement width		+											
Transit facilities													
Number of transit/bus stops			+	+				+	+	+			+
Presence of a subway stop									+				
Bus frequency													+
Accessibility to an underground station		+											
Major generators													
Distance from the CBD/downtown					–				–				
Distance from the nearest body of water					–								
Proximity to a university campus											+	+	
Number of schools									+				

Variables	Behnam and Patel (1977)	Desyllas et al. (2003)	Hankey et al. (2017)	Hankey and Lindsey (2016)	Hankey et al. (2012)	Jones et al. (2010)	Lindsey (2011)	Liu and Griswold (2009)	Miranda-Moreno and Fernandes (2011)	Pulugurtha and Repaka (2008)	Qin and Ivan (2001)	Schneider et al. (2012)	Tabeshian and Kattan (2014)
Weather and environmental													
Precipitation					–								
Recorded temperature							–						
Very warm temperature (max. temperature >32°C)									–				
Land use													
Residential land use	+		–					–				+	
Land-use mix (area of retail, office, and commercial space per housing unit)							+						
Retail area		+		+		–							
Office space area	+												
Industrial area				–									
Cultural and entertainment space area	+												
Storage and maintenance space area	+												
Vacant space area	+												
Open space area				+									
Job accessibility				+									
Dwell count													+
Commercial space	+								+				+
Open space									–				
Schools													+
High-activity zone intersection												+	
Maximum/mean slope								–				–	
Traffic-signal-controlled intersection												+	
Patch richness density								+					
Single-family residential areas										–			
Average visibility within the street network		+											
Tourist and downtown area											+		

Table 4. Influence of Explanatory Variables on Bicycle Volumes.

Variables	Fagnant and Kockelman (2016)	Griswold et al. (2011)	Hankey et al. (2017)	Hankey and Lindsey (2016)	Hankey et al. (2012)	Jones et al. (2010)	Lindsey et al. (2012)	Lindsey (2011)	Schneider et al. (2009b)	Strauss et al. (2013)	Strauss and Miranda-Moreno (2013)	Tabeshian and Kattan (2014)
Demographic												
Population density			+	+								
Percent of non-White residents					+		+					
Percent of Black residents								−				
Percent of residents with a college education					+		+	+				
Percent of residents younger than 5 and older than 65 years								+				
Total population									+			
Socioeconomic												
Household income			−		−		−	−			+	
Employment density	+					+						
Number of jobs									+	+	+	
Network/interaction with vehicle traffic												
Major roads length			+									
Number of lanes	+											−
Speed limit	−											
Off-street trail (of count location)					+		+					
Arterial street (of count location)					+		+					
Pedestrian- or bicycle-specific infrastructure												
On-street bicycle facility length			+	+			+			+	+	
Presence of a bicycle track											+	
Presence of a bicycle lane											+	
Off-street trail length				+				+				
Bicycle-trail access	+											
LAB gold[a]	+											
Curb-lane width	+											
Bike-lane width	+											
Separated path	+											
Minor arterial with a bicycle								+				
Minor arterial without a bicycle								+				
Presence of bicycle markings on any approach		+										
Footage of bicycle network I (Multi-Use Path)						+						
Transit facilities												
Number of bus/transit stops									+		+	+
Presence of a subway station										+	+	
Major generators												
Distance from the CBD					−		−					
Distance from a university		−										
Presence of schools										+	+	

Variables	Fagnant and Kockelman (2016)	Griswold et al. (2011)	Hankey et al. (2017)	Hankey and Lindsey (2016)	Hankey et al. (2012)	Jones et al. (2010)	Lindsey et al. (2012)	Lindsey (2011)	Schneider et al. (2009b)	Strauss et al. (2013)	Strauss and Miranda-Moreno (2013)	Tabeshian and Kattan (2014)
Weather and environmental												
Precipitation				–	–		–					
Temperature				+				+				
Deviation from daily high temperature								–				
Temporal or time related												
Morning period count	–											
Year					+		+					
Land use												
Centrality[b]			+									
Retail area				+								
Industrial area				–								
Open space area				+								
Job accessibility[c]				+								
Bridge	+											
Residential area	–											
Commercial space		+							+	+		+
Low-density residential space												+
Institutional space												+
Land-use mix					+		+	+		+	+	
Average slope (degrees) of terrain		–										
Presence of three approaches										–		
Presence of parking entrance											–	

[a] Gold-level communities are included in LAB bicycle-friendly community listings.
[b] *Centrality* is the number of times a link in a network is used along the path of all shortest paths between all nodes.
[c] All accessibility measures were for the 60-minute time-based buffer.

Table 5. Influence of Explanatory Variables on Trail Traffic Volumes.

Variables	Lindsey et al. (2007)	Wang et al. (2013)
Demographic		
Percent of African-American population	+	+
Percent of other ethnicity population, excluding White and African American	+	
Percent with a college education	+	+
Percent of population over 64 or below 6	–	–
Population density	+	+
Socioeconomic		
Median household income	+	+
Network/interaction with vehicle traffic		
Street length	+	
Weather and environmental		
Recorded high temperature		+
Deviation of temperature from the long-term average	+	–
Precipitation		–
Deviation of precipitation from the long-term average	–	
Deviation of daily snow from the long-term average	–	
Deviation of daily percentage sunshine from the long-term average	+	
Wind speed		–
Temporal or time related		
Weekend	+	
Month	+	
Land use		
Parking lot area	+	
Commercial land use	+	
Mean NDVI value	+	

Demographic

Studies have used a wide array of demographic variables to explain the nonmotorized activity in an area. Population and employment density are two of the easiest to compute and two of the most frequently used variables in these studies. A demographic characteristic such as a high population of residents with a college education tends to have a positive relationship with pedestrian and bicycle activity. A study by Lindsey (2011) expected a high percentage of residents over the age of 65 or under the age of 5 to have a positive effect on pedestrian volume but a negative one on bicycle volume. However, contrary to the expectation, the study found bicycle traffic to be positively associated with the population in those age group. On the other hand, a high population in those age groups had a negative influence on trail traffic (Lindsey et al., 2007; Wang et al., 2013).

Socioeconomic

The volume of pedestrians and bicyclists appears to be associated with income, employment, and vehicle ownership characteristics. Lower-income citizens are expected to walk more, considering their lack of access to cars. A majority of the studies found high income to have a negative impact on pedestrian and bicycle activity. Although not consistent in all of the studies, employment density is generally found to be positively correlated with pedestrian and bicycle activity, as might be expected.

Network/Interaction with Vehicle Traffic

Street network and roadway characteristics tend to influence nonmotorized activity. Studies indicate that communities with greater street connectivity observe higher pedestrian activity compared to communities without street connectivity (Moudon et al., 1997). A steep slope at intersections is inversely correlated with pedestrian traffic, whereas signal-controlled intersections have a positive impact on pedestrians (Schneider et al., 2012). Other traffic characteristics such as vehicular traffic volume and traffic speed influence traffic safety and therefore pedestrian and bicycle activity (Landis et al., 2001). A majority of the studies found a positive correlation between major streets and pedestrian and bicyclist volume. Hankey et al. (2012) explained the association by stating that pedestrians and bicyclists want to reach destinations for which the most efficient access is provided by the major roads.

Pedestrian- or Bicycle-Specific Infrastructure

Pedestrian- and bicycle-specific facilities are one of the most significant determinants of walking and bicycling volume in an area. Studies have indicated that within the same functional type street facility, without controlling for other factors, average bicycle volume tends to be higher on streets with separate bicycle facilities than on streets with none (Lindsey, 2011). However, Hankey et al. (2017) found off-street trails to be negatively correlated with pedestrian volume. The study suggested that the finding might be the result of high pedestrian volume on sidewalks near retail areas and on local roads near university campuses. Studies have included other variables such as median refuge areas, bike-lane width, curb-lane width, sharrows, and crosswalks to investigate their impact on nonmotorized activities and generally found nonmotorized-friendly designs and infrastructures to have a positive impact on the volume.

Transit Facilities

The availability of and accessibility to a transit facility seem to have a significant impact on pedestrian and bicycle activity because people generally walk or bicycle to transit stops. Studies conducted in both Canada and the United States have found high pedestrian activity in places with high transit use (Schneider et al., 2009b; Miranda-Moreno and Fernandes, 2011). The presence and number of transit stops, transit frequency, and accessibility to transit stations have a positive correlation with nonmotorized activity.

Major Generators

Downtown areas, CBDs, and university areas generate a large amount of walking and bicycling traffic partly because they often provide limited parking facilities. Schools and colleges, especially major university campuses, are considered prominent pedestrian and bicycle trip attractors. The presence of water bodies, including seas, lakes, rivers, and creeks, may also attract nonmotorized activity (Hankey et al., 2012).

Weather and Environmental

Studies have used a number of weather variables to explain the variation in nonmotorized activity. Lindsey et al. (2006, 2007) computed variables for temperature, precipitation, snow accumulation, and percentage of sunshine as the deviation from the long-term daily mean. Nordback (2012) computed precipitation, temperature, and solar radiation as daily and hourly values. Studies have indicated that nonmotorized activity has a negative correlation with precipitation and a positive correlation with temperature. However, nonmotorized volume may decrease with extremely high temperatures (Stinson and Bhat, 2004). Studies conducted in places such as San Diego, California, where the weather is mild, did not find weather to be a significant factor (Jones et al., 2010). The pedestrian model developed by Hankey and Lindsey (2016) did not include a weather variable with the assumption that weather has a greater impact on bicyclists than on pedestrians. Findings in a study by Lindsey (2011) corroborate this assumption.

Temporal or Time Related

To control the time effects—including day of the week, month, and year—studies use dummy variables for months, years, weekend days, and so forth. Differences in day-of-the-week nonmotorized activity volumes may account for utilitarian or recreational trips depending on location (Nordback, 2012). Studies have indicated that mean weekend daily traffic tends to be greater than mean weekday daily traffic (Lindsey et al., 2007).

Land Use

The characteristics of surrounding land use may be fundamental in shaping walking and bicycling behavior in an area. Compact, mixed land-use characteristics at employment or commercial areas promote nonmotorized activity (Kuzmyak et al., 2014). Commercial and office areas may attract more pedestrians than industrial areas do (Pulugurtha and Repaka, 2008). Slope, presence of three- or four-way intersections, presence of bridges, and other elements have a significant impact on nonmotorized volume. Studies have used different buffer widths to compute the impact of various land-use characteristics. Liu and Griswold (2009) found a positive correlation between crime level and pedestrian count. Although the result is counterintuitive, the study indicated that crime exposure may be a more suitable variable than raw crime measures. The same study also used a patch richness density and Shannon's diversity index variable as a measure of landscape diversity.

Desyllas et al. (2003) acknowledged the influence of street configuration on nonmotorized activity and argued that the land-use characteristics within a buffer area may represent the attractiveness of origins or destinations but may not capture pedestrians' preference of route choice between a set origin and destination. The model developed by Pushkarev and Zupan (1971) accounted for the unique geometry of the Manhattan street grid by differentiation between streets and avenues. Hillier (1996) suggested that pedestrian movement patterns are influenced by the layout of the street grid of a city area and that measures of the street grid's spatial configuration should be used as an explanatory variable in the pedestrian flow model. Desyllas et al. (2003) developed a pedestrian model where average visibility within the street network was used as one of the explanatory variables to represent the street grid configuration.

Model Benefits and Limitations

Direct-demand models are comparatively simple tools that enable transport planners to predict nonmotorized traffic at relevant locations where count data are not available. Aoun et al. (2015) listed a number of advantages of direct-demand models such as:

- Software requirements to develop and use the models are usually limited to Excel® spreadsheets or common statistical software packages.
- The models can be developed largely using existing data.
- Necessary data are typically available to the public and most often can be found at various geographic levels.

The direct-demand modeling approach also has some advantages over some other modeling approaches such as traditional four-step models. When compared to conventional models, studies indicated that some of the shortcomings of four-step models, such as errors in trip-end totals and errors generated by poorly estimated intra-zonal trips, can be avoided in direct-demand models because they simultaneously calibrate trip generation, distribution, and mode choice steps, including characteristics of other modes and a wide array of level-of-service and activity variables (Ortuzar and Willumsen, 2011; Aldian, 2005).

Ortuzar and Willumsen (2011) also indicated that direct-demand models are more useful for estimating demand in areas where the zones are large, such as those areas depicted in inter-urban studies. While discussing the use and suitability of demand models in developing countries, Timberlake (1988) found that direct-demand models were able to better accommodate the unique traffic characteristics of a corridor in Sudan than the gravity model.

While discussing the shortcomings of conventional modeling approaches, Domencich and McFadden (1974) highlighted the fact that conventional models are basically nonbehavioral and not policy oriented. The models do not necessarily reveal the interaction between system performance and the choices of trip frequency or trip destination. On the other hand, direct-demand models are expected to explain the impact of different factors that influence the demand for people's travel choice and therefore inform decision making.

A few studies have discussed shortcomings and recommended caution in using direct-demand models for nonmotorized traffic estimation. Barnes and Krizek (2005) examined why using predictive models to estimate nonmotorized activity based on explanatory factors such as land-use and sociodemographic factors is unlikely to ever be very precise. The study noted the following shortcomings of direct-demand models:

- People's perceptions of a facility are not considered.
- The pattern of activity varies widely with location.
- The range of sampling error tends to be many times greater than the sample mean due to the low levels of nonmotorized activity.
- Sometimes positive correlation between volume and facilities could be causation in the other direction, such as when a large number of bicyclists justifies an exclusive facility rather than a new facility.

When models are transferred far into the future and for large areas, the difference between people and locations may result in inaccurate estimates. Often, changes in technology and society have an unanticipated influence on user behavior (Katz, 2003). Highlighting the caveats of using this method, Kuzmyak et al. (2014) indicated that while the direct-demand model may establish a strong relationship between nonmotorized traffic volume and explanatory variables, the model cannot readily show the cause of the behavior represented in the count. The authors suggested that because the accuracy of the model is limited due to its aggregated structures, its use should be limited to preliminary analysis or screening until a more comprehensive model is available.

However, noting the relative advantage of direct-demand models over other modeling approaches, Kuzmyak et al. (2014) mentioned the need to be judicious in developing and applying these models and suggested the following guidelines for application:

- Models developed for a specific area cannot be construed as transferable.
- Uncertainties developed due to unaccounted origin-destination, route choice, and trip purpose data may be narrowed down by developing counts and models focused on a specific time period.
- After model calibration, the reliability of the models to predict volume in individual locations and the overall study area should be tested.
- Decisions or recommendations based on the models have to be carefully investigated.

Conclusion

This report presents an in-depth review of the available literature associated with direct-demand modeling to estimate nonmotorized activity. Researchers and various transport practitioners endeavor to use direct-demand modeling to estimate exposure for crash risk analysis. A steady evolution in methodology to overcome previous deficiencies and to yield new insights can be observed.

Studies have shown the need for developing count programs specifically tailored for spatial modeling to decrease the potential for spatial autocorrelation and improve the performance of facility-demand models. Through their exploration of a wide array of modeling techniques, the majority of the studies have acknowledged the suitability of negative binomial models in predicting pedestrian and bicycle volume. The wide range of independent variables include type, unit, and measured buffer width. Studies have suggested that not only should pedestrian and bicycle volumes be modeled separately, but also the factors may need to be interpreted differently for these nonmotorized activities (Lindsey, 2011). The influence of explanatory variables may provide useful insights in the decision-making process. For example, the positive association between major streets and nonmotorized volume may promote initiatives such as Complete Streets, and the significance of independent variables such as bus stops and off-street trails, at a small spatial scale, indicates that targeted improvements in infrastructure have the potential to encourage nonmotorized activities.

Although there is a tendency to associate walking and bicycling potential with a number of sociodemographic attributes (e.g., young or low-income individuals), given the right circumstances, nonmotorized modes can also be chosen by other sociodemographic segments. The choice of independent variables and their magnitude and direction of impact on nonmotorized activity largely depend on community, people, and location. For example, despite inclement weather and complex topography, many countries observe higher rates of pedestrians and bicyclists than the United States (Bassett et al., 2008).

Moreover, while differentiating the primary difference between the factors affecting walking and bicycling trips, studies have indicated that walking trips are more likely to be motivated by availability of sidewalks or land use characteristics, but decisions about bicycling trips may be affected by different factors across spatial areas beyond the trip origin (Winters et al., 2010). The geographic accessibility of destinations may also influence the decision to bicycle, considering the fact that the majority of bicycling trips are less than 10 km (Doyle et al., 2006)

Although a number of studies have established the relationship between bicycling and built environment, a few individual-level studies could not find strong associations between bicycling and built-environment variables (Moudon et al., 2005; Wendel-Vos, 2004). These studies underlined the methodological issue in bicycle studies where a spatial zone is selected to investigate the influence of the built environment. A buffer area of 1 mile may represent the walkable distance from home or origin point, but the activity space for bicycling should cover a larger area (Moudon et al., 2005).

These limitations warrant using emerging technologies such as GPS or mobile phone apps to accurately identify where people travel or engage in physical activity (Dill, 2009). The data from GPS-enabled smartphone apps, wearable tech, interactive websites, or even bike-share systems in a city—often termed *big data*—have the potential to describe the detailed spatio-temporal travel patterns of nonmotorized traffic in an unprecedented level of detail (Misra et al., 2014). Researchers have investigated various data sources to identify their usefulness to understand the travel patterns and behaviors of bicyclists and to eliminate traditional data collection efforts. A number of studies have already used some of the big data sources to estimate volume in a region (Strauss et al., 2015) and to shed light on the factors influencing nonmotorized trips of varying purposes (Griffin and Jiao, 2015).

Finally, although significant progress can be observed in the use of direct-demand modeling to estimate nonmotorized activity, challenges in data collection and model interpretation are yet to be fully resolved. The insights and findings in this report are intended to inform future research in the area of transportation-related policy making.

References

Aldian, A. (2005). *On the Development of Evaluation System and Transport Demand Model for Road Network Planning in Developing Countries (A Case Study of Indonesia)*. Doctoral dissertation, University of South Australia, Adelaide.

Anda, C., Erath, A., and Jacobu, P. (2017). Transport Modelling in the Age of Big Data. *International Journal of Urban Sciences*, Vol. 21, No. Sup 1, pp. 19–42.

Aoun, A., Bjornstad, J., DuBose, B., Mitman, M., Pelon, M., and Fehr and Peers (2015). *Bicycle and Pedestrian Forecasting Tools: State of the Practice.* DTFHGI-11-H-00024. Federal Highway Administration, Washington, DC.

Baran, P., Rodríguez, D. A., and Khattak, A. J. (2008). Space Syntax and Walking in New Urbanist and Suburban Neighbourhoods. *Journal of Urban Design*, Vol. 13, No. 1, pp. 5–28.

Barnes, G., and Krizek, K. (2005). Estimating Bicycling Demand. *Transportation Research Record*, No. 1939, pp. 45–51.

Bassett, D. R., Pucher, Jr., J., Buehler, R., Thompson, D. L., and Crouter, S. E. (2008). Walking, Cycling, and Obesity Rates in Europe, North America, and Australia. *Journal of Physical Activity and Health*, Vol. 5, No. 6, pp. 795–814.

Behnam, J., and Patel, B. (1977). A Method for Estimating Pedestrian Volume in a Central Business District. *Transportation Research Record*, No. 629, pp. 22–26.

Bhat, C. R., Guo, J. Y., and Sardes, R. (2005). *Non-motorized Travel in the San Francisco Bay Area*. Department of Civil Engineering, University of Texas at Austin.

Cao, X., Handy, S., and Mokhtarian, P. (2006). The Influences of the Built Environment and Residential Self-Selection on Pedestrian Behavior: Evidence from Austin, TX. *Transportation*, Vol. 33, No. 1, pp. 1–20.

Centers for Disease Control and Prevention (2016). *Bicycling and Walking in the United States 2016, Benchmarking Report*. Alliance for Biking and Walking, Washington, DC.

Centers for Disease Control and Prevention (2017). *Bicycle Safety*. National Center for Injury Prevention and Control, Atlanta, GA.

City of Greensboro, North Carolina (2015). *2015 Bicycle, Pedestrian, Trails and Greenways Plan Update*. The Greensboro Urban Area Metropolitan Planning Organization.

Desyllas, J. D. D., Duxbury, E., Ward, J., and Smith, A. (2003). *Pedestrian Demand Modelling of Large Cities: An Applied Example from London.* Centre for Advanced Spatial Analysis, London.

Dill, J. (2009). Bicycling for Transportation and Health: The Role of Infrastructure. *Journal of Public Health Policy*, Vol. 30, No. 1, pp. 95–110.

Dill, J., and Carr, T. (2003). Bicycle Commuting and Facilities in Major U.S. Cities: If You Build Them, Commuters Will Use Them. *Transportation Research Record,* No. 1828, pp. 116–123.

Dill, J., and Voros, K. (2007). Factors Affecting Bicycling Demand: Initial Survey Findings from the Portland Region. *Transportation Research Record*, No. 2031, pp. 9–17.

Domencich, T., and McFadden, D. (1974). *Urban Travel Demand: A Behavioral Analysis.* Charles River Associates, North-Holland, Amsterdam.

Doyle, S., Kelly-Schwartz, A., Schlossberg, M., and Stockard, J. (2006). Active Community Environments and Health: The Relationship of Walkable and Safe Communities to Individual Health. *Journal of the American Planning Association*, Vol. 72, No. 1, pp. 19–31.

El Esawey, M., Mosa, A., and Nasr, K. (2015). Estimation of Daily Bicycle Traffic Volumes Using Sparse Data. *Computers, Environment and Urban Systems*, Vol. 54, pp. 195–203.

Fagnant, D. J., and Kockelman, K. (2016). A Direct-Demand Model for Bicycle Counts: The Impacts of Level of Service and Other Factors. *Environment and Planning B: Planning and Design,* Vol. 43, No. 1, pp. 93–107.

Gosse, C., and Clarens, A. (2014). Estimating Spatially and Temporally Continuous Bicycle Volumes by Using Sparse Data. *Transportation Research Record*, No. 2443, pp. 115–122.

Griffin, G. P., and Jiao, J. (2015). Where Does Bicycling for Health Happen? Analysing Volunteered Geographic Information through Place and Plexus. *Journal of Transport and Health*, Vol. 2, No. 2, pp. 238–247.

Griswold, J., Medury, A., and Schneider, R. (2011). Pilot Models for Estimating Bicycle Intersection Volumes. *Transportation Research Record*, No. 2247, pp. 1–7.

Handy, S., and McCann, B. (2010). *The Regional Response to Federal Funding for Bicycle and Pedestrian Projects*. University of California, Davis, Institute of Transportation Studies.

Hankey, S., and Lindsey, G. (2016). Facility-Demand Models of Peak Period Pedestrian and Bicycle Traffic: Comparison of Fully Specified and Reduced-Form Models. *Transportation Research Record*, No. 2586, pp. 48–58.

Hankey, S., Lindsey, G., Wang, X., Borah, J., Hoff, K., Utecht, B., and Xu, Z. (2012). Estimating Use of Nonmotorized Infrastructure: Models of Bicycle and Pedestrian Traffic in Minneapolis, MN. *Landscape and Urban Planning*, Vol. 107, No. 3, pp. 307–316.

Hankey, S., Lu, T., Mondschein, A., and Buehler, R. (2017). *Merging Traffic Monitoring and Direct-Demand Modeling to Assess Spatial Patterns of Annual Average Daily Bicycle and Pedestrian Traffic.* Transportation Research Board 2017 Annual Meeting.

Hedlund, J. (2000). *NHTSA/FHWA Pedestrian and Bicycle Strategic Planning Research Workshops.* Workshop conducted by the Federal Highway Administration and National Highway Traffic Safety Administration.

Hillier, B. (1996). *Space Is the Machine: A Configurational Theory of Architecture*. Cambridge University Press, Cambridge, UK.

Hunter, W., and Huang, H. (1995). User Counts on Bicycle Lanes and Multiuse Trails in the United States. *Transportation Research Record*, No. 1502, pp. 45–57.

Jones, M. G., Ryan, S., Donlon, J., Ledbetter, L., Ragland, D. R., and Arnold, L. S. (2010). *Seamless Travel: Measuring Bicycle and Pedestrian Activity in San Diego County and Its Relationship to Land Use, Transportation, Safety, and Facility Type.* University of California, Berkeley, Safe Transportation Research and Education Center.

Katz, R. (2003). *Forecasting Demand for Bicycle Facilities*. Australian Road Research Board 21st Transport Research Conference.

Kim, N., and Susilo, N. (2011). Comparison of Pedestrian Trip Generation Models. *Journal of Advanced Transportation,* Vol. 47, No. 4, pp. 399–412.

Kitamura, R., Mokhtarian, P., and Laidet, L. (1997). A Micro-analysis of Land Use and Travel in Five Neighborhoods in the San Francisco Bay Area. *Transportation*, Vol. 2, No. 2, pp. 125–158.

Kraft, G. (1963). *Demand for Intercity Passenger Travel in the Washington-Boston Corridor*. Part V, Northeast Corridor Project, U.S. Department of Transportation, Washington, DC.

Krizek, K. (2006). *Guidelines for Analysis of Investments in Bicycle Facilities*. National Cooperative Highway Research Program, Transportation Research Board, Washington, DC.

Kuzmyak, J. R., Walters, J., Bradley, M., and Kockelman, K. M. (2014). *Estimating Bicycling and Walking for Planning and Project Development: A Guidebook.* Transportation Research Board, Washington, DC.

Landis, B., Vattikuti, V., Ottenberg, R., McLeod, D., and Guttenplan, M. (2001). Modeling the Roadside Walking Environment: Pedestrian Level of Service. *Transportation Research Record*, No. 1773, pp. 82–88.

Levinson, D., Boies, A., Cao, J., and Fan, Y. (2016). *The Transportation Futures Project: Planning for Technology Change.* Minnesota Department of Transportation, St. Paul, MN.

Lindsey, G. H. (2011). *Forecasting Use of Nonmotorized Infrastructure: Models of Bicycle and Pedestrian Traffic in Minneapolis, Minnesota.* Transportation Research Board 90th Annual Meeting.

Lindsey, G., Han, Y., Wilson, J., and Yang, J. (2006). Neighborhood Correlates of Urban Trail Use. *Journal of Physical Activity and Health*, Vol. 3, pp. S139–S157.

Lindsey, G., Hoff, K., Hankey, S., and Wang, X. (2012). *Understanding the Use of Nonmotorized Transportation Facilities.* University of Minnesota, Center for Transportation Studies, Intelligent Transportation Systems Institute, Minneapolis, MN.

Lindsey, G., Wilson, J., Rubchinskaya, E., Yang, J., and Han, Y. (2007). Estimating Urban Trail Traffic: Methods for Existing and Proposed Trails. *Landscape and Urban Planning*, Vol. 81, No. 4, pp. 299–315.

Liu, X., and Griswold, J. (2009). Pedestrian Volume Modeling: A Case Study of San Francisco. *Association of Pacific Coast Geographers Yearbook*, Vol. 71, No. 1, pp. 164–181.

McKenzie, B. (2014). *Modes Less Traveled—Bicycling and Walking to Work in the United States: 2008–2012 American Community Survey Reports.* U.S. Census Bureau, Washington, D.C.

Miranda-Moreno, L., and Fernandes, D. (2011). Modeling of Pedestrian Activity at Signalized Intersections: Land Use, Urban Form, Weather, and Spatiotemporal Patterns. *Transportation Research Record*, No. 2264, pp. 74–82.

Misra, A., Gooze, A., Watkins, K., Asad, M., and Le Dantec, C. (2014). Crowdsourcing and Its Application to Transportation Data Collection and Management. *Transportation Research Record*, No. 2414, pp. 1–8.

Molino, J., Kennedy, J., Johnson, P., Beuse, P., Emo, A., and Do, A. (2009). Pedestrian and Bicyclist Exposure to Risk: Methodology for Estimation in an Urban Environment. *Transportation Research Record*, No. 2140, pp. 145–156.

Moudon, A., Hess, P., Snyder, M., and Stanilov, K. (1997). Effects of Site Design on Pedestrian Travel in Mixed-Use, Medium-Density Environments. *Transportation Research Record*, No. 1578, pp. 48–55.

Moudon, A. V., Lee, C., Cheadle, A. D., Collier, C. W., Johnson, D., Schmid, T. L., and Weather, R. D. (2005). Cycling and the Built Environment, a US Perspective. *Transportation Research Part D: Transport and Environment*, Vol. 10, No. 3, pp. 245–261.

National Highway Traffic Safety Administration (2016). *2015 Motor Vehicle Crashes: Overview.* U.S. Department of Transportation, Washington, DC.

Nordback, K. (2012). *Estimating Annual Average Daily Bicyclists and Analyzing Cyclist Safety at Urban Intersections.* Doctoral dissertation, University of Colorado at Denver.

Nordback, K., Kothuri, S., Figliozzi, M., Phillips, T., Gorecki, C., and Schrope, A. (2016). *Investigation of Bicycle and Pedestrian Continuous and Short Duration Count Technologies in Oregon.* Oregon Department of Transportation, Salem, OR.

Ortuzar, J. d. D. and Willumsen, L. G. (2011). *Modelling Transport*. Fourth Edition, John Wiley & Sons, Ltd, Chichester, UK.

O'Toole, K., and Piper, S. (2017). *Innovation in Bicycle and Pedestrian Counts: A Review of Emerging Technology*. White Paper. Alta Planning and Design, Charlotte, NC.

Pikora, T., Giles-Corti, B., Bull, F., Jamrozik, K., and Donovan, R. (2003). Developing a Framework for Assessment of the Environmental Determinants of Walking and Cycling. *Social Science and Medicine*, Vol. 56, No. 8, pp. 1693–1703.

Porter, C., Suhrbier, J., and Suhrbier, W. (1999). Forecasting Bicycle and Pedestrian Travel: State of the Practice and Research Needs. *Transportation Research Record*, No. 1674, pp. 94–101.

Pulugurtha, S., and Repaka, S. (2008). Assessment of Models to Measure Pedestrian Activity at Signalized Intersections. *Transportation Research Record*, No. 2073, pp. 39–48.

Pushkarev, B., and Zupan, J. (1971). Pedestrian Travel Demand. *Transportation Research Record*, No. 377, pp. 37–53.

Qin, X., and Ivan, J. (2001). Estimating Pedestrian Exposure Prediction Model in Rural Areas. *Transportation Research Record*, No. 1773, pp. 89–96.

Ranaiefar, F., Chow, J. Y., McNally, M. G., and Ritchie, S. G. (2014). *A Structural Direct-Demand Model for Inter-Regional Commodity Flow Forecasting*. Transportation Research Board 93rd Annual Meeting.

Schepers, P. (2012). Does More Cycling Also Reduce the Risk of Single-Bicycle Crashes? *Injury Prevention,* Vol. 18, No. 4, pp. 240–245.

Schneider, R. J., Arnold, L. S., and Ragland, D. R. (2009a). Methodology for Counting Pedestrians at Intersections: Use of Automated Counters to Extrapolate Weekly Volumes from Short Manual Counts. *Transportation Research Record*, No. 2140, pp. 1–12.

Schneider, R. J., Arnold, L. S., and Ragland, D. R. (2009b). A Pilot Model for Estimating Pedestrian Intersection Crossing Volumes. *Transportation Research Record*, No. 2140, pp. 13–26.

Schneider, R., Henry, T., Mitman, M., Stonehill, L., and Koehler, J. (2012). Development and Application of a Pedestrian Volume Model in San Francisco, California. *Transportation Research Record*, No. 2299, pp. 65–78.

Schwartz, W. L., Porter, C. D., Payne, G. C., Suhrbier, J. H., Moe, P. C., and Wilkinson III, W. L. (1999). *Guidebook on Methods to Estimate Nonmotorized Travel: Overview of Methods.* FHWA-RD-98-165. Research, Development, and Technology Turner-Fairbank Highway Research Center, McLean, VA.

Sjafruddin, A., Frazils, R. B., and Astuti, R. D. (1999). Regional Freight Transport Demand Modeling in the Java Island. *Journal of the Eastern Asia Society for Transportation Studies*, Vol. 3, No. 3, pp. 303–313.

Stinson, M. A., and Bhat, C. R. (2004). Frequency of Bicycle Commuting: Internet-Based Survey Analysis. *Transportation Research Record*, No. 1878, pp. 122–130.

Strauss, J., and Miranda-Moreno, L. (2013). Spatial Modeling of Bicycle Activity at Signalized Intersections. *The Journal of Transport and Land Use*, Vol. 6, No. 2, pp. 47–58.

Strauss, J., Miranda-Moreno, L. F., and Morency, P. (2013). Cyclist Activity and Injury Risk Analysis at Signalized Intersections: A Bayesian Modelling Approach. *Accident Analysis and Prevention*, Vol. 59, pp. 9–17.

Strauss, J., Miranda-Moreno, L. F., and Morency, P. (2015). Mapping Cyclist Activity and Injury Risk in a Network Combining Smartphone GPS Data and Bicycle Counts. *Accident Analysis and Prevention*, Vol. 83, pp. 132–142.

Tabeshian, M., and Kattan, L. (2014). Modeling Nonmotorized Travel Demand at Intersections in Calgary, Canada: Use of Traffic Counts and Geographic. *Transportation Research Record*, No. 2430, pp. 38–46.

Timberlake, R. (1988). Traffic Modeling Techniques for the Developing World: Case Studies. *Transportation Research Record*, No. 1167, pp. 28–34.

Turner, S., Sener, I. N., Martin, M., Das, S., Shipp, E., Hampshire, R., Fitzpatrick, K., Molnar, L., Wijesundera, R., Colety, M., and Robinson, S. (2017). *Synthesis of Methods for Estimating Pedestrian and Bicyclist Exposure to Risk at Areawide Levels and on Specific Transportation Facilities.* U.S. Department of Transportation, Federal Highway Administration, Office of Safety, Washington, DC.

Wang, J., Hankey, S., Wu, S., and Lindsey, G. (2016). Monitoring and Modeling of Urban Trail Traffic: Validation of Direct-Demand Models in Minneapolis, Minnesota, and Columbus, Ohio. *Transportation Research Record*, No. 2593, pp. 47–59.

Wang, X., Lindsey, G., Hankey, S., and Hoff, K. (2013). Estimating Mixed-Mode Urban Trail Traffic Using Negative Binomial Regression Models. *Journal of Urban Planning and Development,* Vol. 140, No. 1. http://dx.doi.org/10.1061/(ASCE)UP.1943-5444.0000157.

Wardman, M. (1997). Inter-urban Rail Demand, Elasticities and Competition in Great Britain: Evidence from Direct-Demand Models. *Transportation Research Part E: Logistics and Transportation Review*, Vol. 33, No. 1, pp. 15–28.

Wardman, M., Whelan, G. A., and Toner, J. P. (1994). *Direct-Demand Models of Air Travel: A Novel Approach to the Analysis of Stated Preference Data*. Working Paper. Institute of Transport Studies, University of Leeds, Leeds, UK.

Wendel-Vos, G. C., Schuit, A. J., De Niet, R., Boshuizen, H. C., Saris, W., and Kromhout, D. A. A. N. (2004). Factors of the Physical Environment Associated with Walking and Bicycling. *Medicine and Science in Sports and Exercise*, Vol. 36, No. 4, pp. 725–730.

Winters, M., Brauer, M., Setton, E. M., and Teschke, K. (2010). Built Environment Influences on Healthy Transportation Choices: Bicycling Versus Driving. *Journal of Urban health*, Vol. 87, No. 6, pp. 969–993.