Routledge
Taylor & Francis Group

Transport Reviews

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ttrv20

A systematic scoping review of methods for estimating link-level bicycling volumes

Debjit Bhowmick, Meead Saberi, Mark Stevenson, Jason Thompson, Meghan Winters, Trisalyn Nelson, Simone Zarpelon Leao, Sachith Seneviratne, Christopher Pettit, Hai L. Vu, Kerry Nice & Ben Beck

TRANSPORT REVIEWS
https://doi.org/10.1080/01441647.2022.2147240

Routledge
Taylor & Francis Group

Check for updates

A systematic scoping review of methods for estimating link-level bicycling volumes

Debjit Bhowmick[a], Meead Saberi [b], Mark Stevenson[c,d], Jason Thompson[c], Meghan Winters [e], Trisalyn Nelson [f], Simone Zarpelon Leao[g], Sachith Seneviratne[c], Christopher Pettit [h], Hai L. Vu [i], Kerry Nice[c] and Ben Beck[a]

[a]School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia; [b]School of Civil and Environmental Engineering, Research Centre for Integrated Transport Innovation (rCITI), University of New South Wales, Sydney, Australia; [c]Transport, Health and Urban Design Research Lab, Melbourne School of Design, The University of Melbourne, Parkville, Australia; [d]Faculty of Engineering and Information Technology, The University of Melbourne, Parkville, Australia; [e]Faculty of Health Sciences, Simon Fraser University, Vancouver, Canada; [f]Department of Geography, University of California, Santa Barbara, CA, USA; [g]School of Built Environment, Faculty of Arts, Design and Architecture, University of New South Wales, Sydney, Australia; [h]City Futures Research Centre, School of Built Environment, University of New South Wales, Sydney, Australia; [i]Institute of Transport Studies, Monash University, Clayton, Australia

ABSTRACT

Estimation of bicycling volumes is essential for the strategic implementation of infrastructure and related transport elements and policies. Link-level volume estimation models (models that estimate volumes on individual street segments) allow for understanding variation in bicycling volumes across an entire network at higher spatial resolution than area-level models. Such models assist transport planners to efficiently monitor network usage, to identify opportunities to enhance safety and to evaluate the impact of policy and infrastructure interventions. However, given the sparsity and scarcity of bicycling data as compared to its motorised counterparts, link-level bicycling volume estimation literature is relatively limited. This paper conducts a scoping review of link-level bicycling volume estimation methods by implementing systematic search strategies across relevant databases, thereby identifying appropriate studies for the review. The review resulted in some interesting findings. Among all the methods implemented, direct demand modelling was the predominant one. Not a single study implemented multiple modelling approaches in the same study area, thereby not allowing for comparison of these approaches. Most studies were conducted in the United States. It was also observed that there exists a lot of heterogeneity in the reporting of basic study characteristics and validation results, sometimes to the extent of not reporting these at all. The study presents the different types of data used in modelling (count, travel survey, GPS data) along with an array of popular explanatory variables that can inform future studies about data collection and variable selection for modelling. The study discusses the strengths and limitations of different methods and finally presents recommendations for future research.

ARTICLE HISTORY
Received 23 June 2022
Accepted 2 November 2022

KEYWORDS
Bicycling; link-level; volume; modelling; estimation; review

CONTACT Debjit Bhowmick debjit.bhowmick@monash.edu School of Public Health and Preventive Medicine, Monash University, Level 3, 553 St Kilda Road, Melbourne, VIC 3004, Australia

Supplemental data for this article can be accessed online at https://doi.org/10.1080/01441647.2022.2147240.

1. Introduction

1.1. Background

Bicycling is an active and sustainable mode of transportation that offers an array of public health and environmental benefits. At an individual level, bicycling improves the physical health of the rider via physical exercise (De Geus et al., 2007; Wen & Rissel, 2008). At a community scale, more bicycling leads to reduced traffic congestion thereby minimising air pollution, therefore improving public health in the longer term (De Hartog et al., 2010; Grabow et al., 2012; Lindsay et al., 2011). To enable planners and decision-makers to develop targeted investment strategies for improving the uptake of bicycling, data on patterns of bicycling across the entire area of interest namely, at the municipality-wide scale or city-wide scale, is needed. However, collecting such spatially comprehensive data requires frequent implementation of costly data collection infrastructure, such as dense placement of sensors, running computationally expensive vehicle classifier programmes, and traditional observational surveys, and is therefore not convenient. To fill this gap, transport models that can estimate relevant traffic parameters, such as traffic volumes, are developed (ATAP, 2022; Kraft & Wohl, 1967).

Researchers and planners develop travel demand models (the most common form of transport models) to understand traffic volumes (motor vehicles or other modes such as bicycles or pedestrians) within a geographic area. For example, Mohamad et al. (1998) developed a model to predict traffic volumes across 40 counties in Indiana, U.S., Zhao and Chung (2001) estimated annual average daily traffic in a Florida county, and Apronti and Ksaibati (2018) estimated traffic volumes on rural low-volume roads in Wyoming, U.S., while Cooper (2017) developed link-level models for active travel in Cardiff, Wales. These travel demand models attempt to replicate ground conditions of existing traffic flow, such as traffic volume along street segments (also referred to as links). Validated models are then used to estimate changes in flow dynamics due to changes in parameter settings, estimate volumes at locations without any known volume information in the same study area (spatial transferability), or predict future traffic volumes at locations with known volume information across multiple time points using time-series methods (temporal transferability). Spatial and temporal models can work together as well if a model is transferred over both space and time (Fox & Hess, 2010). Therefore, travel demand models help relevant authorities in monitoring, evaluating and in strategic implementation of policy, program, or infrastructural interventions by either estimating current demand at locations without demand information, or forecast future demand at locations with demand information across multiple time points, or both. Demand can be specific to one mode of travel, or be multi-modal. Travel demand models can cater to small, large, and multiple geographical areas. To support pro-bicycling policies, decision makers often need access to bicycling volume data to determine patterns of bicycling, create denominators to measure bicycling safety, and evaluate the impact of policies and infrastructure interventions aimed to increase bicycling. Therefore, travel demand models that focus on modelling bicycling volumes are important. The focus of this review is to conduct a systematic review of studies that have developed models to estimate bicycling volumes at link-level.

1.2. Motivation

Travel demand models are primarily data-driven and their performance heavily depends on the data they use. However, bicycling-related data is sparse and lacks comprehensive coverage spatially and temporally, as compared to its motorised counterparts (DiGioia et al., 2017; Roy et al., 2019; Winters & Branion-Calles, 2017). Crowdsourced bicycling data has filled some of the voids by leveraging current advances in smartphone technology (Jestico et al., 2016; Kwigizile et al., 2019) such as providing data at a high spatial and temporal resolution, offsetting limitations related to spatial coverage of traditional datasets, and real-time monitoring of mobility. However, the disparity in quantity between bicycling data and motor vehicle data is still prominent. This limits the scope, resolution and representativeness of bicycling-related models. Hence, traditionally, transportation planning has been more focused on motorised vehicles, resulting in far greater numbers of associated research efforts. In contrast, bicycling-specific demand modelling (especially link-level volume modelling) efforts are far fewer in number and less mature, with authorities gradually, only recently, making more serious investments in relation to active and sustainable mobility.

Moreover, the mobility patterns of cyclists differ from motorists as the mobility behaviour of cyclists depends on a host of factors such as available infrastructure, safety and safety perception, weather and other disaggregated factors, which are not significant drivers of mobility behaviour of motor vehicle users (Dill & Gliebe, 2008). For example, motor vehicle drivers usually opt for the fastest routes as they tend to prefer highways with greater posted speed limits over local roads (Winters et al., 2010). However, whilst trip time and route length are important factors for bicyclists as well, choice of route is also influenced by safety and the availability of bike-related infrastructure, such as preference for lower-volume local roads and off-road bike paths separated from motorised traffic (Stinson & Bhat, 2003; Tilahun et al., 2007; Winters et al., 2010). Area-level transport demand models aggregate over large geographic areas which are divided into large analysis zones. However, bicycling trips are usually, significantly shorter in distance compared to car trips, and therefore, bicycling trips need to be captured at a greater spatial resolution, such as at link-level (Liu et al., 2012). Models which are based on the mobility behaviour of motor vehicle users, or aggregate over large traffic analysis zones, are not appropriate for modelling bicycling in their current state (McDaniel et al., 2014). This is because bicycling volume estimates for aggregated zones are not appropriate for strategic implementation of additional infrastructure, or identification of the impacts of such changes across the streets in the network. Given the mobility behaviour of bicyclists, link-level bicycling volume estimates are essential for planners to understand the dynamics of bicycling flows at the finest spatial resolution. While disaggregated travel demand models are data intensive, computationally expensive and complex, such models are gaining prominence with improvement in computational capacities and the availability of more high-resolution data (McDaniel et al., 2014; Wang et al., 2011).

Furthermore, we have limited the scope of our review to studies that developed demand models that can estimate bicycling volumes at locations without any volume information. The scope of this review excludes studies that had implemented time-series models to forecast bicycling volumes only at locations where current and past volume information is known. Our aim is to review studies which develop a model using bicycling volumes at known locations and consequently expand this to the rest of the study network producing estimates at locations where there is no volume data.

1.3. Objectives

There have been limited research efforts for modelling link-level bicycling volumes that overcome the challenges of the traditional travel demand models. However, to the best of our knowledge, no comprehensive review of such modelling efforts exists in the current literature. In this paper, we present a scoping review of studies that have estimated link-level bicycling volumes, and highlight the strengths and limitations of the modelling approaches.

The first set of objectives of this review are to present:

(a) The reported study characteristics
(b) The approaches taken for modelling link-level bicycling volumes
(c) The types of data used for modelling link-level bicycling volumes
(d) Frequently used variables in models
(e) The reported modelling accuracies and validation methods.

Furthermore, the review intends to:

(a) Discuss the variation in reporting of study characteristics, modelling accuracies and validation methods
(b) Discuss the strengths and limitations of the major modelling approaches
(c) Make recommendations for future studies.

1.4. Structure

The rest of the paper is structured as follows. Section 2 briefly discusses the predominant travel demand modelling approaches in the literature. Section 3 presents the systematic approach undertaken to search, screen and select the studies relevant for this review. Section 4 presents important information related to the studies included in this review such as year of publication, location of the study, types of data used, modelling approach and accuracy measures, discusses the findings, highlights the strengths and limitations of the different methods, and makes recommendations for future research. Section 5 concludes the paper by presenting a summary of all the previous sections.

2. Types of travel demand modelling approaches

There are multiple approaches to modelling traffic volume and/or travel demand that exist in the literature. Researchers and practitioners predominantly use (a) four-step demand modelling, (b) direct demand modelling and (c) agent-based modelling. These approaches are presented in the order of their development/first usage to showcase the historical evolution of different modelling approaches. There are some alternative methods which are presented as well. These are briefly discussed as follows.

2.1. Traditional four-step demand modelling

Conventionally, travel demand modelling is a four-step process. (i) Trip generation – number of trips generated by each travel analysis zone (origins). (ii) Trip distribution – share of generated trips from each travel analysis zone (origins) distributed to same/other travel analysis zones (destinations). (iii) Mode choice – share of these trips distributed among existing travel modes. (iv) Trip assignment – assigning trips to specific routes. These four-step demand models are key to developing long-term city-wide strategic frameworks and therefore, have been extensively used by city planning authorities. However, four-step models are often associated with greater errors as a result of error accumulation across multiple steps (Choi et al., 2012).

2.2. Direct demand modelling

The primary alternative to the four-step model is the direct demand modelling approach, which bypasses some of the limitations of sequential four-step models and potentially increases accuracy. Direct demand models (DDMs) predict traffic volumes by combining all four steps of the conventional demand model into one step (McFadden, 1974) and therefore, do away with cumulative errors across multiple steps. Direct demand models directly associate known traffic volumes with relevant attributes (de Dios Ortúzar & Willumsen, 2011). They employ statistical and machine learning models to first fit the model to known volumes using surrounding road network, socio-demographic, land-use and other relevant information which significantly affect these volumes, and then estimate volumes at locations (where volume information is unavailable) using the same attributes (Anderson et al., 2006; Cooper, 2016, 2017, 2018; McDaniel et al., 2014). However, DDMs do not account for either travel behaviour and perception of travellers, or the interaction between travellers and between travellers and their environment.

2.3. Simulation-based demand modelling

Simulation-based models have gained popularity in the travel demand modelling space with increases in computational capacities (Turner et al., 2017). Microsimulation and agent-based simulation models provide an appropriate simulation environment to produce spatially and temporally explicit traffic flows. Agent-based models (ABMs) simulate actions of and interactions between agents (individual elements in the model) that act individually based on their characteristic attributes and behaviour (Wilensky & Rand, 2015). In a transportation-focused ABM, agents might represent heterogenous individuals who make travel decisions based on their assigned unique characteristics and related behavioural traits. ABMs are considered a suitable approach when recognising the heterogeneity of cyclists with individual demands, behaviour patterns, and interpersonal interactions (Kaziyeva et al., 2021). ABMs can reproduce the travel behaviour of real individuals, and therefore aim to imitate the real-world heterogeneity of persons and external conditions (Heppenstall et al., 2011). This results in generation of traffic flows and mobility patterns of the population at a very fine spatial and temporal resolution (Leao & Pettit, 2016). However, evidence-based behaviour assignment of agents is critical for developing a representative and useful ABM.

2.4. Other modelling approaches

Spatial network analysis models are variations of DDMs that use the concepts of graph theory, space syntax and spatial network analysis, with network measures such as centrality measures and intersection density to estimate traffic volumes across the network (Cooper, 2016, 2017, 2018; McDaniel et al., 2014; Raford & Ragland, 2004, 2006). However, in comparison to assignment-free DDMs, spatial network analysis models incorporate the trip assignment process. Therefore, the built environment and demographic variables serve as independent variables to formulate trip assignment, which in turn, acts as the governing variable to estimate travel demand (Cooper et al., 2021). *GIS-based models* rely heavily on GIS (Geographic Information System) to depict relevant land-use and infrastructure configurations and thereby estimate their effects on trip generation and modal split, consequently estimating travel demand (often focused for specific modes, such as walking) (Turner et al., 2017). They are useful for scenario planning, such as comparison of the impacts of different configurations of relevant land-use and infrastructure on travel demand (Kuzmyak et al., 2014).

While we presented a brief overview of these modelling approaches in this section, it is only through this systematic review, thorough synthesisation and critique of the reviewed studies and other related literature, that we can robustly compare the strengths and limitations of these modelling approaches. These are presented later in Section 4.6.

3. Method

To address the objectives mentioned in section 1.3, we conducted a systematic literature review on studies that estimated link-level bicycling volumes. We developed a comprehensive search strategy based on similar transport and bicycling reviews, content matter expertise. We reviewed multiple articles that estimated link-level bicycling volumes to identify relevant keywords. We employed our search strategy on multiple databases. There were no restrictions placed on the type of article or the year of publication. However, only articles published in English were included.

3.1. Study eligibility

Studies were included if they met the following criteria:

- Modelled link-level bicycling volumes separately (study has to include a separate bicycling volume model).
- Conducted a primary study involving their own data collection and modelling.
- Published in English.

Studies were excluded if:

- The spatial granularity of the models developed were coarser than link-level, such as, a model that produces volume estimates at an area-level.

- They developed a time-series model to forecast future link-level bicycling volumes only at locations with known volumes, and do not predict volumes on links without any count information and therefore, the rest of the study network.
- They developed mixed-mode models, where volumes of transport modes other than bicycles are grouped together in the same model, such as a combined volume model for pedestrians and bicyclists.
- They were related to bike-sharing programmes.
- Their full-text was not accessible.

3.2. Search strategy

The search was conducted in December 2021 using Scopus, Compendex, Inspec, GEOBASE, GeoRef, U.S. Patents, EP Patents & WO Patents, and TRID (Transport Research International Documentation). Scopus is the largest abstract and citation database of peer-reviewed literature (Abduljabbar et al., 2022; Biswas et al., 2021; Pritchard, 2018), while "Ei Compendex is the broadest and most complete engineering literature database available in the world" (Bian et al., 2022; Biswas et al., 2021; Elsevier, 2022a). Inspec is another comprehensive engineering dataset (Elsevier, 2022c), while TRID "focuses on transportation research and contains more than 1.1 million records worldwide and is maintained by the Transportation Research Board of the U.S. National Academies" (Bian et al., 2022; Kamaluddin et al., 2018; Pritchard, 2018). While a search of the four databases mentioned above is expected to cover all relevant transport related literature, the other databases (also mentioned above) were searched anyway as they are included on the Engineering Village, a search and discovery platform (Elsevier, 2022b), with Compendex and Inspec.

We used free-text keywords for all the databases with index/MeSH terms relevant to the database. Major search concepts were bicycling/bicyclists, demand/volume modelling (excluding bike-sharing and related terms). A detailed description of the systematic search strategy (conducted by the primary author) is provided in the supplementary material. The queries listed in the supplementary material were implemented in the respective databases and then imported in Covidence, "A web-based collaboration software platform that streamlines the production of systematic and other literature reviews". The references of the imported articles were stored in Covidence as it facilitated organised screening (abstract screening and full-text screening) (Covidence, 2022). Screening was conducted solely by the primary author.

3.3. Title, abstract screening and full-text review

All the identified and unique articles underwent title and abstract screening. All articles that passed the title and abstract screening were reviewed by studying their full-texts. Studies whose full-texts could not be retrieved were excluded, and so were duplicate studies and studies that developed only mixed-mode models. The reference lists of the full-texts were also screened for any additional studies that were not returned by the search, but fitted into the scope of this review.

3.4. Data extraction

We extracted basic study characteristics such as title, year of publication, primary author's name, source where the study was published (journal/conference), and geographic location of the study. Next, we recorded information related to the data and methods employed by the study which included the types and quantity of data used, modelling approach (e.g. DDM, ABM), type of models (subclass of model employed, i.e. type of statistical or machine learning model such as negative binomial regression, support vector machines), and independent variables used in the model. Finally, we extracted information on the results of the studies, such as reported accuracy and error of the model, and whether model validation was performed.

4. Results and discussion

Our search strategy produced 6865 studies for screening, out of which 2592 were identified as duplicates by Covidence, and were therefore excluded. Title and abstract screening were conducted on the remaining 4273 studies, which led to the exclusion of 4135 irrelevant studies. Full-texts of the remaining 138 studies were reviewed out of which 98 studies were excluded. Reasons for exclusion included studies not estimating link-level bicycling volumes ($n = 56$), temporal-only forecasting using time-series models ($n = 17$), mixed mode models which estimate collective volumes of more than one transport mode ($n = 2$), non-primary review studies ($n = 3$), duplicate studies ($n = 4$), irrelevant study design and outcomes ($n = 8$) and full texts not found ($n = 8$). One, additional relevant study was identified from the list of references of another study while reviewing its full text, leading to a final set of 41 studies within the scope of this review. This search process is illustrated in Figure 1. In the following subsections, we present the information extracted from the studies in detail.

4.1. Basic study characteristics

Forty-one studies were included for the final review. The publication year of these studies range from 2010 to 2021. Thirty-one studies were conducted in single cities while 10 studies were conducted in multiple cities, sometimes across multiple countries ($n = 2$). The majority of studies were conducted in North America ($n = 30$; $n = 26$ in the United States of America), with others in Europe ($n = 9$), Oceania ($n = 2$) and South America ($n = 1$). There is a clear geographical bias when it comes to the location of the studies. The 26 studies conducted in the U.S. reflected models in 55 unique cities. This is in contrast to other study locations that reflected models from 16 cities. The greater number of U.S.-based models may be explained by greater availability of bicycling data and data collection infrastructure in the U.S. as compared to the rest of the world. Only one study was conducted in a developing nation (Arellana et al., 2020). With the increased uptake of bicycles around the globe, municipal authorities should be promoting enhanced bicycling data collection in their respective areas.

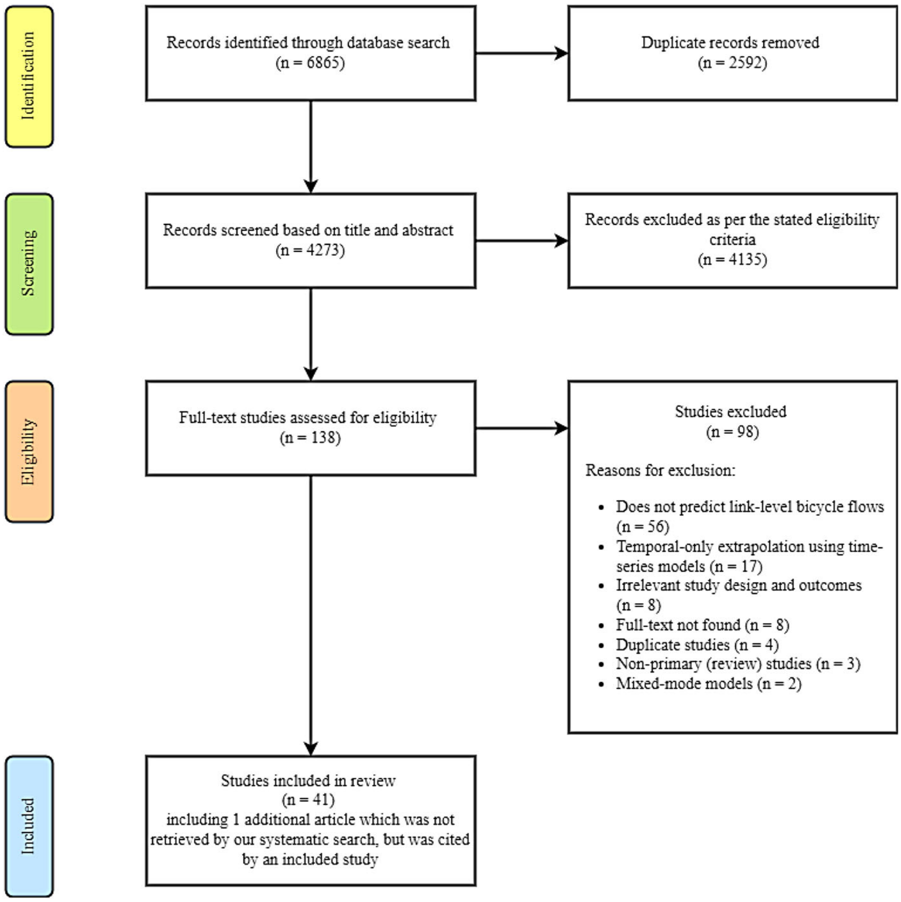

Figure 1. Study selection flowchart.

4.2. Types of modelling approaches

Among the modelling approaches undertaken by the authors of the studies, DDM was the most predominant ($n = 29$). Mixed methods, using a combination of Spatial Network Analysis and DDM, were used in four studies. The traditional four-step demand modelling procedure was employed by three studies, while ABM, and two-stage bicycle origin and destination (OD) matrix estimation process was employed by two studies each. Finally, there was one study that employed microsimulation. The basic study characteristics and the corresponding modelling approaches have been detailed in Table 1.

4.3. Types of data (and variables) used in the models

The most popular type of data used was bicycle count data (short-term and long-term counts of bicycles obtained using automated sensors or a manual counting process) ($n = 39$), followed by road infrastructure data (presence of protected bicycle lanes, density of bicycling facility present, presence of bike signals, functional class of road, pavement width, number of lanes in the carriageway, presence of on-road parking lanes, posted

Table 1. Study characteristics.

Author (year)	Study area	Countries	Number of count locations	Modelling approach
Arellana et al. (2020)	Barranquilla	Colombia	27 locations	Direct demand modelling
Bargh et al. (2012)	Hastings and Havelock North	New Zealand		Microsimulation
Cooper (2016)	Cardiff	U.K.	107 on-road locations, 14 traffic-free paths	Mixed method – Spatial Network Analysis + Direct demand modelling
Cooper (2017)	Cardiff	U.K.	107 on-road locations, 14 traffic-free paths	Mixed method – Spatial Network Analysis + Direct demand modelling
Cooper (2018)	Cardiff	U.K.	107 on-road locations, 14 traffic-free paths	Mixed method – Spatial Network Analysis + Direct demand modelling
Dadashova et al. (2020)	Austin, Brownsville, Corpus Christi, Dallas, Houston, League City, Lubbock, Midland, Odessa, Plano, San Antonio, and Wichita Falls (All TX)	U.S.A.	100 locations across 12 cities	Direct demand modelling
Dadashova and Griffin (2020)	Austin, Brownsville, Corpus Christi, Dallas, Houston, League City, Lubbock, Midland, Odessa, Plano, San Antonio, and Wichita Falls (All TX)	U.S.A.	350 locations in 12 cities	Direct demand modelling
Ermagun et al. (2018)	Portland, ME; Arlington, VA; Miami, FL; New Orleans, LA; Minneapolis, MN; Duluth, MN; Fort Worth, TX; Houston, TX; Albuquerque, NM; Colorado Springs, CO; Billings, MT; Seattle, WA; San Diego, CA	U.S.A.	32 locations in 13 cities	Direct demand modelling
Fagnant and Kockelman (2016)	Seattle, WA	U.S.A.	251 locations	Direct demand modelling
Fan and Lin (2019)	Charlotte, NC	U.S.A.	7 locations	Direct demand modelling
Gehrke and Reardon (2021)	Cambridge, MA	U.S.A.	91 manual counts across 19 intersections	Direct demand modelling
Gosse and Clarens (2014)	Charlottesville, VA	U.S.A.	18 locations	Temporal factoring using Markov chain Monte Carlo (MCMC) sampling, and Spatial factoring
Hankey et al. (2012)	Minneapolis, MN	U.S.A.	259 locations	Direct demand modelling
Hankey et al. (2021)	Blacksburg, VA; Boston, MA; Champaign Urbana, IL; Cleveland, OH; Columbus, OH; Denver, CO; Hartford, CT; Lawrence, KS; Los Angeles, CA; Madison, WI; Manhattan, KS; Minneapolis, MN; New York City, NY; Philadelphia, PA; Portland, OR; San Francisco, CA; Seattle, WA; St Louis, MO; Tucson, AZ; Washington, DC	U.S.A.	4145 locations across 20 cities	Direct demand modelling

(*Continued*)

Table 1. Continued.

Author (year)	Study area	Countries	Number of count locations	Modelling approach
Haworth (2016)	London	U.K.	298 locations	Direct demand modelling
Hochmair et al. (2019)	Miami-Dade County, FL	U.S.A.	32 locations	Direct demand modelling
Jacyna et al. (2017)	Warsaw	Poland		Four-step modelling procedure
Jahangiri et al. (2019)	San Diego, CA	U.S.A.	88 locations	Direct demand modelling
Jestico et al. (2016)	Victoria, BC	Canada	18 locations	Direct demand modelling
Jones et al. (2010)	San Diego County, CA	U.S.A.	80 locations	Direct demand modelling
Kaziyeva et al. (2021)	Salzburg	Austria	9 locations	Agent-based simulation modelling
Kwigizile et al. (2019)	Ann Arbor, MI; Grand Rapids, MI	U.S.A.	19 locations	Direct demand modelling
Le et al. (2017)	Blacksburg, VA; Boston, MA; Champaign Urbana, IL; Cleveland, OH; Columbus, OH; Denver, CO; Hartford, CT; Lawrence, KS; Los Angeles, CA; Madison, WI; Manhattan, KS; Minneapolis, MN; New York City, NY; Philadelphia, PA; Portland, OR; San Francisco, CA; Seattle, WA; St Louis, MO; Tucson, AZ; Washington, DC	U.S.A.	9870 locations	Direct demand modelling
Lin and Fan (2020a)	Charlotte, NC	U.S.A.		Direct demand modelling
Lin and Fan (2020b)	Charlotte, NC	U.S.A.	7 locations	Direct demand modelling
Lindsey et al. (2018)	Minneapolis, MN	U.S.A.	471 locations	Direct demand modelling
Liu et al. (2021)	New York, NY	U.S.A.	112 locations	Direct demand modelling
Lu et al. (2018)	Blacksburg, VA	U.S.A.	101 locations	Direct demand modelling
McDaniel et al. (2014)	Moscow, ID	U.S.A.	14 intersections	Mixed method – Spatial Network Analysis + Direct demand modelling
Munira et al. (2021)	Austin, TX	U.S.A.	44 locations	Direct demand modelling
Nelson et al. (2021)	Boulder, CO; Ottawa, ON; Phoenix, AZ; San Francisco, CA; Victoria, BC	U.S.A., Canada	1236 locations	Direct demand modelling

(*Continued*)

Table 1. Continued.

Author (year)	Study area	Countries	Number of count locations	Modelling approach
Orvin et al. (2021)	Auckland; Kelowna, BC	New Zealand, Canada	25 locations (Auckland), 12 locations (Kelowna)	Direct demand modelling
Oskarbski et al. (2021)	Gdynia	Poland	221 locations	Four-step modelling procedure
Rupi et al. (2019)	Bologna	Italy	46 locations	Direct demand modelling
Ryu et al. (2019)	Utah State University campus, Logan, UT	U.S.A.	46 locations	Two-stage bicycle origin and destination (OD) matrix estimation process
Sanders et al. (2017)	Seattle, WA	U.S.A.	403 locations	Direct demand modelling
Schoner et al. (2021)	Seattle, WA	U.S.A.	65 locations	Direct demand modelling
Ryu (2020)	Winnipeg, MB	Canada	19 locations	Two-stage bicycle origin and destination (OD) matrix estimation process
Strauss et al. (2015)	Montreal, QC	Canada	1065 locations	Direct demand modelling
Wallentin and Loidl (2015)	Salzburg	Austria	3 locations	Agent-based simulation modelling
Wang et al. (2016)	Minneapolis, MN; Columbus, OH	U.S.A.	86 locations (Minneapolis) and 41 locations (Columbus)	Direct demand modelling

Table 2. Data types by study.

Study	Bike count data	Infrastructure data	Census data	Network characteristics data	Land-use data	Strava data	Weather data	Travel survey data	Traffic flow data	GPS data	Crash data	Google Earth bicycle facility data	Google StreetView imagery	OD data	POI data
Arellana et al., (2020)	✓	✓						✓	✓						
Bargh et al., (2012)	✓	✓							✓						
Cooper (2016)	✓			✓											
Cooper (2017)	✓		✓	✓							✓				
Cooper (2018)	✓			✓											
Dadashova et al. (2020)	✓	✓	✓			✓									
Dadashova and Griffin (2020)	✓	✓	✓			✓	✓								
Ermagun et al. (2018)	✓	✓			✓		✓								
Fagnant and Kockelman (2016)	✓	✓	✓												
Fan and Lin (2019)	✓	✓	✓	✓		✓									
Gehrke and Reardon (2021)	✓	✓	✓												
Gosse and Clarens (2014)	✓	✓					✓								
Hankey et al. (2012)	✓	✓	✓				✓								
Hankey et al. (2021)	✓	✓	✓				✓					✓	✓		✓
Haworth (2016)	✓					✓									
Hochmair et al. (2019)	✓	✓	✓	✓	✓	✓									
Jacyna et al. (2017)		✓	✓		✓			✓							
Jahangiri et al. (2019)	✓	✓	✓												
Jestico et al. (2016)	✓	✓				✓									
Jones et al. (2010)	✓	✓	✓					✓							

(*Continued*)

Table 2. Continued.

Study	Bike count data	Infrastructure data	Census data	Network characteristics data	Land-use data	Strava data	Weather data	Travel survey data	Traffic flow data	GPS data	Crash data	Google Earth bicycle facility data	Google StreetView imagery	OD data	POI data
Kaziyeva et al. (2021)	✓	✓	✓					✓							
Kwigizile et al. (2019)	✓	✓	✓		✓	✓	✓	✓							
Le et al. (2017)	✓	✓	✓	✓	✓		✓								
Lin and Fan (2020b)		✓	✓			✓									
Lin and Fan (2020a)	✓	✓	✓	✓											
Lindsey et al. (2018)	✓	✓	✓	✓	✓		✓								
Liu et al. (2021)	✓		✓	✓	✓										
Lu et al. (2018)	✓	✓	✓	✓	✓										
McDaniel et al. (2014)	✓	✓		✓	✓									✓	
Munira et al. (2021)	✓	✓	✓	✓	✓										
Nelson et al. (2021)	✓		✓		✓	✓	✓		✓		✓				
Orvin et al. (2021)	✓	✓	✓	✓	✓		✓								
Oskarbski et al. (2021)	✓	✓	✓		✓			✓		✓					
Rupi et al. (2019)	✓									✓					
Ryu (2020)	✓	✓		✓				✓	✓						
Ryu et al. (2019)	✓	✓	✓	✓				✓	✓						
Sanders et al. (2017)	✓	✓	✓			✓									
Schoner et al. (2021)	✓		✓			✓									
Strauss et al. (2015)	✓									✓					
Wallentin and Loidl (2015)	✓		✓					✓							
Wang et al. (2016)	✓	✓	✓	✓	✓										

speed limit, slope of road segment, proximity to transit stops, etc.) ($n = 31$) and census data (aggregated socio-demographic information such as median age, percentage of people > 65 years, percentage of male, female, male bicyclists, percentage of population with at least a high-school degree, more highly-educated people, employment density, job accessibility, unemployment rate, percentage of white people, percentage of African-Americans, population density, households within 0.25 miles of count location, number of housing units in each census block, median household income, number of households with income > $125,000, car ownership, bicycle mode share, mode share for bicycling, walking and transit combined, etc.) ($n = 29$). Network characteristics data (such as betweenness centrality, angular distance, segment length, number of through lanes, size of street polygon in a block group, multimodal network density, intersection density, road-connectivity index, etc.) were employed in 16 studies, land use data (percentage of industrial, retail, residential and open space area, if commercial land use is greater than residential land use, proximity to university and other educational institutes, proximity to water bodies, etc.) in 14 studies, while crowdsourced Strava data (link-level Strava counts, percentage of Strava trips that are commuting trips, etc.) was used in 11 studies. Weather data (mean temperature, maximum precipitation, etc.) was used in 10 studies, travel survey data (travel diaries recording trips of participants over a specific time period) in nine studies, traffic flow data (link-level Annual Average Daily Traffic, percentage of right turning traffic, percentage of trucks, etc.) was used in five studies, while bicycle crash data (bicycle crash density, number of crashes) was used in two studies. Points-of-interest data (POI) (locations that serve as probable destinations of cyclists), origin-destination data (OD) (aggregated trip counts from one travel analysis zone to another in a study area), Google Street View imagery (for characterising street environment near the count locations) and Google Earth data (derive bicycle facility data derived from manual coding of imagery data) were used in one study each. The types of data used by individual studies are presented in Table 2. It is important to note that our search strategy is not designed to be biased towards any particular trip tracking application. However, the studies included in this review that used crowdsourced mobility data had only used Strava Metro data for this purpose. There are several other tracking applications such as Endomondo, MapMyRide, MapMyFitness and Garmin Connect (Romanillos et al., 2016). Figure 2 shows the types of datasets used vs their corresponding modelling approaches.

4.4. Reporting study and modelling details

Although the studies included in this review aimed to estimate bicycling volumes at link-level, there were inconsistencies observed in reporting of modelling details and results. First, of the 41 studies, only three studies objectively reported (in terms of a number) the size of the study area (Lu et al., 2018; Ryu et al., 2019; Wallentin & Loidl, 2015). We acknowledge that most of the other studies had mentioned the names of these geographical locations such as cities and counties which indicate the extent of the study area. However, the real extent of study areas is often ambiguous. For example, Melbourne, City of Melbourne and Greater Melbourne are areas with different sizes. Reporting the size of the study area size objectively in terms of an approximate number is a trivial task, and can help the reader judge the extrapolation capability of the model, i.e. how well it is able

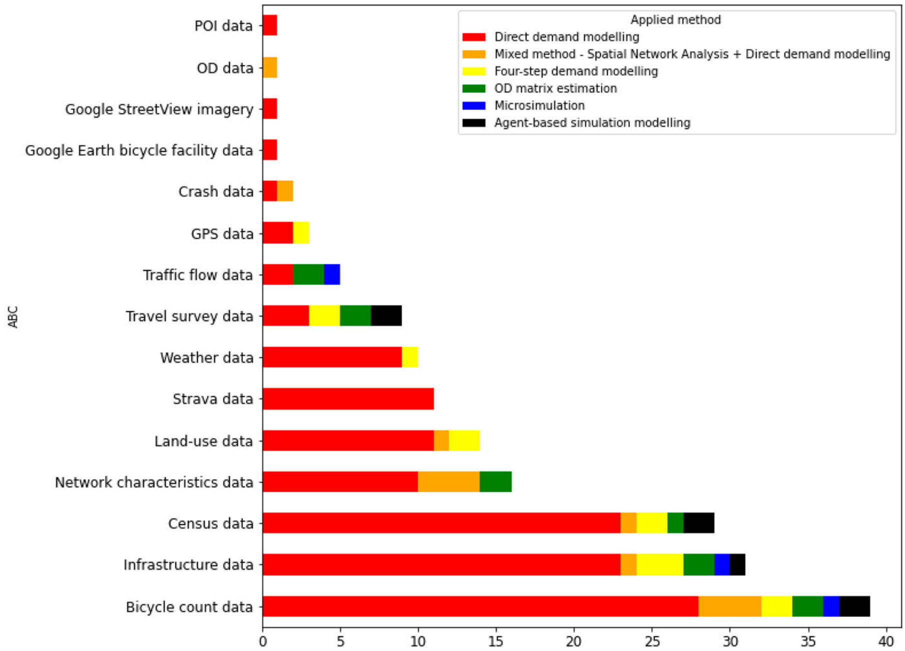

Figure 2. Number of studies by data type and applied method.

to predict bicycling volumes at locations where there is no count information available. For example, 15 count locations across a suburb would have a greater probability of being representative of the entire suburb (i.e. the study area), than 15 count locations spread across a large city being representative of the entire city. Second, there were only three studies out of the 41 that reported the number of links or the number of nodes that were part of their study (Jahangiri et al., 2019; Oskarbski et al., 2021; Rupi et al., 2019; Ryu, 2020) while three out of the 39 studies that used bicycle count data did not detail the number of bicycle count locations (Bargh et al., 2012; Oskarbski et al., 2021; Ryu et al., 2019). Bicycle counts is the most critical data that is used by the studies as the dependent variable in their models. Therefore, reporting the number of count sites is crucial to help the reader understand the foundation of the model. Reporting of the number of nodes (intersections) or links (street segments) covered by the study area helps the reader to understand the proportion of known data (number of nodes or links with known counts) and the extent of extrapolation (number of nodes or links with unknown counts where volumes have been estimated).

4.5. Model validation and reporting of accuracies and errors

There was variation in reporting of accuracy and errors of the developed models. Model effectiveness was reported in the form of accuracy, *R*-squared (Ermagun et al., 2018; Haworth, 2016), log-likelihood (Fagnant & Kockelman, 2016; Gehrke & Reardon, 2021), Mean Absolute Error (MAE) (Jahangiri et al., 2019; Munira et al., 2021), Mean Absolute Percentage Error (MAPE) (Dadashova & Griffin, 2020; Jestico et al., 2016), Mean Relative Percentage Error (MRPE) (Ermagun et al., 2018), and/or Root Mean Squared Error (RMSE)

(Jahangiri et al., 2019; Ryu, 2020). DDMs were more standardised in this respect, with all studies but two (Arellana et al., 2020; Schoner et al., 2021) reporting model effectiveness using at least one metric. Other approaches did not report model accuracy consistently.

The three studies which only report log-likelihood values (Fagnant & Kockelman, 2016; Gehrke & Reardon, 2021; Liu et al., 2021) have compared the goodness of fit of multiple models (not model types) such as a Poisson model and negative binomial model and nested models (comparing one model with a subset of predictor variables of another). However, log-likelihood values are not helpful in terms of assessing the accuracy or the representativeness of a model. Therefore, this should have been supported with better metrics such as *R*-squared and mean errors.

Model validation is an important step that reveals the representativeness of the model for the study area, especially when trying to estimate bicycling volumes for an area which is relatively large in size (such as a city or a county). Validation efforts were conducted by 23 studies included in this review. Most of the studies (18 out of 23) conducted either used k-fold cross validation, random *n*% hold-out cross validation or leave-one-out cross-validation. Pattern-oriented modelling (POM) framework was adopted by both the ABM studies to validate their models. Multiple patterns were used to test the validity of model outputs such as spatio-temporal distributions of cyclists over the study area and their relative frequencies from observed datasets. One study reported a cross-city validation method, where the model developed on one city was applied to another city (and vice versa) and checked for estimation accuracy. Ten out of 41 studies reported neither estimation accuracy nor errors of their models (see Table 3). Interestingly, a DDM study used cross-validation techniques to achieve the best-fitting model (Schoner et al., 2021). However, the study did not report metrics reflecting the accuracy for the bicycling model. Furthermore, some studies did not report the units of their errors. Hankey et al. (2021) described count data collection for peak periods, but their results did not mention the unit of Mean Absolute Error (bicycles per hour or bicycles per peak period, consisting of two hours). Table 3 presents the accuracy/error metrics used, accuracy/error values and validation details reported by the studies included in the review.

4.6. Strengths and limitations of different modelling approaches

In this section, the order of presenting the modelling approaches is based on their current usage for link-level bicycling volume modelling. Direct demand modelling is presented first because of its predominance, while the adjacency of the four-step model and the agent-based model helps the reader understand how activity-based models address the limitations of four-step trip-based models.

4.6.1. Direct demand modelling and related mixed methods

Direct demand modelling was the most popular modelling approach ($n = 29$) identified in this review.(Arellana et al., 2020; Orvin et al., 2021; Schoner et al., 2021) Using count data (dependent variable) from a set of locations, and variables that influence bicycling volumes such as socio-demographics, network characteristics, and land-use patterns (independent variables), direct demand models apply statistical methods (Sanders et al., 2017; Schoner et al., 2021; Wang et al., 2016) or machine learning algorithms (Hankey et al., 2021; Kwigizile et al., 2019; Nelson et al., 2021) to estimate link level

Table 3. Modelling types vs. accuracy/errors.

Method	Study	Validation	No. of count sites	Validation method	Accuracy	R-squared	Log-likelihood	MAE	MAPE	MRPE	RMSE
Direct demand modelling	Arellana et al. (2020)	No	27								
	Dadashova and Griffin (2020)	Yes	350	Leave-one-out cross-validation with 100 repetitions					29%		
	Dadashova et al. (2020)	Yes	100	Random 20% hold-out cross validation	70–75%				29%		
	Ermagun et al. (2018)	Yes	32	Leave-one-out cross-validation		0.63				65.4%	
	Fagnant and Kockelman (2016)	No	251				−970.05				
	Fan and Lin (2019)	No	7			0.61					
	Gehrke and Reardon (2021)	No	91				−603.31				
	Hankey et al. (2012)	Yes	259	Comparing estimated 12-h counts to 12-h counts imputed from scaling factors for the year 2010 at 85 locations (46 new and 39 previously sampled locations).		0.48					
	Hankey et al. (2021)	Yes	4145	Random hold-out cross validation		0.84		34–63 (units missing)			
	Haworth (2016)	Yes	298	Cross-validation		0.67					72.9 cyclists per hour
	Hochmair et al. (2019)	No	32			0.45–0.50					
	Jahangiri et al. (2019)	Yes	88	10-fold cross validation		0.67		87.68 AADB			105.93 AADB
	Jestico et al. (2016)	Yes	18	Random 10% hold-out cross validation with 100 repetitions	76–85%				38%		
	Jones et al. (2010)	No	80			0.47					
	Kwigizile et al. (2019)	Yes	19	10-fold cross-validation with 10 repetitions		0.71					
	Le et al. (2017)	Yes	9870	Random 10% hold-out cross validation with 100 repetitions		0.19–0.56					
	Lin and Fan (2020a)	No				0.22					
	Lin and Fan (2020b)	No	7			0.61					
	Ermagun et al. (2018)	Yes	471	Monte Carlo-based 10% hold-out analysis		0.58					

	Liu et al. (2021)	No	112				−659.36, −646.51		
	Lu et al. (2018)	Yes	101	Monte Carlo-based 20% holdout analysis		0.49			
	Munira et al. (2021)	Yes	44	10-fold cross validation		0.7		132 AADB	171 AADB
	Nelson et al. (2021)	Yes	1236	10-fold cross validation		0.08–0.92			
	Orvin et al. (2021)	No	37			0.996–0.997			
	Rupi et al. (2019)	No	46			0.73			
	Sanders et al. (2017)	No	403			0.57–0.62			
	Schoner et al. (2021)	Yes	65	Leave-one-out cross-validation					
	Strauss et al. (2015)	Yes	1065	Correlation between AADB values obtained from GPS data and from count data		0.48–0.76			
	Wang et al. (2016)	Yes	127	Within city validation – random 10% hold out; Cross city validation – Model for one city is applied to the other city and vice versa		0.58–0.64			
Mixed methods	Cooper (2016)	Yes	121	Generalised cross-validation		0.65–0.78			
	Cooper (2017)	No	121			0.61			
	Cooper (2018)	Yes	121	7-fold cross validation with 50 repetitions		0.78			
	McDaniel et al. (2014)	Yes	14	Random 10% hold-out cross validation		0.45–0.61			
Agent-based modelling	Kaziyeva et al. (2021)	Yes	9	Pattern oriented modelling framework – Relative frequencies of spatial and temporal distributions of cyclists over the study area.				1002.27 cyclists per day	
	Wallentin and Loidl (2015)	Yes	3	Relative frequencies of cyclist count from 3 locations					
OD matrix estimation	Ryu (2020)	No	19						37.91 (unit missing)
	Ryu et al. (2019)	No	46						18.57 (unit missing)
Four-step demand modelling	Oskarbski et al. (2021)	No	221		0.77				
	Jacyna et al. (2017)	Yes		Sampling rate distributions in each speed interval were compared for the model dataset and the control observation dataset					
	Gosse and Clarens (2014)	No	18						
Microsimulation	Bargh et al. (2012)	No							

AADB = Annual Average Daily Bicycling traffic.

counts, first at known locations, and finally applying the resultant model to predict link level counts for all the links in the rest of the study area where count information is unavailable. DDMs are simpler to interpret, as the associations between the explanatory variables and the bicycling volumes are apparent from the magnitude and direction of the model coefficients and their significance values. This is unlike more complex models, where the cause and effect have to be traced back via a longer chain of model elements, which is challenging. Therefore, DDMs have been extensively employed by researchers to inform urban transport planners about the elements that drive volumes at locations across a study area. State and local government authorities often use results from the DDMs for operational purposes, such as to introduce an infrastructural element at more locations that is known to improve bicycling patronage (Jahangiri et al., 2019; Jones et al., 2010).

However, there are limitations of direct demand modelling approach. Variables that influence bicycling volumes at link level often vary across cities. For example, *distance to sea shore* was a unique explanatory variable in the model developed by Nelson et al. (2021). Some models considered *ethnicity* as a significant explanatory variable influencing bicycling volumes while others did not (Hochmair et al., 2019; Kwigizile et al., 2019; Nelson et al., 2021; Wang et al., 2016). Some studies have tried to overcome this limitation by developing a single direct demand model for multiple cities to expand the spatial coverage of their models (Ermagun et al., 2018; Hankey et al., 2021; Le et al., 2017). While that improves the transferability of the model across different cities, such multi-city models lose accuracy when compared to multiple city-specific models. Nelson et al. (2021) stated that the R-squared values of their generalised models ranged from 0.08 to 0.80 (when applied on individual cities), while that of their city-specific models ranged from 0.68 to 0.92. On the other hand, separate studies developing direct demand models for the same study area came up with different models. For example, Sanders et al. (2017) and Schoner et al. (2021) developed DDMs for Seattle, and the included independent variables differed between the two studies. Sanders et al. (2017) included unique variables such as *number of bike markings on a segment* and *presence of school within 0.25 miles of either end of a segment* while Schoner et al. (2021) included unique variables such as *households within 0.25 miles* and *type of street and available bicycle facility*. Often, DDMs are developed using a small number of known count locations (Fan & Lin, 2019; Jestico et al., 2016; Kwigizile et al., 2019; Lin & Fan, 2020b) which may be biased towards high-frequency bicycling locations (majority of count locations being on dedicated bicycle paths, and fewer on roads and streets). DDM predicts volumes for links without count data, and often assumes uniformity in the relationship between the predictor variables and the dependent variables (one set of coefficients for the entire study area), which may lead to inaccuracy in volume predictions. Most models do not account for temporal variability and seasonality observed in bicycle count data (Dadashova & Griffin, 2020). DDMs are unable to account for travel behaviour and perceptions of travellers (Kaziyeva et al., 2021). The models are simply statistical associations between observed counts and ambient information around a count site relevant for bicycling, however, they do not incorporate the nuance of travel behaviour of bicyclists. Hence, they are not capable of considering the interactions between the travellers and their environment, or the interactions between the travellers themselves, based on the characteristics of the travellers. Modelling this heterogeneity for bicycling is extremely important as bicycling preferences is dependent on a wider range of individual and environmental factors and is heterogeneous (Loidl et al., 2019).

4.6.2. Four-step demand modelling

Traditional demand models (trip-based models) are usually developed based on detailed data collected via household travel surveys. Such models are commonly macroscopic in nature, covering large study areas. Unlike DDMs, the major advantage of traditional four-step models is that they are employed for large-scale, long-term planning purposes. For example, Oskarbski et al. (2021) developed a bicycle traffic model for the city of Gdynia, Poland to support the planning and decision making with regard to changes in the transport network.

However, disaggregated versions of large-scale trip-based demand models that have higher spatial granularity, are data-intensive and computationally expensive. As a result, to apply them in the bicycling context becomes challenging as they become computationally expensive for high resolution bicycling networks, requiring a significant increase in the cost of calibration. Despite such challenges, Oskarbski et al. (2021), Jacyna et al. (2017) and Gosse and Clarens (2014) had developed trip-based models for modelling link-level bicycling volumes. However, these high-resolution bicycling models use area-level origin-destination (OD) survey data, which is of far lower spatial and temporal resolution. Oskarbski et al. (2021) simulated trips from centroids of travel-analysis zones (TAZs) due to low-resolution of the OD data. Depending on the size of the TAZs, this can lead to large inaccuracies in modelling, especially in local non-arterial roads, given link-level volumes are being estimated. Also, adequate information related to different steps in the modelling procedure was not clearly communicated. For example, Jacyna et al. (2017) only stated the importance of link classification based on road characteristics for route choice, but did not present any detail on how the route assignment was conducted in their Warsaw model. Furthermore, trip assignment of these studies were not robust. For example, Oskarbski et al. (2021) had considered surface type and longitudinal gradient to be the only factors influencing bicycling route choice. While Gosse and Clarens (2014) had considered the importance of dedicated bicycling infrastructure on route assignment, they employed only two attribute levels with coarse values. Models of this nature are unable to account for heterogeneity of route preferences among cyclists.

However, there are inherent limitations of such traditional trip-based demand models. While these models have ensured the popularity of simulation in transport planning, their inability to associate trips to individuals, capture their heterogeneous behaviour and their interactions, and inter-dependencies between different components of the transport system such as infrastructure, congestion, mode and route choice, and activity chains, calls for more sophisticated models that address these limitations (Rasouli & Timmermans, 2014). Activity-based models (an alternative to trip-based models) simulate the individual behaviour of each traveller in the system, their interactions with the environment and other travellers.

4.6.3. Agent-based modelling

In the transport demand modelling space, agent-based modelling approaches are appropriate for producing spatially and temporally explicit traffic flows. Agent-based models (ABMs) are congruent with disaggregated activity-based models (different to trip-based models of traditional four-step demand modelling). Activity-based models are more promising for active travel modes such as bicycling, as unlike trip-based models and DDMs, activity-based models focus on individual travellers, their mode and route choice decisions being influenced by their characteristics (Liu et al., 2012). Agent-based models for traffic flows are relatively recent (due to advancement in computational capacity). These models have the ability to incorporate the travel behaviour of individual people (agents in the model). Also, parameter tweaking in agent-based models allows for scenario analysis, and therefore test the system's sensitivity to plausible hypothetical scenarios.

There still exists challenges with agent-based models when attempting to calibrate to real-world scenarios. First, generating an agent population that appropriately represents the travel behaviour of the real-world population is challenging. ABMs rely on robust inputs to the model related to the agent population, and their mobility behaviour when their goal is to accurately estimate volumes. Researchers often rely on census data to assign spatially driven socio-demographic characteristics to the agents, and based on travel survey data, assign activities and destinations. However, behaviour assignment is mostly assumption-based, since individual or household-level population-wide information about travel behaviour is rarely available, particularly in low-income countries. Wallentin and Loidl (2015) categorised their agents into "working cyclists", "student cyclists", and "leisure cyclists" and assigned only two trips (home to destination, and destination to home) to each agent. Kaziyeva et al. (2021) increased the level of complexity and assigned age, gender and employment status to their agents. Second, the route choice heuristics assigned to the agents are often too simplistic and are not based bicycling route choice models. Kaziyeva et al. (2021) implemented a safest route assignment in their model, and so did Wallentin and Loidl (2015). However, Wallentin and Loidl (2015) accounted for the variations of perception and preferences across individuals by introducing a stochastic variation in routing weights. Third, representative ABMs are highly complex and therefore are computationally expensive. Study areas of bicycling models commonly span across entire cities. An ABM designed to accurately represent real-world travel demand needs to assign characteristics to individual agents that govern their travel behaviour. Bicycling ABMs that intend to estimate link-level volumes require high spatial resolution to capture fine-grained movement of these agents moving across the network. As the ABM attempts to replicate the actual travel behaviour of an entire city, complexity increases, and the model becomes computationally expensive.

4.6.4. Summary

The Table 4 summarises the strengths and limitations of the aforementioned modelling approaches.

4.7. Method vs. accuracy

While we sought to understand the effectiveness of different methods at estimating link-level bicycling volumes, it was not possible to do so, given the range of confounding factors present across these studies. We have already presented the details of modelling accuracy and errors of all the studies included in this review in Table 3. Despite the use of standardised metrics, it is challenging to compare these methods using the metrics alone.

First, model performance depends highly on data quality and quantity, which varies across studies. Sourcing of bicycling travel data from travel surveys, counts, and nonmotorized travel infrastructure databases remains a challenge (Liu et al., 2012). Since traditional monitoring methods are resource-intensive, bicycling travel data is usually limited by small sample sizes, low spatial and temporal coverages, and infrequent updates (Lee & Sener, 2020). Therefore, there is a lack of standardisation in not just data quantity (e.g. number of count sites ranging from 3 (Wallentin & Loidl, 2015) to 9870 locations (Le et al., 2017)), but also in methods of data collection (e.g. manual counts (Ryu et al., 2019) vs automatic counts (Schoner et al., 2021); counts on AM and PM peaks (Rupi et al., 2019) vs continuous counts (Cooper, 2018); short-term counts (Liu et al., 2021) vs long-term counts (Hankey et al., 2021)), which may affect model accuracy. Studies are often reliant on using a combination of counts with different temporal coverages and resolutions (Hankey et al., 2012; Munira et al., 2021; Strauss et al., 2015). Second, we noted that each study only employed a single modelling approach, and no study compared the accuracy of different modelling approaches, limiting our understanding of the most effective modelling approach.

4.8. Exclusion of research related to bike-share programmes

Bike-sharing programmes have gained popularity over the years. They could serve as a rich source of information for bicycling related research. However, they were outside the scope of this review for the following reasons. Bike-sharing data is fundamentally different from all other forms of bicycling data as it is normally complete and more detailed. Most bike-sharing datasets have entire GPS traces, and therefore bike-sharing data contains detailed spatial and temporal information about all the bikes involved in that particular bike-share platform, and all from a single user base. Link-level (bike-sharing) volume modelling using bike-sharing data is not necessary (as we already have all necessary information). Therefore, such studies do not apply any transport demand modelling that would be of interest to this review, and hence, the scope of this review excludes modelling bike-sharing volumes.

This is also reflected in existing bike-sharing literature. Bike-sharing data is almost always used to develop models that predict bike-sharing demand at stations, temporal demand analysis and forecasting of bike-sharing, bike-sharing route choice and others. Conducting link-level (overall) volume modelling is not appropriate with bike-sharing counts and therefore, the scope of this review excludes studies related to bike-sharing. Only Pogodzinska et al. (2020) had expanded bike-share GPS traces to estimate overall bicycling volumes with high accuracy, however, they had only made estimations at 5 locations, and their model was not validated.

4.9. Recommendations and possible research directions

We make the following recommendations for future studies to consider.

- Studies could undertake multiple modelling approaches in the same study area using the same data leading the way to comparison of accuracy across different types of models.

Table 4. Strengths and limitations of major modelling approaches.

	Direct demand modelling	Four-step demand modelling	Agent-based modelling
Strengths	• Low computational complexity • Easier to infer the cause-effect relationships • Useful for operational purposes • Model validation is easier	• Suitable for long-term urban and traffic planning purposes • Can account for route choice of bicyclists	• Able to capture heterogeneity in rider behaviour • Able to capture rider-rider and rider-environment interaction • Suitable for long-term urban and traffic planning purposes • Can account for route choice of bicyclists • Can be used to develop and test future scenarios
Limitations	• Unable to capture heterogeneity in rider behaviour • Unable to capture rider-rider and rider-environment interaction • Unable to account for route choice of bicyclists • Not suitable for long-term urban and traffic planning purposes	• Unable to capture heterogeneity in rider behaviour • Unable to capture rider-rider and rider-environment interaction • Hard to infer cause-effect relationships • Implementation of this model is not simple • Data intensive • Model is validation is challenging	• Most complex of all modelling approaches • Hard to infer cause-effect relationships • Data intensive • Model is validation is challenging

- In cases where studies will have access to temporal patterns of data (daily, monthly, yearly) across a large enough timespan, they should explore both spatial and temporal expansion of the model (estimating current demand and forecasting future demand at locations where volume information is absent).
- Studies should ensure that they report basic study characteristics, thereby allowing greater readability and reproducibility of their research methods. Most important, they should report the number of count sites which are used to develop their model, and the extent of the study area, where the developed model estimates volumes.
- In case of direct demand models, studies should report all possible accuracy and error metrics. Errors should be provided with units.
- Studies should validate their models and report the same, otherwise the validity of such models is difficult to judge by the reader. In cases where validation is not being done, this should be explicitly mentioned with justifications.
- Agent-based modelling studies should base the route assignments of their agents on a data-driven route choice model, where possible.
- More studies should be encouraged in developing nations.
- Appropriate state-of-the-art machine learning techniques (deep learning methods) such as graph neural networks (which can capture graph-structured data such as road networks), which are already applied for traffic prediction and forecasting, should be translated to estimate link-level bicycling volumes and consequently,

comparisons of the results could be drawn with traditional approaches. However, it must be noted that such deep learning methods have high data demand (Yin et al., 2022), which is a challenge in the bicycling space.

5. Conclusions

Given the importance of sustainable urban goals, the importance of active mobility modes such as bicycling has become paramount. Strategic infrastructural and policy interventions require city-wide high-resolution models of bicycling. In this review of methods for modelling link-level bicycling volumes, we identified a variety of methodological approaches, all of which have intrinsic limitations. Direct demand modelling was most frequently employed, but the inability to account for heterogeneity in travel behaviour and the inability to leverage this approach for forecasting are major limitations of this method. Agent-based simulation models may overcome these limitations, but existing studies for modelling link-level bicycling volumes were infrequent and further validation is required. Further research is required to contrast modelling approaches within the same study area to determine the most robust approaches for modelling link-level bicycling volumes.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

BB was supported by an Australian Research Council Future Fellowship [FT210100183]. The CYCLED (CitY-wide biCycLing Exposure modelling) Study is funded by an Australian Research Council Discovery Project [DP210102089].

ORCID

Meead Saberi http://orcid.org/0000-0002-6526-239X
Meghan Winters http://orcid.org/0000-0001-9932-0255
Trisalyn Nelson http://orcid.org/0000-0003-2537-6971
Christopher Pettit http://orcid.org/0000-0002-1328-9830
Le Hai Vu http://orcid.org/0000-0001-6984-2060

References

Abduljabbar, R. L., Liyanage, S., & Dia, H. (2022). A systematic review of the impacts of the coronavirus crisis on urban transport: Key lessons learned and prospects for future cities. *Cities*, *127*, 103770. https://doi.org/10.1016/j.cities.2022.103770

Anderson, M. D., Sharfi, K., & Gholston, S. E. (2006). Direct demand forecasting model for small urban communities using multiple linear regression. *Transportation Research Record*, *1981*(1), 114–117. https://doi.org/10.1177/0361198106198100117

Apronti, D. T., & Ksaibati, K. (2018). Four-step travel demand model implementation for estimating traffic volumes on rural low-volume roads in Wyoming. *Transportation Planning and Technology*, *41*(5), 557–571. https://doi.org/10.1080/03081060.2018.1469288

Arellana, J., Saltarín, M., Larrañaga, A. M., González, V. I., & Henao, C. A. (2020). Developing an urban bikeability index for different types of cyclists as a tool to prioritise bicycle infrastructure investments. *Transportation Research Part A: Policy and Practice*, *139*, 310–334. https://doi.org/10.1016/j.tra.2020.07.010

ATAP. (2022). *Australian Transport Assessment and Planning (ATAP) guidelines: Overview of transport modelling*. Australian Transport Assessment and Planning (ATAP). https://www.atap.gov.au/tools-techniques/travel-demand-modelling/2-overview

Bargh, A., Kelly, J., & Li, V. (2012). Transportation modelling for modal walking and cycling communities. *Traffic Engineering & Control*, *53*(2), 65–70. https://trid.trb.org/view/1136947

Bian, J., Li, W., Zhong, S., Lee, C., Foster, M., & Ye, X. (2022). The end-user benefits of smartphone transit apps: A systematic literature review. *Transport Reviews*, *42*(1), 82–101. https://doi.org/10.1080/01441647.2021.1950864

Biswas, R. K., Friswell, R., Olivier, J., Williamson, A., & Senserrick, T. (2021). A systematic review of definitions of motor vehicle headways in driver behaviour and performance studies. *Transportation Research Part F: Traffic Psychology and Behaviour*, *77*, 38–54. https://doi.org/10.1016/j.trf.2020.12.011

Choi, J., Lee, Y. J., Kim, T., & Sohn, K. (2012). An analysis of Metro ridership at the station-to-station level in Seoul. *Transportation*, *39*(3), 705–722. https://doi.org/10.1007/s11116-011-9368-3

Cooper, C. (2016). *Spatial network analysis as a low cost land use-transport model of city wide cyclist flows*. European Transport Conference 2016. https://aetransport.org/past-etc-papers/conference-papers-2016https://trid.trb.org/view/1452158

Cooper, C. (2017). Using spatial network analysis to model pedal cycle flows, risk and mode choice. *Journal of Transport Geography*, *58*, 157–165. https://doi.org/10.1016/j.jtrangeo.2016.12.003

Cooper, C. (2018). Predictive spatial network analysis for high-resolution transport modeling, applied to cyclist flows, mode choice, and targeting investment. *International Journal of Sustainable Transportation*, *12*(10), 714–724. https://doi.org/10.1080/15568318.2018.1432730

Cooper, C., Harvey, I., Orford, S., & Chiaradia, A. J. F. (2021). Using multiple hybrid spatial design network analysis to predict longitudinal effect of a major city centre redevelopment on pedestrian flows. *Transportation*, *48*(2), 643–672. https://doi.org/10.1007/s11116-019-10072-0

Covidence. (2022). *Covidence systematic review software, Veritas Health Innovation, Melbourne, Australia*. www.covidence.org

Dadashova, B., & Griffin, G. P. (2020). Random parameter models for estimating statewide daily bicycle counts using crowdsourced data. *Transportation Research Part D: Transport and Environment*, *84*(11), 390–402. https://doi.org/10.1016/j.trd.2020.102368

Dadashova, B., Griffin, G. P., Das, S., Turner, S., & Sherman, B. (2020). Estimation of average annual daily bicycle counts using crowdsourced strava data. *Transportation Research Record*, *2674*(11), 390–402. https://doi.org/10.1177/0361198120946016

de Dios Ortúzar, J., & Willumsen, L. G. (2011). *Modelling transport*. John Wiley & Sons.

De Geus, B., De Smet, S., Nijs, J., & Meeusen, R. (2007). Determining the intensity and energy expenditure during commuter cycling. *British Journal of Sports Medicine*, *41*(1), 8–12. https://doi.org/10.1136/bjsm.2006.027615

De Hartog, J. J., Boogaard, H., Nijland, H., & Hoek, G. (2010). Do the health benefits of cycling outweigh the risks? *Environmental Health Perspectives*, *118*(8), 1109–1116. https://doi.org/10.1289/ehp.0901747

DiGioia, J., Watkins, K. E., Xu, Y., Rodgers, M., & Guensler, R. (2017). Safety impacts of bicycle infrastructure: A critical review. *Journal of Safety Research*, *61*, 105–119. https://doi.org/10.1016/j.jsr.2017.02.015

Dill, J., & Gliebe, J. (2008). Understanding and measuring bicycling behavior: A focus on travel time and route choice.

Elsevier. (2022a). *Ei Compendex*. https://www.elsevier.com/solutions/engineering-village/content/compendex

Elsevier. (2022b). *Engineering village: Empowering engineers to solve the world's greatest challenges*. https://www.elsevier.com/solutions/engineering-village

Elsevier. (2022c). *Inspec: Engineering research database*. https://www.elsevier.com/solutions/engineering-village/content/inspec

Ermagun, A., Lindsey, G., & Hadden Loh, T. (2018). Bicycle, pedestrian, and mixed-mode trail traffic: A performance assessment of demand models. *Landscape and Urban Planning*, *177*, 92–102. https://doi.org/10.1016/j.landurbplan.2018.05.006

Fagnant, D. J., & 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*, *43*(1), 93–107. https://doi.org/10.1177/0265813515602568

Fan, W., & Lin, Z. (2019). *Evaluating the potential use of crowdsourced bicycle data in North Carolina*. https://cammse.uncc.edu/sites/cammse.uncc.edu/files/media/CAMMSE-UNCC-2018-UTC-Project-Report-03-Fan-Final.pdfhttps://trid.trb.org/view/1669971

Fox, J., & Hess, S. (2010). Review of evidence for temporal transferability of mode-destination models. *Transportation Research Record*, *2175*(1), 74–83. https://doi.org/10.3141/2175-09

Gehrke, S. R., & Reardon, T. G. (2021). Direct demand modelling approach to forecast cycling activity for a proposed bike facility. *Transportation Planning and Technology*, *44*(1), 1–15. https://doi.org/10.1080/03081060.2020.1849959

Gosse, C. A., & Clarens, A. (2014). Estimating spatially and temporally continuous bicycle volumes by using sparse data. *Transportation Research Record*, *2443*(1), 115–122. https://doi.org/10.3141/2443-13

Grabow, M. L., Spak, S. N., Holloway, T., Stone, B., Mednick, A. C., & Patz, J. A. (2012). Air quality and exercise-related health benefits from reduced car travel in the midwestern United States. *Environmental Health Perspectives*, *120*(1), 68–76. https://doi.org/10.1289/ehp.1103440

Hankey, S., Lindsey, G., Wang, X., Borah, J., Hoff, K., Utecht, B., & Xu, Z. (2012). Estimating use of non-motorized infrastructure: Models of bicycle and pedestrian traffic in Minneapolis, MN. *Landscape and Urban Planning*, *107*(3), 307–316. https://doi.org/10.1016/j.landurbplan.2012.06.005

Hankey, S., Zhang, W., Le, H. T. K., Hystad, P., & James, P. (2021). Predicting bicycling and walking traffic using street view imagery and destination data. *Transportation Research Part D: Transport and Environment*, *90*, 102651. https://doi.org/10.1016/j.trd.2020.102651

Haworth, J. (2016). Investigating the potential of activity tracking app data to estimate cycle flows in urban areas. *XXIII ISPRS Congress, Prague, Czech Republic*.

Heppenstall, A. J., Crooks, A. T., See, L. M., & Batty, M. (2011). *Agent-based models of geographical systems*. Springer Science & Business Media.

Hochmair, H. H., Bardin, E., & Ahmouda, A. (2019). Estimating bicycle trip volume for Miami-Dade county from Strava tracking data. *Journal of Transport Geography*, *75*, 58–69. https://doi.org/10.1016/j.jtrangeo.2019.01.013

Jacyna, M., Wasiak, M., Kłodawski, M., & Gołębiowski, P. (2017). Modelling of bicycle traffic in the cities using VISUM. *10th International Scientific Conference Transbaltica 2017: Transportation Science and Technology*.

Jahangiri, A., Hasani, M., Sener, I. N., Munira, S., Owens, J., Appleyard, B., Ryan, S., Turner, S. M., & Machiani, S. G. (2019). *Data mining to improve planning for pedestrian and bicyclist safety*. https://safed.vtti.vt.edu/wp-content/uploads/2020/08/01-003_Final-Research-Report_Final.pdf, https://trid.trb.org/view/1672601. https://doi.org/10.15787/VTT1/IUTNDS

Jestico, B., Nelson, T., & Winters, M. (2016). Mapping ridership using crowdsourced cycling data. *Journal of Transport Geography*, *52*, 90–97. https://doi.org/10.1016/j.jtrangeo.2016.03.006

Jones, M., Ryan, S., Donlon, J., Ledbetter, L., Ragland, D., & Arnold, L. (2010). *Measuring bicycle and pedestrian activity in San Diego County and its relationship to land use, transportation, safety, and facility type*.

Kamaluddin, N. A., Andersen, C. S., Larsen, M. K., Meltofte, K. R., & Várhelyi, A. (2018). Self-reporting traffic crashes – A systematic literature review. *European Transport Research Review*, *10*(2), 26. https://doi.org/10.1186/s12544-018-0301-0

Kaziyeva, D., Loidl, M., & Wallentin, G. (2021). Simulating spatio-temporal patterns of bicycle flows with an agent-based model. *ISPRS International Journal of Geo-Information*, *10*(2), 88. https://doi.org/10.3390/ijgi10020088

Kraft, G., & Wohl, M. (1967). New directions for passenger demand analysis and forecasting. *Transportation Research*, *1*(3), i–viii. https://doi.org/10.1016/0041-1647(67)90017-2

Kuzmyak, J. R., Walters, J., Bradley, M., & Kockelman, K. M. (2014). *Estimating bicycling and walking for planning and project development: A guidebook*.

Kwigizile, V., Oh, J.-S., & Kwayu, K. (2019). *Integrating crowdsourced data with traditionally collected data to enhance estimation of bicycle exposure measure*. https://wmich.edu/sites/default/files/attachments/u883/2019/TRCLC_RR_17_03.pdfhttps://wmich.edu/transportationcenter/trclc17-3https://trid.trb.org/view/1483416

Le, H. T. K., Buehler, R., & Hankey, S. (2017). *Multi-city, national-scale direct-demand models of peak-period bicycle and pedestrian traffic*. http://www.matsutc.org/wp-content/uploads/2017/06/Multi-City-Direct-Demand-Models-of-Peak-Period-Bicycle-and-Pedestrian-Traffic.pdfhttps://trid.trb.org/view/1501753

Leao, S. Z., & Pettit, C. (2016). Mapping bicycling patterns with an agent-based model, census and crowdsourced data. *International Workshop on Agent Based Modelling of Urban Systems*.

Lee, K., & Sener, I. N. (2020). Emerging data for pedestrian and bicycle monitoring: Sources and applications. *Transportation Research Interdisciplinary Perspectives*, *4*, 100095. https://doi.org/10.1016/j.trip.2020.100095

Lin, Z., & Fan, W. (2020a). Bicycle ridership using crowdsourced data: Ordered probit model approach. *Journal of Transportation Engineering Part A: Systems*, *146*(8), 04020076. https://doi.org/10.1061/JTEPBS.0000399

Lin, Z., & Fan, W. (2020b). Modeling bicycle volume using crowdsourced data from Strava smartphone application. *International Journal of Transportation Science and Technology*, *9*(4), 334–343. https://doi.org/10.1016/j.ijtst.2020.03.003

Lindsay, G., Macmillan, A., & Woodward, A. (2011). Moving urban trips from cars to bicycles: Impact on health and emissions. *Australian and New Zealand Journal of Public Health*, *35*(1), 54–60. https://doi.org/10.1111/j.1753-6405.2010.00621.x

Lindsey, G., Wang, J., Pterka, M., & Hankey, S. (2018). *Modeling bicyclist exposure to risk and crash risk: Some exploratory studies*. http://www.cts.umn.edu/Publications/ResearchReports/reportdetail.html?id=2702https://conservancy.umn.edu/bitstream/handle/11299/199776/CTS%2018-15.pdf https://trid.trb.org/view/1540891

Liu, B., Bade, D., & Chow, J. Y. J. (2021). Bike count forecast model with multimodal network connectivity measures. *Transportation Research Record*, *2675*(7), 320–334. https://doi.org/10.1177/03611981211021849

Liu, F., Evans, J. E., & Rossi, T. (2012). Recent practices in regional modeling of nonmotorized travel. *Transportation Research Record*, *2303*(1), 1–8. https://doi.org/10.3141/2303-01

Loidl, M., Werner, C., Heym, L., Kofler, P., & Innerebner, G. (2019). Lifestyles and cycling behavior – Data from a cross-sectional study. *Data*, *4*(4), 140. https://www.mdpi.com/2306-5729/4/4/140 https://doi.org/10.3390/data4040140

Lu, T., Mondschein, A., Buehler, R., & Hankey, S. (2018). Adding temporal information to direct-demand models: Hourly estimation of bicycle and pedestrian traffic in Blacksburg, VA. *Transportation Research Part D: Transport and Environment*, *63*, 244–260. https://doi.org/10.1016/j.trd.2018.05.011

McDaniel, S., Lowry, M. B., & Dixon, M. (2014). Using origin–destination centrality to estimate directional bicycle volumes. *Transportation Research Record*, *2430*(1), 12–19. https://doi.org/10.3141/2430-02

McFadden, D. (1974). The measurement of urban travel demand. *Journal of Public Economics*, *3*(4), 303–328. https://doi.org/10.1016/0047-2727(74)90003-6

Mohamad, D., Sinha, K. C., Kuczek, T., & Scholer, C. F. (1998). Annual average daily traffic prediction model for county roads. *Transportation Research Record*, *1617*(1), 69–77. https://doi.org/10.3141/1617-10

Munira, S., Sener, I. N., & Zhang, Y. (2021). Estimating bicycle demand in the Austin, Texas area: Role of a Bikeability Index. *Journal of Urban Planning and Development*, *147*(3), 3. https://doi.org/10.1061/(ASCE)UP.1943-5444.0000725

Nelson, T., Roy, A., Ferster, C., Fischer, J., Brum-Bastos, V., Laberee, K., Yu, H., & Winters, M. (2021). Generalized model for mapping bicycle ridership with crowdsourced data. *Transportation Research Part C: Emerging Technologies*, *125*, 102981. https://doi.org/10.1016/j.trc.2021.102981

Orvin, M. M., Fatmi, M. R., & Chowdhury, S. (2021). Taking another look at cycling demand modeling: A comparison between two cities in Canada and New Zealand. *Journal of Transport Geography*, *97*, 103220. https://doi.org/10.1016/j.jtrangeo.2021.103220

Oskarbski, J., Birr, K., & Żarski, K. (2021). Bicycle traffic model for sustainable urban mobility planning. *Energies*, *14*(18), 5970. https://doi.org/10.3390/en14185970

Pogodzinska, S., Kiec, M., & D'Agostino, C. (2020). Bicycle traffic volume estimation based on GPS data. *Transportation Research Procedia*, *45*, 874–881. https://doi.org/10.1016/j.trpro.2020.02.081

Pritchard, R. (2018). Revealed preference methods for studying bicycle route choice – A systematic review. *International Journal of Environmental Research and Public Health*, *15*(3), 470. https://www.mdpi.com/1660-4601/15/3/470 https://doi.org/10.3390/ijerph15030470

Raford, N., & Ragland, D. (2004). Space syntax: Innovative pedestrian volume modeling tool for pedestrian safety. *Transportation Research Record*, *1878*(1), 66–74. https://doi.org/10.3141/1878-09

Raford, N., & Ragland, D. (2006). *Pedestrian volume modeling for traffic safety and exposure analysis: The case of Boston, Massachusetts*.

Rasouli, S., & Timmermans, H. (2014). Activity-based models of travel demand: Promises, progress and prospects. *International Journal of Urban Sciences*, *18*(1), 31–60. https://doi.org/10.1080/12265934.2013.835118

Romanillos, G., Zaltz Austwick, M., Ettema, D., & De Kruijf, J. (2016). Big data and cycling. *Transport Reviews*, *36*(1), 114–133. https://doi.org/10.1080/01441647.2015.1084067

Roy, A., Nelson, T. A., Fotheringham, A. S., & Winters, M. (2019). Correcting bias in crowdsourced data to map bicycle ridership of all bicyclists. *Urban Science*, *3*(2), 62. https://doi.org/10.3390/urbansci3020062

Rupi, F., Poliziani, C., & Schweizer, J. (2019). Data-driven bicycle network analysis based on traditional counting methods and GPS traces from smartphone. *ISPRS International Journal of Geo-Information*, *8*(8), 322. https://doi.org/10.3390/ijgi8080322

Ryu, S. (2020). A bicycle origin-destination matrix estimation based on a Two-stage procedure. *Sustainability*, *12*(7), 2951. https://doi.org/10.3390/su12072951

Ryu, S., Su, J., Chen, A., & Choi, K. (2019). Estimating bicycle demand of a small community. *KSCE Journal of Civil Engineering*, *23*(6), 2690–2701. https://doi.org/10.1007/s12205-019-0415-5

Sanders, R. L., Frackelton, A., Gardner, S., Schneider, R., & Hintze, M. (2017). Ballpark method for estimating pedestrian and bicyclist exposure in Seattle, Washington: Potential option for resource-constrained cities in an age of big data. *Transportation Research Record*, *2605*(1), 32–44. https://doi.org/10.3141/2605-03

Schoner, J., Proulx, F., Orvañanos, K. K., & Almdale, B. (2021). Prioritizing pedestrian and bicyclist count locations for volume estimation. *Transportation Research Record*, *2675*(10), 277–290. https://doi.org/10.1177/03611981211011164

Stinson, M. A., & Bhat, C. R. (2003). Commuter bicyclist route choice: Analysis using a stated preference survey. *Transportation Research Record*, *1828*(1), 107–115. https://doi.org/10.3141/1828-13

Strauss, J., Miranda-Moreno, L. F., & Morency, P. (2015). Mapping cyclist activity and injury risk in a network combining smartphone GPS data and bicycle counts. *Accident Analysis and Prevention*, *83*, 132–142. https://doi.org/10.1016/j.aap.2015.07.014

Tilahun, N. Y., Levinson, D. M., & Krizek, K. J. (2007). Trails, lanes, or traffic: Valuing bicycle facilities with an adaptive stated preference survey. *Transportation Research Part A: Policy and Practice*, *41*(4), 287–301. https://doi.org/10.1016/j.tra.2006.09.007

Turner, S., Sener, I. N., Martin, M. E., Das, S., Hampshire, R. C., Fitzpatrick, K., Molnar, L. J., Colety, M., Robinson, S., & Shipp, E. (2017). *Synthesis of methods for estimating pedestrian and bicyclist exposure to risk at areawide levels and on specific transportation facilities*.

Wallentin, G., & Loidl, M. (2015). Agent-based bicycle traffic model for Salzburg city. *GI_Forum Journal for Geographic Information Science*, *3*, 558–566. https://doi.org/10.1553/giscience2015s558

Wang, J., Hankey, S., Wu, X., & Lindsey, A. G. (2016). Monitoring and modeling of urban trail traffic: Validation of direct demand models in Minneapolis, Minnesota, and Columbus, Ohio. *Transportation Research Record*, *2593*(1), 47–59. https://doi.org/10.3141/2593-06

Wang, L., Waddell, P., & Outwater, M. L. (2011). Incremental integration of land use and activity-based travel modeling: Workplace choices and travel demand. *Transportation Research Record*, *2255*(1), 1–10. https://doi.org/10.3141/2255-01

Wen, L. M., & Rissel, C. (2008). Inverse associations between cycling to work, public transport, and overweight and obesity: Findings from a population based study in Australia. *Preventive Medicine*, *46*(1), 29–32. https://doi.org/10.1016/j.ypmed.2007.08.009

Wilensky, U., & Rand, W. (2015). *An introduction to agent-based modeling: Modeling natural, social, and engineered complex systems with NetLogo*. MIT Press.

Winters, M., & Branion-Calles, M. (2017). Cycling safety: Quantifying the under reporting of cycling incidents in Vancouver, British Columbia. *Journal of Transport & Health*, *7*, 48–53. https://doi.org/10.1016/j.jth.2017.02.010

Winters, M., Teschke, K., Grant, M., Setton, E. M., & Brauer, M. (2010). How far out of the way will we travel? Built environment influences on route selection for bicycle and car travel. *Transportation Research Record*, *2190*(1), 1–10. https://doi.org/10.3141/2190-01

Yin, X., Wu, G., Wei, J., Shen, Y., Qi, H., & Yin, B. (2022). Deep learning on traffic prediction: Methods, analysis, and future directions. *IEEE Transactions on Intelligent Transportation Systems*, *23*(6), 4927–4943. https://doi.org/10.1109/TITS.2021.3054840

Zhao, F., & Chung, S. (2001). Estimation of annual average daily traffic in a Florida county using GIS and regression. *Transportation Research Record*, *1769*(1), 113–122. https://doi.org/10.3141/1769-14