MAPPING BICYCLE LEVEL OF SERVICE

by

Dane Atkins, M.A.

A dissertation submitted to the Graduate Council of
Texas State University in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
with a Major in Geography
May 2022

Committee Members:

Osvaldo Muñiz-Solari, Chair

Jennifer Devine

Edwin Chow

Feng Wang

DEDICATION

For every bicyclist who took their lane.
For every bicyclist who won't.

For every bicyclist who rides in the rain.
And for those who don't.

For every bicyclist, to whom pain
Is a friend.

For every bicyclist stopped for a train,
Cause they're speedin' on the flyway.

For every bicyclist ridin' on glass,
While they're readin' on the highway.

For every bicyclist, watching 'em pass,
into oncoming traffic on the ol' country bi-way.

To every motorist. For every bicyclist.
We are traffic; w*e play in traffic.*

ACKNOWLEDGEMENTS

I am indebted with deepest gratitude to Dr. Osvaldo Muñiz-Solari, for the opportunity to receive his lessons as both a student of geography and pedagogy. It was Dr. Jennifer Devine who gave empathetic guidance through this project, and graduate school more generally, and for that I am also deeply grateful. More thanks to Dr. Edwin Chow and Dr. Feng Wang, who provided their critical eyes to these methods and models; I am grateful to them, and the committee, for seeing this project through to the end.

I would also like to thank Dr. Eric Sarmiento, Dr. Russel Weaver, and Dr. Ronald Hagelman for orientating me to the Department of Geography. I must also extend a deeply sincere thanks to Allison Glass-Smith and Angelica Wahl, who were among the first to welcome me to Texas. Another thank you I owe to Dr. Yongmei Lu, for her thoughtful discussion on what the University ought to be. More thanks still, to Dr. Injeong Jo for her lessons in pedagogy, to Mark Carter for his discussions on research methods, and to the supervision of Dr. Christi Townsend – their influence on my teaching helped maintain the precarious balance between instruction and research. Then there is Charles, Dan, Joyce, and so many others in our school's community – it is thanks to all of the staff and students at Texas State for making this work possible.

I am thankful, too, for my peers in geography. We all take different paths through graduate school, but we take them all together. It is my peers who seem to most clearly understand such ramblings, or at least tolerate them, whether at noon or by midnight.

I cannot understate the importance of the San Marcos bicyclists who contributed to this work; to them I am deeply grateful for their time and the perceptions they have developed over thousands of hours bicycling. There is also a special thanks owed to those who work diligently to maintain our vibrant bicycling community: Kerry, Elroy, Alex V., Joe S., Alex B., and many more. More thanks still to anyone who would or will ride a bicycle in San Marcos; it is by their energies that I believe San Marcos can lead the way in the promotion of the bicycle – I believe in strength in numbers and in the hundreds of bicyclists who believe in their right to the road.

I am grateful for the support of my best friends. In many cases, our relationships have withstood the tests of time and distance and become all the stronger for it. Thank you, Doug, J.P, and Mujahid; thank you for just being there (and for all the games)!

Without my family, this project would not be possible. I would like to thank my wife, Elizabeth, for understanding the delicate balance between the roots and the wild, and to thank my mother, Marsha, for reading every word, even when we disagree on their order or pluralities. I would also like to thank Tom and Kate, for their support of both education and geography, learned and lived.

Last but not least, I would like to thank *everyone*. I am grateful to live in a society which values education as highly as we do; I am honored to have been given the opportunity to pursue my own education to the fullest. I am indebted to you all.

Thanks y'all.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure **Page**

ABSTRACT

Bicycle level of service (BLOS) offers transportation planners a mathematical measurement for evaluating roadway infrastructure. Over the last several decades, there have been many formulations used to calculate BLOS; this research aims to improve upon these models. This necessitates the development of a useful roadway dataset, replication of existing models, and the scrutinization of additional model parameters, such as urban density and hillslope. In past studies, it was common for the researcher to select roads for which data attributes were known in advance – in this work, the sample of roads were provided by participating bicyclists familiar with the local study area; participants also discussed their perception of each roads' level of service, enabling validation of various BLOS models. Finally, an improved BLOS model was fitted to a comprehensive regional road network dataset (including over 40,000 individual links). The results clarify that, while there is yet much work to do, a BLOS model can provide a useful tool for identifying areas of the road network needing improvement for bicyclists. Closer evaluation of the model results and the roads themselves, along with participant commentary, suggest a broader need to reevaluate the use of space along American roadways, the high-speed design of these roads, as well as the protections and consideration afforded to vulnerable road users, such as bicyclists.

1 INTRODUCTION

> *A word in advance. Every cycle path is a protest against bad roads, a sort of public notice that the public wagonways are unfit for public travel, a wit-sharpener to every highway officer who has seven holes in his head, and a splendid example of the charming relations which the wheel and the roadway may be made to sustain to each other. It is a declaration of independence which, for the time being, lifts the bicycle out of the mud and puts the wheelman on a firmer ground of argument for good roads, takes from his critics the charge that the cyclist's warfare is a selfish one, and supplies to every traveler an impressive exhibition of the value of a good wheelway. – I.B. Potter*

In the United States, bicyclists were among the first to advocate for improved roadways both in and between major cities (McCullough 2015); for the League of American Wheelmen, the matter was simple: "good roads for all" (Potter 1898). A century gone by, however, and the public roadway has become an exclusive environment – a public space fully dedicated to the personal automobile (Kay 1998; Ladd 2008). It is within this environment – the public roadway – in which 4 billion metric tons of carbon dioxide are emitted annually (Schipper et al. 2011) and nearly 40,000 Americans are killed each year (NHTSA 2021). Of course, bicyclists are (at least legally) welcome to "share the road".

Some roads are better than others. In transportation planning, such a statement is typically qualified by assigning each road an estimated level of service. Methods for estimating level of service for automotive traffic are well defined, and in principle, these measures are useful for a variety of transportation services (Stopher and Stanley 2014). In reality, the level of service models for alternative modes of transportation – such as walking, bicycling, and transit – are all underdeveloped (Devasurendra et al. 2020; Kazemzadeh et al. 2020; Raad and Burke 2018; Zuniga-Garcia et al. 2018). Adequately mapping safe bicycle routes and planning future infrastructure improvements for bicyclists must be supported by robust models such as bicycle level of service (BLOS).

In efforts to stave off climate change, to reduce obesity, to prevent unnecessary roadway fatalities, to limit noise pollution, and to minimize natural resource extraction, the bicycle has regularly been suggested as a sustainable alternative to motorization (Gorobets 2016; Katteler and Roosen 1989; Lowe 1990; Parkin 2012; Tomlinson 2003). Some argue that encouraging bicycle use will depend on improving built infrastructure, such as by constructing dedicated bicycle lanes (DiGioia et al. 2017; Hull and O'Holleran 2014); these approaches, while valuable to future bicyclists, overlook the millions of people that ride their bicycles on existing roadway infrastructure, just as it is. Reconsidering the value of a BLOS rating – as well as how such ratings can be estimated and applied – offers another tool for improving the bicycling experience, and hopefully, for drawing new bicyclists to the road.

1.1 Problem Statement

The key problem with convincing more people to bicycle instead of driving is road safety. Globally, over 1.35 million lives are lost annually, with over half of those deaths being pedestrians, bicyclists, and motorcyclists; roadway collisions are *the leading cause of death* for those between 5 and 29 years of age (WHO 2018).

In the United States, the roadway fatality rate hovers around 12.4 deaths per million residents. By comparison, Vietnam, South Africa, and Saudi Arabia hold traffic fatality rates of 26.4, 25.9, and 28.8 deaths per million residents, respectively; on the other hand, Canada, Australia, and the United Kingdom boast low traffic fatality rates of 5.8, 5.6, and 3.1 deaths per million residents, respectively (WHO 2018, 302-313). Overall, the United States (US) could be considered a relatively safe place to drive, if also factoring the increased crash risk created by the disproportionately high number of miles driven by US drivers (Ecola et al. 2014, 8).

The US traffic fatality rate also includes pedestrians and bicyclists – together, non-drivers account for over 20% of all traffic deaths, and that proportion is trending upwards (Bogel-Burroughs 2019). While driver-deaths are declining, pedestrians' and bicyclists' deaths are on the rise; bicyclist fatalities rose starkly by 6.3% from 2017 to 2018 (NHSTA 2019). This is particularly startling because bicycling is not as popular in the US as in other countries: when adjusting for bicycle miles traveled per capita, the US' bicycling fatality rate is exceedingly high (*see* Figure 1; OECD 2013). Despite this risk, over 22 million Americans still ride their bicycle weekly (FHWA 2017).

Bicyclists are among the most vulnerable road users; their speeds and distances almost guarantee interaction with motor vehicle traffic, even on short trips. Current research and advocacy have identified physical separation between bicycles and motor traffic as the *best* solution to the problem at hand (Wegman et al. 2012); unfortunately, such large-scale infrastructural projects are overly idealistic for many American settings. A more immediate solution should be considered for the near term.

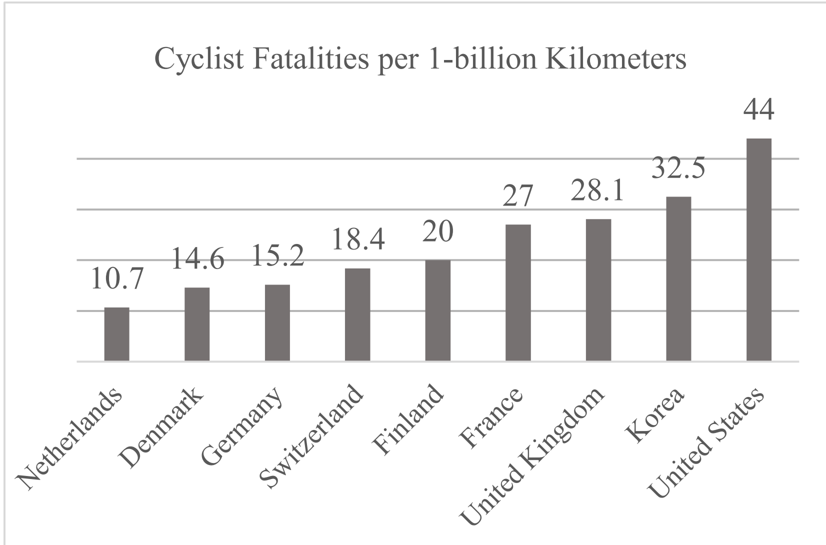

Figure 1: Bicyclist fatalities per 1-billion kilometers (OECD, 2013).

One suggestion for improving bicycle safety – both real and perceived – has been to develop and share evaluations of regional road-networks' BLOS (Huff and Ligget 2014; Klobucar and Fricker 2007; Landis et al. 1997). While level-of-service is a routine metricization for motorized planning, a strong consensus about the model specification for BLOS has been more elusive (Kazemzadeh et al. 2020; Veillette et al. 2019). Continued research is needed to clarify the variables supporting a network-scale BLOS model, both for bicyclists and planners alike (Ridgway et al. 2013).

1.2 Purpose

This research aims to develop a BLOS model suitable for use by both bicyclists and professional planners. The development of this model necessarily requires clarification of the model specification, including the selection of significant variables, their appropriate weights, and an accounting of potential interactions. In many previous studies, researchers have selected network segments for analysis based on the availability of high-fidelity network attribute data. In the proposed study, the aim is to improve upon existing BLOS models by acknowledging a wider range of segments within a regional road network; this new model will be more closely tied to the expertise of active bicyclists to select and rank road segments. These participant-driven ratings can then be utilized to evaluate existing BLOS models and to propose the most suitable BLOS model for the regional study area.

1.3 Study Questions

1. Which variables (describing the road network) are most essential for representing bicyclists' assessments of roadways, and how should those variables be mathematically arranged?

2. Which BLOS modeling strategy produces the most representative map of bicyclists' reported experiences in the local study area?

3. How do local bicyclists' experience and perceive their regional road networks level of service, beyond those attributes reflected in the BLOS model?

4. Does current regional transportation planning address infrastructural inadequacies highlighted by the proposed BLOS model?

1.4 Significance

Bicycling is the most efficient mode of overland travel; the bicycle increases human speed potential over threefold while simultaneously reducing the total energy expended (Wilson 1973). Bicycling also generates few negative externalities, with no air pollution, minimal noise pollution, and minimal risk of injury (Hurst 2009; Litman 2013, 59). Bicycling also benefits individual riders by constraining obesity (Xu 2019) and stimulating cardio-vascular health (Thorin 2017).

Though the relative modal share of the bicycle in the United States remains marginal (at around 0.5 to 1% of all trips), the absolute number of daily commuter bicyclists has risen dramatically: from about 488,000 in 2000 up to 786,000 as of the 2010 Census (McKenzie 2014). Not only has the number of bicyclists in the US risen, but so has the federal governments' investment in bicycle transportation, from a low of $50 million in 1992, now to over $1 billion in 2020 (FHWA 2020). These numbers suggest the potential for further increases in bicycle use around the country.

It is clear that we should continue to support bicycling as one component of a sustainable transportation system (Geels et al. 2017). In order to provide better guidance to current and new bicyclists and to better guide large-scale governmental investments in infrastructure, reliable metrics – and maps – about existing roadway infrastructure are required. Mapping BLOS could provide this much needed guidance.

2 RESEARCH SETTING

2.1 Scope of the Study

Improving on existing BLOS models, and by extension, developing improved bicycle network maps is suggested as a generalized solution to aiding bicyclists and urban planners to more readily identify safe and comfortable bicycling routes. To proffer improvements for existing BLOS models, this research will focus on a limited areal extent, and the participant pool will be limited to experienced bicyclists – those who either bicycle 500 or more miles per year, or travel locally by bicycle at least once per week – who are also familiar with the road network in the study area.

2.1.1 Study Area

The study area is intended to include any link in the road network which is within a day's ride of the city of San Marcos, in Texas. For most bicyclists, a reasonable ride is expected to include an area around the city roughly 50 kilometers in any given direction. This areal extent provides many interesting geographical variations (Figure 2).

First and foremost, San Marcos does not have any significant length of bicycle lanes or pathways. The city is, instead, a prototypical mid-urban US city, with a population of around 60,000. Although additional bicycle infrastructure has been planned, the new facilities have received the common criticisms expected of car drivers (Cavagnaro 2019); it is unclear whether the city's early forays into bicycle infrastructure will gain momentum or fall to the wayside. Furthermore, it is even more uncertain how city infrastructure will integrate with the larger area outside of the city's limits.

San Marcos' physiography is dominated by two distinctive geological regions. The city is centered upon the geological Balcones Escarpment: to the east lie the rolling Blackland Prairies and meandering perennial stream basins, while to the west rises

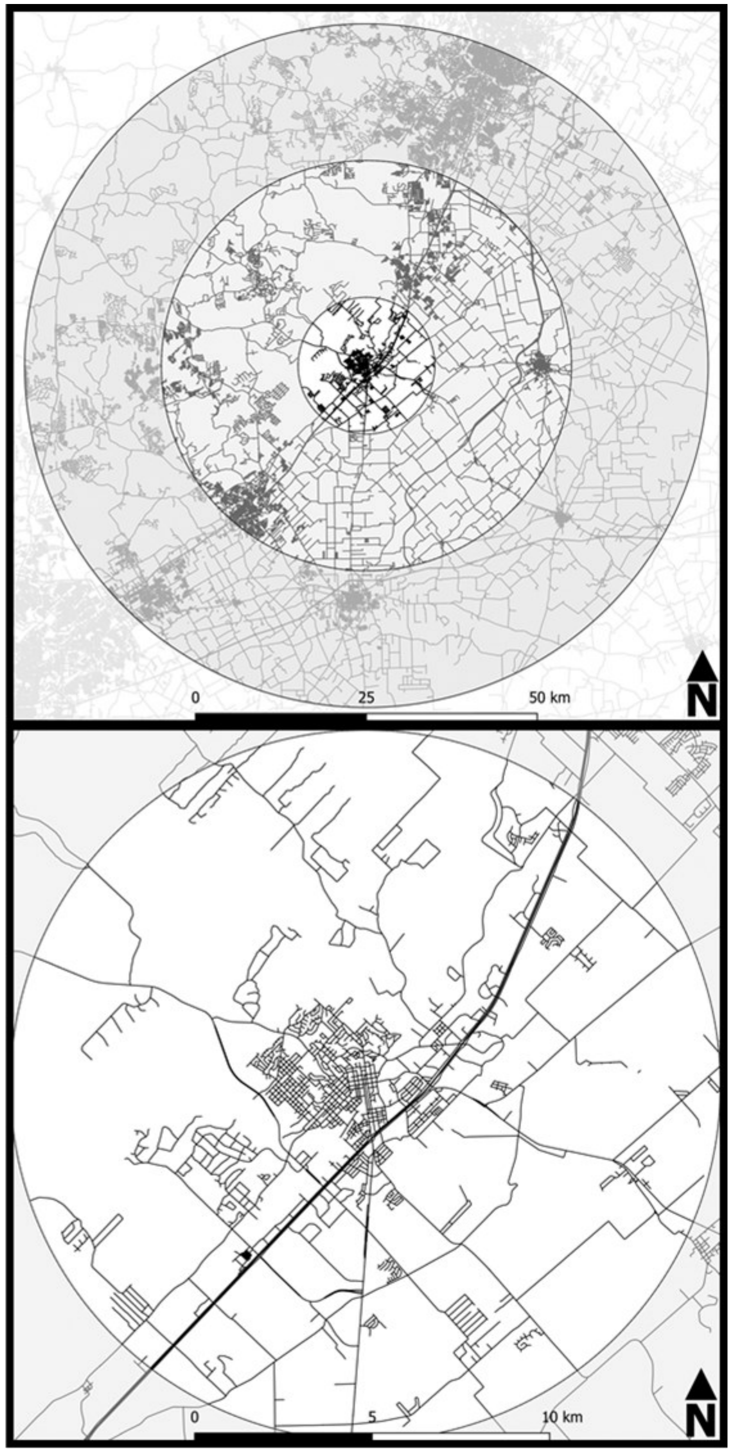

Figure 2: San Marcos and outlying areas with 10, 30, and 50km buffers.

the Edwards Plateau, more commonly known as the Texas Hill Country. Though the effect on the road network is self-evident, whether the differences in topographic roughness are meaningful to bicyclists is a matter of inquiry.

Within this areal extent, the present study aims to focus exclusively on the bicycling route network. Although this network is similar to existing road networks, it is expected to include many more connections, links, and routes than found in traditional road network datasets. Whereas automobiles have strict legal and physical limitations as to their operation, bicyclists are much more capable of going beyond curbs and pavements. Therefore, existing route network data may need to be reevaluated to account for the potential extent of the bicycle route network.

2.1.2 Limitations

There are several limitations in the development of a BLOS model. First, is that the universality of the model cannot be guaranteed. Roadway infrastructure in the United States is often more comprehensive, extensive, and of a higher quality than in many developing countries, which may limit the applicability of any US-developed model for other areas. Secondly, the expertise of active bicyclists may not be representative of the perspectives of infrequent or non-cyclists. Thirdly, the collected participant responses may not fully appreciate the demand or preference for dedicated, physically separated bicycle facilities: the local study area does not enjoy significant mileage of dedicated facilities. Nonetheless, the present work is expected to produce a usable BLOS model for mid-density urban areas, suitable for use by bicyclists and planners, and especially for United States cities and regions presently lacking in dedicated bicycling facilities.

3 LITERATURE REVIEW

3.1 Planning for Bicycles

The call for including bicycles in transportation planning began in the ‘70s but failed to gain significant traction for many decades thereafter; one of the key difficulties (and limitations) then, as it remains today, was forecasting demand to aid in planning:

> *… there are no recognized methods and little experience in forecasting such a demand and almost no experience in forecasting latent demand… Experience in cities that have provided new bikeway facilities suggests that substantial latent demand may exist (Germano et al. 1973, 17).*

It was around this time that planners began to recognize the four distinct steps of the travel demand model, including “decisions to make a trip, choice of destination, choice of travel mode, and choice of route” (Stopher 1977, 70; *see also* Turner et al. 1997a). There was also growing recognition that early models – derived from aggregated zonal data – were overgeneralized and underspecified; in reality, our travel decisions are highly individualized spatial behaviors (Ben-Akiva 1973; Golledge and Garling 2002).

One of the earliest bicycle demand models was developed in Davis, California; a stated preference survey was used to determine individual, disaggregated estimates of the choice to bicycle (over other modes) for a trip (Lott et al. 1977). Attempts to understand individual’s bicycling route choices were not far behind, but the environmental variables describing the underlying transportation network in these early works were limited to simplified and subjective categorical classifications (Axhausen and Smith 1986).

By the early 1990’s there were a handful of bicycle-specific route choice demand models, each more clearly defined than those that had come before. Although modal choice modeling was (and is) still popular, research interests began examining one of the more difficult steps of the demand model: assigning a chosen trip and mode to any one of several potential routes. Whereas automotive traffic can typically be assumed to follow the least-time routing, bicyclists are shown to be more responsive to hazards – real and perceived – and less sensitive to time impedance (Hunt and Abraham 2007). Operating under this theory, modelers in the '90s began developing safety and hazard models describing individual links of the transportation network (Landis 1994). These models would become the foundations of the various BLOS models in use today.

3.1.1 Level of Service

Level of Service (LOS) is a traditional transportation planning metric. It is a qualitative, categorical measure including designations that range from "A" to "F". Although LOS can be affected by factors such as lane width, gradient, and roadside hazards, the primary consideration for vehicular models is traffic density. An "A" rating indicates low traffic density and free-flow conditions meeting (or exceeding) the roadway's designed speeds. A "C" rating, by contrast, suggests that free-flow speeds are maintained, but that traffic density may limit safe spacing and maneuvering. Grades of "E" suggest near-peak capacities, where minor incidents might degrade service to an "F": extreme congestion or absolute gridlock (*see* Stopher and Stanley 2014). Aside from high traffic roadways, typical roads maintain a high level of service.

3.1.2 Bicycle Level of Service

The development of a BLOS measurement is ongoing, few of which include concern about bicyclists' traffic density. The majority of BLOS models center around bicycling suitability criterion, as characterized by the road and environment. Variables such as average vehicular traffic count are almost universally used in these BLOS models, but there is also a range of other variables included. The majority of these models are limited in scale – they are typically constrained to a small selection of linear segments, or links, for analysis. Fewer studies have addressed BLOS at a network scale, and participants have been primarily used for validation of existing models (using links chosen by the researcher) rather than using participants' input for model development.

3.2 Road Link BLOS variables

BLOS modeling relies on, at the very least, details about the physical roadway and the anticipated traffic; many models incorporate additional variables, including details about local area's socio-demographics, topography, and dedicated bicycle facilities. There has been little consistency in the formatting of these variables.

3.2.1 Motor Vehicle Traffic

Traffic counts are ubiquitous in bicycle demand and level of service models, though no clear standards have emerged (Kazemzadeh et al. 2020; Majumdar and Mitra 2018). In the United States, the Federal Highway Administration provides one method for the estimation of motorized average annual daily traffic (AADT), which is then measured by various state-level offices of the Department of Transportation; in some cases, regional and metropolitan planning agencies may also provide local traffic counts.

In addition to motor vehicle traffic counts, the roadway speed is very commonly used for modeling. In some instances, this might be approximated by designed roadway speed limits (Griswold et al. 2018; Landis 1994; Lowry et al. 2012). However, where data are available, the 85th percentile of observed traffic speed is recommended for modeling (Lefeve 1954; Majumdar and Mitra 2018). It is noteworthy to highlight that while the designated speed limit is an objective measure, there may be variance in statistical estimates of the 85th percentile of observed speeds (Abbas et al. 2011).

Several studies also find significant effects on BLOS resulting from the composition of motor vehicle traffic. In addition to a measure of AADT, traffic composition is typically generalized by the inclusion of either an absolute count of heavy vehicles (Jones and Carlson 2003; Petritsch et al. 2007) or as the percentage of heavy vehicles expected in daily traffic (Lowry et al. 2012; SFDPH 2010). In all models, heavy vehicle traffic negatively influences the measure of BLOS.

Motor vehicle traffic is highly dynamic, year to year, day to day, and hour to hour. While average daily traffic, 85th percentile travel speeds, and heavy vehicle counts are suggestive of actual road conditions, these measurements typically fail to differentiate traffic conditions between rush and non-rush hours, or between weekdays and weekends (Sorton and Walsh 1994). More importantly, for longer terms, these variables are not able to anticipate major adjustments in vehicle traffic – this may include reductions in miles-traveled when oil prices spike (Gillingham 2014), or from emerging modal alternatives such as motorized electric-bicycles and scooters (Plazier et al. 2017).

3.2.2 Roadway Characteristics

Variables describing the roadway are more static than motor vehicle traffic. Measurable characteristics typically include road widths, lane counts, shoulder widths, and in more extensive datasets, may also include variables such as sight distance, access-drive counts, curb-cut frequencies, and reported pavement conditions. There is limited consensus about which of these variables are most important, or how these variables are included in models.

Lane width and number of lanes are two of the most commonly used attributes of the roadway included in BLOS models. Early models evaluated just the width of the outside travel lane (the lane where bicyclists are presumed to be riding) and the number of lanes (Landis 1994). This approach was quickly modified to *effective width*, a measure that considers the width of the outside lane along with the width of the shoulder, and then adjusted for possible encroachments in the outside lane (Landis et al. 1997; Petritsch et al. 2007). Others have imputed only the highway shoulder width or potential buffer area (outside of the primary travel lanes) to estimate BLOS (Jones and Carlson 2003).

The provision of on-street parking is generally shown to reduce the bicycling suitability of a network link. Though bicyclists appear to prefer routes without on-street parking, where necessary, angled parking is preferred to parallel parking (Sener et al. 2009a); nonetheless, no precise estimate of this effect has been determined. Epperson (1994) and Landis (1994) assign a penalty to network links that have parking, using a penalty similar to that of links with moderate grades. Lowry and others (2012) utilize the proportion of occupied on-street parking within a regression model; Majumdar and Mitra (2018) also use the proportion of occupied on-street parking, but in an ordered probit model; their findings suggest that on-street parking occupancy has the strongest effect on bicyclists' perceived BLOS. This is a critical consideration in highly urbanized environments with dense on-street parking; it is important to keep in mind, however, that for many links in the transportation network, on-street parking is not permitted.

Road grade continues to be a matter of contention among BLOS modelers. Common wisdom suggests that hills – requiring more effort to bicycle – should be penalized in both the initial BLOS and derived demand models. This, at least, was the approach taken by early BLOS modelers (Epperson 1994; Landis 1994); the grade in these models is classified as either flat, moderate, or severe – it is a subjective measure when applied in these examples (Emery et al. 2003). Against the common wisdom, some research suggests that bicyclists *prefer* moderate hills, and only modestly discount steep hills (Kaplan 1975; Sener et al. 2009a). There has been limited use of grade estimates or terrain roughness indices (Lindsay et al. 2019), despite the more precise nature of these data compared to the subjective classifications often substituted for gradient.

Additional model variables describing the roadway – such as pavement conditions or qualities, roadway sight distances, access densities, and traffic-stop frequencies – are generally not available in statewide road inventories (Lowry et al. 2012; Turner et al. 1997b). Where available, the collection and coding of these data are generally limited to a single metropolitan planning region, or coded by the BLOS modeler themselves. High access densities and frequent traffic control devices reduce free-flow speeds and decrease BLOS; this term can be included in a model either as an absolute count, relative density, or factored with other attributes to form a measure of link impedance. Short sight distances are presumed to create hazardous situations, thereby reducing BLOS; sight distance determinations are from a single point, and therefore most practical in nodal analyses and often excluded from link and network analyses. Lastly, good pavement conditions may not bolster BLOS ratings, however poor pavement conditions can significantly degrade expected BLOS (Landis et al. 1997). There have been significant efforts to better understand the effects of pavement quality on BLOS (Bíl et al. 2015; Li et al. 2015; Thigpen et al. 2015), but much like sight distance, true estimates of pavement quality are highly variable within networks and within links – pavement quality potentially varies even within the area of a single node in the network.

3.2.3 Dedicated Bicycle Facilities

The presence or absence of a dedicated bicycle facility is also commonly included as a parameter in BLOS models. The definition of a bicycle facility includes a wide range of amenities, from a minimum of roadside signage, to bicycle sharrows, painted lanes, protected lanes, and at best, out-of-the-roadway dedicated paths. As Kazemzadeh and others (2020) suggest, "there is a limited body of research addressing the heterogeneity of these facilities" (*see also* Veillette et al. 2019). Despite a popular conclusion that dedicated facilities are worth the investment (Dill and Carr 2003), several studies have found that experienced bicyclists prefer lane-sharing to dedicated facilities (Hunt and Abraham 2007; *see also* Broach et al. 2012, 1738; Forester 1993). This may be, in part, attributable to the reduced passing distances afforded by drivers when bicyclists make use of on-road bicycle lanes (Beck et al. 2019; Parkin and Meyers 2010).

Although we should expect bicyclists to prefer paths without motor vehicle traffic, this would not always be true if the traffic-less facility required more stopping and starting, lengthy detouring, or an increased risk of pedestrian conflicts. Furthermore, analysis of dedicated facilities is typically limited to selected nodes and links in the transportation network – the results thus imply an idealistic view of bicyclists' stated preferences. Contemporary transportation planners should immediately recognize that a comprehensive network of dedicated facilities is rare in *most* metropolitan areas; although such facilities may represent the highest levels of service, an up-to-date BLOS *network* model cannot be limited to an evaluation of dedicated bicycle facilities.

3.2.4 Land Use

Land use attributes are included in many BLOS models as a proxy for local traffic demand and density. The earliest models applied simple penalties to BLOS for both industrial and commercial land uses; commercial areas were penalized similarly to moderate grades, while industrial areas were rated similar to severe grades (Landis 1994; Epperson 1994). These simple estimates were later substituted with more precise rankings of trip demand and cross traffic associated with different land use classes (Landis et al. 1997). Rather than penalize commercial land areas, Harkey and others (1998) produced a model which boosts BLOS for residentially zoned areas, leaving other land use classes unaffected. The binary differentiation of residential and non-residential areas has remained a popular modeling strategy (Majumdar and Mitra 2018), though more extensive classification systems have also been used (Jensen 2007).

Many BLOS models make do without land use parameters (Griswold et al. 2018; Lowry et al. 2012). In fact, this may be preferable – in effect, land use classes are intended as a proxy for traffic, however, the most basic BLOS models already include a direct estimate of traffic volume. Likewise, where the land use class is considered as a proxy for cross-traffic generation, access density and curb cut frequency would provide a more empirical measure. In 1997, Landis and others retain the land use factor "for institutional reasons" (p. 123). Jensen (2007) finds a statistically significant interaction between land use and BLOS, but ultimately concludes that the magnitude of the effect is relatively small when compared with the effects of traffic and road width on BLOS. Nonetheless, the potential significance of land use and cross-traffic should not be ignored in future BLOS model development.

3.2.5 Urban Density and Sociodemographic Factors

Urban density has not been fully examined in the context of BLOS; this may be partly because, where BLOS models are typically constructed at micro scales (of individuated links and nodes), density measures are typically aggregated to larger areal zones (Asadi-Shekari et al. 2013). These areal density estimates are used to predict the macro-demand model, both in terms of real demand and potential latent demand. Dill and Carr (2003) begin with a population density variable, though their regressions do not find the variable to be significant. On the other hand, Dadashova and Griffin (2020) find that areas with higher population densities exhibit higher rates of resident bicyclists; Griffin and Jiao (2015) also find significant interactions between bicycling traffic and the variables of activity density and regional diversity, which are independent variables reported in the Environmental Protection Agency's Smart Location database (*see also* Ramsey and Bell 2014). Turner and others (1997a) also found various density measures significant, including population density, employment density, and the percentage of land area devoted to employment use. Wang and others (2016) reaffirm the significant effects of population and employment densities on bicycle demand and use, whereas they find a suite of other sociodemographic variables to be insignificant.

The role of urban density in the BLOS model (and in the macro-demand model) deserves more attention. Although higher density suggests increased demand, results are somewhat inconclusive as to whether the increased demand is more or less than expected (given an underlying absolute increase in population). Furthermore, most models do not address variations in population density associated with changes in scale (also known as the modifiable areal unit problem); models which summarize and compare population densities of whole cities cannot be directly compared to models which estimate the population density of census block groups (*see* Buzzelli 2020). Furthermore, the use of density in zonal demand models is problematic because most bicycle trips are not work trips, hence marginalizing the traditional assumption that traffic flows from high residential density to high employment density in the morning and contra in the evening. Finally, there is a certain circularity in the models here: population density and residential density may be used to estimate total bicycling demand at the macro scale, and then BLOS models are used to assign predicted bicycle traffic to a modeled network of routes – but oftentimes, as discussed, the BLOS model *also* uses a proxy of density (land use or access frequency) to adjust the level-of-service. The risk here is that if the BLOS equation is later incorporated within a more complete demand and traffic assignment model, that a resulting unobservable and unreportable multicollinearity will exist.

3.3 Bicycle Level of Service Modeling Equations

The various BLOS models used in research and planning typically include a similar suite of data variables, however, the interaction between variables is often modeled in widely divergent forms. In most instances, the variable coefficients and factorizations are predetermined and remain static, though more recent modeling and validation efforts utilize statistical estimation of the model parameters (such as through regressions or machine learning algorithms) yielding dynamic equations.

3.3.1 Static Parameterization

By predetermining the configuration of the model, the minima and maxima of the model outputs can be defined by the analyst, and coefficients can be readily modified. Six variations of a static parameter BLOS model were considered for investigation.

Among the most streamlined of the equations (1) is given by Sorton and Walsh (1994) and restated by Turner and others (1997b). The equation utilizes a rank-classification of three key variables – traffic volume, lane width, and the speed limit. A more thorough evaluation was developed earlier, by Davis in 1987. The original equation was reproduced by both Epperson (1994, Eq. 2) and Landis (1994, Eq. 3), though adapted for metric and standard units, respectively.

Epperson (1994) and Landis (1994) both offered their own revisions to Davis' original road link evaluation. In Epperson's revised equations (2,4), the units of measure remain in the metric system (i.e., 56kph, 4.25m); meanwhile, in Landis' equations (3,5) the parameters are given in standard units. Application of either formula requires consideration of the input units and whether transformations are needed.

There are several noteworthy observations in the given equations. In Epperson's revision of Davis' equation (4), the speed factor is reapplied in the width factor, penalizing narrow and high-speed roads, while lessening the penalty for narrow roads having low speeds. Landis' first interpretation of Davis' equation (3) has a requirement that for roads over 14 feet wide the entire width factor is treated as 0, nullifying any benefit from wider roads. Epperson's interpretation (4) of the width factor avoids this concern entirely, but at the expense of two added static parameters. Landis' proposed solution (5) is rather elegant: lane-widths of roughly 14 feet do not influence the BLOS, but significantly narrower or wider lane-widths are given substantive effects by squaring.

The most recent development among static parameter BLOS models is published in the Highway Capacity Manual (HCM) (Transportation Research Record, 2010). This BLOS equation begins with 10 data variables which are then combined and transformed algebraically into 4 summative factors – width, traffic, speed, and pavements (Huff and Liggett 2014, 47). An approximation of this equation is given in equation 6 (lacking only preprocessing variables which subtly modify lane width and average daily traffic).

$$BLOS = sl(AADT) + sl(W) + sl(S) \tag{1}$$

$$BLOS = \frac{AADT}{L * 2500} + \frac{S}{56} + [(4.25 - W) * 1.635] + \sum PF + \sum LF \tag{2}$$

$$BLOS = \frac{AADT}{L * 2500} + \frac{S}{35} + \frac{14 - W}{2} + \sum PF + \sum LF \tag{3}$$

$$BLOS = \frac{AADT}{L * 3100} + \frac{S}{48} + \{S/48 * (4.25 - W) * 1.635\} + \sum PF + \sum LF \tag{4}$$

$$BLOS = \{ \frac{AADT}{L} * \left(\frac{14}{W}\right)^2 * \left[\frac{S}{30} * (1 + \%HV)^2 + \sum PF\right] + \sum LF * CCF \} * \frac{1}{10} \tag{5}$$

$$BLOS = 0.760 + (-.005 * \mathrm{W}^2) + (.507 * \log(AADT/\, 24/(4 * \mathrm{L})) + .199 * (1.1199 * \log(\mathrm{S} - 20) + .8103) * (1 + .1038 * (\%\mathrm{HV}^2) + (\frac{7.066}{\mathrm{PF}^2}) \tag{6}$$

Where,

sl() *is a function indicating rank-classification of each variable from 1 to 5, so that 1 indicates the highest suitability and 5 the least*

AADT = *Average Annual Daily Traffic*
L = *Number of Traffic Lanes*
S = *Speed Limit*
W = *Curb Lane Width*
PF = *Sum of Pavement Factors*
LF = *Sum of Location Factors*
%HV = *Percent heavy vehicles*
CCF = *Curb cut frequency*

The complexity of the HCM model makes its sensitivities more difficult to interpret. Lowry and others (2012) were among the first to test the sensitivity of the HCM BLOS, evaluating how hypothetical road improvements might improve BLOS ratings in an urban area. Huff and Liggett (2014) performed a more comprehensive assessment of the HCM BLOS, even uncovering an error in the published equations (p. 56); their results suggest the HCM model is sensitive to traffic volume, though this effect is countered by increasing road widths (the width term includes exponentiation, whilst traffic volume factors require a floor because of logarithmic transformations). However, the HCM model does not evaluate grade, is insensitive to emerging multi-modal roadway designs, and was derived from a visual preference survey of 120 participants – the Transportation Research Record deemed that improvement of the current HCM BLOS model was a key research need as early as 2013 (Ridgway et al. 2013). Performance of the HCM BLOS has not yet been tested against the established models described.

The static parameter models, as described, are advantageous in several regards. The weighting of individual factors is open to adjustment, and calibration results are readily shared among analysts. In recent decades however, these rigid models have been less frequently examined in academic literature, replaced instead with dynamic model parameterizations derived from statistical methods.

3.3.2 Model Validation and Dynamic Parameterization

What initially began as attempts to validate the early static models have evolved into an intricate formulation of contemporary models; these new-era models estimate BLOS parameters from an input dataset – typically a survey or questionnaire – to define coefficients for the available or included BLOS variables.

In 1994, Epperson proposed the videotaping of roadway segments for evaluation by participant bicyclists, a procedure which Sorton and Walsh (1994) utilized to provide confidence in the streamlined 'ranked stress-levels' index (BLOS equation 1). Thirty-two participants were divided into four categories according to their responses to questions about their bicycling experience and were then asked to rank video recordings of road segments from least to highest stress in terms of traffic, width, and speed. Traffic volume appeared to have the strongest effect on participants' responses; no significant differences were detected between the participants of differing bicycling experience, in part because of the small groupwise samples (Sorton and Walsh 1994, 23).

Landis and others (1997) set out to determine potential coefficients for factors outlined in equation 5: 150 bicyclists provided 4300 observations of an urban route featuring variable road and traffic characteristics. The best fitting model included 6 regression coefficients adjusting for factors of AADT, speed, traffic composition, land use, road width, and pavement conditions; unfortunately, the best-fit model also relied on natural logarithmic transformations of input variables, such that direct interpretation of the coefficients was impossible. To understand the model requires experimenting with various input scenarios and examining the resulting change (i.e., sensitivity analysis).

Despite the commonality of using standardized regression models, several other dynamic modeling approaches have been examined. BLOS, despite its calculation as a real number, is typically reported as ordinal ranks (where level of service can vary from A through F, or 1 through 5); this suggests a reasonable case for an ordinal regression model. Such models cannot fully define the linear relationships between the dependent and independent variables (as a coefficient describing the absolute magnitude of change in *y* given *x*), but rather, the ordinal model produces maximum-likelihood estimates – the relative odds that a change in the dependent variable (BLOS) can be expected from a measurable change in the independent variable under scrutiny (Becker and Kennedy 1992); this may still provide useful information for the development of a generalized static model. For instance, Majumdar and Mitra (2018) utilized an *in-situ* intercept survey to assess bicyclists' perceptions of road conditions; their results suggested that, at least in small Indian cities, vehicular speeds have a less significant effect than earlier presumed, and that the presence of on-street parking is more significant for BLOS than has been weighted in traditional models. Griswold and others (2018) also used an ordinal model and found that speed limits are less influential than either traffic volumes or roadway width, despite the high weighting of speed factors in earlier static models.

There does appear to be an emerging consensus as to which variables are most significant on the overall BLOS (as modeled, and as validated by bicycling participants). However, there is little consensus on how these variables should be incorporated in the model, or if all variables are even necessary: the earliest three-variable index performed relatively well in participant validation (Sorton and Walsh 1994). Although most BLOS models up to this point have undergone some participant validation (and with relative consistency), the validation methods have varied dramatically, ranging from stated preference questionnaires to video-lab evaluations, and have gone as far as on-road bicycling evaluations and *in situ* intercept surveys. While the models and variables become more consistent, there appears to be an increasing recognition in the heterogeneity of bicyclists and their preferences – these findings impose a significant challenge to calibrating BLOS models.

3.4 Challenging Models

The cross-examination of various BLOS models – including their variables and mathematical composition – reveals a series of unresolved challenges. These challenges provide the impetus for further research. A more thorough understanding of these challenges is hoped to improve BLOS estimation, and ultimately, to support a more widespread incorporation of the bicycle in the field of transportation planning – furthermore, these results may also support the production of end-user bicycle road maps and route-planning utilities, tools which may foster a modal shift towards bicycling.

3.4.1 Traffic, Speed, and Separation

Drivers and bicyclists can agree that *any* traffic is less than ideal, therefore, traffic is presumed to degrade BLOS. However, the effect of traffic on BLOS must also be considered in relation to the width of the roadway and the potential separation between motorized and non-motorized users. This relation is obvious in discussion; however, it is difficult to incorporate within the multi-factor models typically used (wherein traffic, road-width, and separation are often assigned independent coefficients). Some models do provide for interactions between these terms, such as multiplying the speed factorization by one plus the percentage of heavy vehicles (*see* equation 5), but these rarely extend to a full consideration of the extrema of interactions (Cox 1984).

The effect of these interactions in the BLOS model is best illustrated using a simple example (Box 1990). Imagine a road link within a network dataset which represents the supremum width and the supremum of separation – traffic factors, including the absolute volume and relative vehicle compositions, as well as the speed factors will both approach factor-coefficients of infima; there are no models in which this interaction is fully realized (and such realization may be impossible).

The above example is further compounded when considering the potential interaction between speed and traffic. When roadway speeds approach infimum, then traffic becomes significantly less concerning for BLOS: the traffic coefficient should also approach infimum. This exemplifies another interaction missing from most models: BLOS potentially increases dramatically as traffic approaches gridlock. This reiterates another concern with these parameters: the temporal heterogeneity of traffic volume and traffic speed conditions, in contrast with the presupposition and utilization of fixed estimates of peak-hour or average-daily traffic volumes.

Examination of roadway width, traffic volume, and speeds using a functional-network algorithm perfectly captures not only the missing interactions, but also the resulting conflict in parameter results. As the model complexity is allowed to increase, the algorithm produces both negative and positive coefficients for both road width, traffic, and speed (Beura and Bhuyan 2017); sensitivity analysis of a composite model (including all terms and coefficients from each of their four models) suggests that road width is the most significant variable, nearly double that of the effect of peak hour traffic volumes (these estimates are, yet again, presented without any interaction). Conversely, using the same dataset with a random-forest algorithm finds traffic to be the most significant model variable, with road-width a far second – the potential for interaction remains underappreciated in the final composite model (Beura et al. 2017, 92).

3.4.2 Hillslope and Grade

Few models incorporate measures of hillslope. Early models used simple classification systems, assigning penalties to hilly routes (Epperson 1994; Landis 1994). Planners in San Francisco specify little difference between grades less than 9%, but strongly penalize BLOS for segments with slopes exceeding 15%; nonetheless, their modeling still idealizes that "the lowest possible grade is always a priority" (SFDPH 2009, 9). Considering the preferences of surveyed bicyclists, conflicting results suggest that: hills impede bicyclists (Rietveld and Daniel 2004), hills have no effect on bicyclists (Moudon et al. 2005), and that some bicyclists seek out moderately hilly terrain (Kaplan 1975; Sener et al. 2009a). At the least, cartographers of bicycling maps must recognize the importance of informing map users – bicyclists – of significant grades.

Estimates of hillslope are not straightforward, immediately raising another challenge for BLOS modelers. Simply changing the scale and recording interval of the slope analysis is expected to significantly alter the resulting slope estimates (Gerrard and Robinson 1971). Even given an equal scale and interval, however, many other computational considerations remain (Jones 1998; Warren et al. 2004). Furthermore, common nomenclature varies dramatically; although *slope* and *gradient* are often used interchangeably, in mathematical convention, slope is typically an angular measure handled as a scalar variable, while gradient is expressed as the simplification of rise over run – therefore, in its fullest, gradient is understood as a vector which depends upon the direction of travel. For bicyclists, the most commonly used slope referents include the maximum gradient obtained (sustained over a short distance, such as 30 meters) or as the elevation gained per mile across an entire segment or route; still, several other trigonometric referents are conceivable. The various referents are not easily transformed (very few grade measures maintain a linear relationship). Which slope measure is best understood by map users and which measure is most appropriate for BLOS modeling has not been widely discussed.

3.4.3 Heterogeneity in Route Preferences

The majority of BLOS models are universalizing in regards to the bicycling experience – the parameters are fixed, and each potential user is assumed to agree with the output. Human spatial behaviors – including route-choices and wayfinding heuristics – are much more complex (Golledge and Garling 2002; *see also* Bovy and Stern 1990; Golledge 1999). Bicyclists are no different (Lawrence and Oxley 2019).

The response to the heterogeneity in route preferences has largely been to adopt a typology of bicyclists, for which the BLOS model could then be adapted to suit each type of bicyclist. Among the earliest of typologies was developed using in-depth interviews of 30 road-users (drivers and bicyclists), then validated with 788 survey respondents (Jensen 1999). The resulting typology derived three types of bicyclists: the passionate, the everyday, and the leisure bicyclists. Although the study did not interrogate factors underlying bicyclists' route-preferences or perceptions, the most important aspect of this typology is the clear lack of ordinality: there was no presumption of bicycling frequency or of habit, only clear evidence of oft-unaccounted heterogeneity among bicyclists.

Unsupervised statistical methods have gained significant popularity for the typologizing of bicyclists. Clustering algorithms, such as latent class choice modeling, reaffirm significant heterogeneity in bicyclists' route preferences. The latent class method simultaneously estimates respondents' most likely class membership and the parameter estimates for BLOS factor-preferences of each class; findings from this typology suggest differences between classes, but again, with no clear ordinality between class and comfort (Griswold et al. 2018). Stepwise statistical procedures have also been shown to provide robust typologies. As a first step, principal component analysis is commonly used to identify related question-blocks to represent groupwise factors; the derived factors are then used to power *k*-means clustering of participants (Chaloux and El-Geneidy 2019; Damant-Sirois et al. 2014; Gatersleben and Haddad 2010; Veillette et al. 2019). Although principal component analysis is, in theory, a useful preprocessing step for *k*-means (Ding and He 2004), the difficulty of selecting the appropriate number of principal components (Rao 1964) and the appropriate number of *k*-clusters is often overlooked (Bradley and Fayyad 1998); many typologies use subjective calibrations.

Among the most widely cited (and under-scrutinized) of typologies is the "Four Types of Cyclists" (Figure 3). The typology emerged from the Portland Office of Transportation, authored by Roger Geller in 2005 (and later revised to its current form in 2009); the typology is intended to characterize an entire population (of bicyclists, and non-cyclists as well). Attempts to reproduce the typology in participant surveys have been modestly successful, but with the looming caveat that the classification of survey respondents is itself a supervised and subjective procedure. In 2013, Dill and McNeil developed a questionnaire to classify respondents into the four types, with resulting class memberships approximating Geller's four types. In 2016, Dill and McNeil expanded their work to a national survey, and discovering that their method had wrongly classified

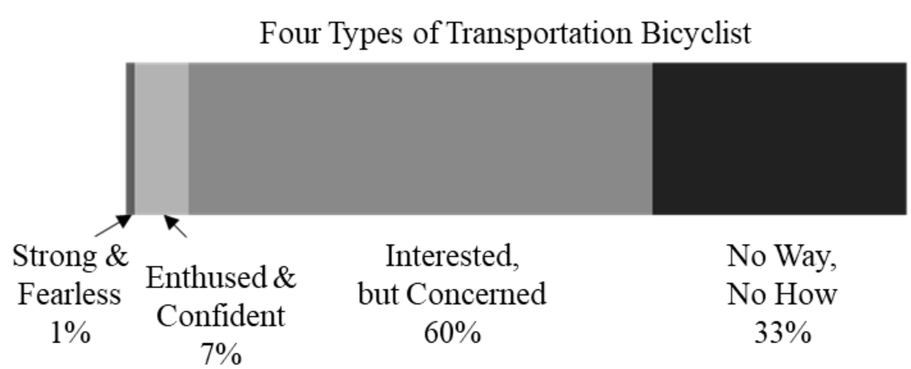

Figure 3: Four types of transportation cyclist (adapted from Geller 2009).

80 respondents (2.6% of their sample) as “No Way, No How” (non-cyclists), despite those respondents *being active transportation bicyclists*. The discrepancy resulted from the fact that *experienced bicyclists* had expressed both discomfort and concerns about bicycling safety in the questionnaire. Furthermore, a key limitation in Geller’s typology is the stated intention that the typology only be applied for utilitarian bicyclists, and not for recreational bicyclists: in validation, active *recreational bicyclists* were present in each of the four types, including the “No Way, No How” type (Dill and McNeil 2013; Dill and McNeil 2016). Similarly, in a crosstabulation including bicycling frequency and bicycling comfort set against the “Four Types” typology, only 121 of 435 respondents classified as “No Way, No How” were fully uninterested in bicycling (Cabral 2019, 56; Cabral and Kim 2020). Regardless of these potential issues (Damant-Sirois et al. 2014, 1155-1156), the four types typology has largely been interpreted as the ‘untapped potential market demand for bicycling’, leading to some interesting conclusions:

> *The goal of future policy and research would be to find the optimal ways to improve the safety perceptions of non-bicyclists when they are riding through intersections. Non-bicyclists are the largest segment of the total population. Designing and providing adequate bicycle infrastructure for non-bicyclists may eventually encourage them to bicycle for transportation (Wang and Akar 2018, 79).*

The obvious risk of advocating for expensive infrastructural improvements for users who *may never* use them is rarely discussed in such conclusions. This reifies a key aspect of this research: which roadways already represent high levels of service and what improvements might raise roadway level of service *for bicyclists*?

4 RESEARCH METHODS

Comprehensive network datasets describing BLOS are uncommon. Where it has been investigated in the past, BLOS is traditionally derived from pre-existing network attributes, such as those provided by state level transportation planning agencies; the exact model specifications have varied widely over the last several decades. This research aimed to examine the underlying model variables when applied to a mid-size United States metropolitan city, blending fieldwork with computer models (Appendix A). Alongside road network data from the Department of Transportation, experienced local bicyclists provided input and validation measures to aid in the development of a BLOS model. Model results were compared with bicyclists' assessments using pairwise differences and best fit (through least squares regression). This approach was intended to provide new insight into various model specifications and to provide future guidance for urban planners seeking to incorporate bicycling within existing road infrastructures.

4.1 Data Acquisition

BLOS is a data-driven analytical model. The most minimalist of approaches include traffic volume, road width, and motor-vehicle speeds. Additional variables may improve the BLOS model, including hillslope and urban density (proxied with measures of network density, population density, or residential and employment densities). Many of these variables are obtainable as secondary data (Table 1). Primary data was ascertained from participant interviews; this included, first, a list of road links guiding the analysis, and second, ratings of each selected link according to bicyclists' perception of traffic, road width, and speeds. The objective was to replicate and improve upon existing BLOS models while validating each model using primary participant ratings.

Table 1: Key Data Variables from Secondary Sources

Name	Source	Source Column	Description
Link_ID	Primary		Internal ID for selected links
Overall Rating	Primary		Perceived bicycling suitability
Traffic Rating	Primary		Perceived volume of traffic
Speed Rating	Primary		Perceived speed of traffic
Width Rating	Primary		Perceived width of bicycle-space
Hill Rating	Primary		Perceived hillslope impediment
Bike Lane	Primary		Painted or delineated Lane
Lane Width	TxDOT	LANE_WIDTH	Design width in feet
Shoulder Width	TxDOT	S_WID_O	Width of outside shoulder, feet
Shoulder Type	TxDOT	S_TYPE_O	Surfacing of outside shoulder
Shoulder Use	TxDOT	S_USE_O	Design use of outside shoulder (parking, bike, emergency)
Traffic Volume	TxDOT	ADT_CUR	Average daily traffic
Truck Traffic	TxDOT	TRK_AADT_P	% of heavy trucks in ADT
Number of Lanes	TxDOT	NUM_LANES	# of continuous travel lanes
Residential Density	EPA/Census	D1a	# of housing units per unprotected acre
Population Density	EPA/Census	D1b	# of people per unprotected acre
Employment Density	EPA/Census	D1c	# of jobs per unprotected acre
Road Network Density	EPA/HERE	D3a	Facility miles per square mile
Intersection Density	EPA/HERE	D3b	Intersections per square mile (auto-centric crossings omitted)
Gradient	USGS NED	*Raster*	Average gradient along link
Max. Grade 30m	USGS NED	*Raster*	Maximum grade sustained over 30 meters
Max. Grade 50m	USGS NED	*Raster*	Maximum grade sustained over 50 meters
Feet per mile	USGS NED	*Raster*	Elevation gain per mile, cumulative in both directions

4.1.1 Primary Data

In the past, research about bicycling infrastructure has typically begun with a researcher-selected list of roadway sites and links for analysis. A pivotal component of the current research was to challenge this approach by allowing experienced bicyclists to guide the research project by selecting the local road links worthy of analysis. In addition to the participants' road-links sample, they also provided perceived estimates (Likert ratings) of traffic, road width, and traffic speed, as well as some qualitative commentary; these ratings were used as validation for a suite of potential BLOS models – both those described in the literature and for those models revised during this research.

4.1.1.1 *Participant Interviews*

The interviewer guided the interviewee throughout the discussion, making regular reference to the survey instrument (Appendix B). The interviews were organized into three primary sections and completed within one hour. The first section addressed basic participant qualifications and demographic details – these details were critical for understanding the current demographic of bicyclists in the local area at the time of the study and for assessing the representativeness of this study for other geographic locales.

The second section constituted the majority of the interview, in which participants provided discussion and analysis of roads of their choosing and presumably, were most familiar. The interviewer began this discussion with a set of pre-selected sites frequented by many local bicyclists – these sites also provided for cross-validation of results amongst participants and served as a priming exercise for participant-selected roadways. Rather than determine a set number of roads set for each interview, each participant was encouraged to discuss as few or as many links as they were comfortable, as time allowed.

The only limitation asked of participants was to focus on road links within San Marcos and the surrounding region, or a radius of about 50 kilometers.

The third section allowed for participants to make more direct comments regarding the variables under discussion (road width, traffic, speed limits, and hillslopes). This enabled the direct comparison of participants' stated preferences regarding these variables with their revealed preferences (as determined from their roadway evaluations). Furthermore, their commentary highlighted insufficiencies in traditional metrics and offered enlightening discussion about, and solutions for, mapping BLOS in the future.

4.1.2 Secondary Data

Many of the necessary datasets utilized were not natively congruent. For instance, the Texas Department of Transportation (TxDOT) had maintained road line and attribute data (Appendix C), though links often differed from those defined by participants. Therefore, in some cases, secondary datasets required manual intervention and adjustment for their practical use in BLOS models.

4.1.2.1 *Road Network Data*

The TxDOT road network dataset offers the primary base map for a regional BLOS model; however, the road data in a 50-kilometer buffer around San Marcos required significant modification for modeling and analysis (Appendix D). Although the TxDOT data included traffic volume estimates (as of 2019) and road widths, maximum speed limits were absent for many roads outside of the department's jurisdiction (primarily, these were local and county roads); topological errors also prevented estimation of conventional network metrics, such as the count of intersections along links – this reinforced the importance of considering alternative sources of model variables.

Missing speed limit attributes were corrected for participants' links. The HCM institutes a minimum speed limit of 21 mph for all links, which includes missing values (Huff and Liggett 2014); this approach almost certainly underestimates the travel speeds of unsigned county and rural links, but was needed for operating the regional model.

Topological errors in road network data prevented some network-wide attribute estimates. Beyond common disjunctions and dangles, errors included disjunctions within single links. This latter issue created problems for hillslope estimation, which requires an input of continuous, single-part lines. Manual digital editing of the road network was necessary, and this editing required familiarity with advanced digitizing tools.

Other attributes utilized from the Texas roadway inventory data included the number of travel lanes, the shoulder width, use and type, as well as the surface type. Fortuitously, these variables included near complete coverage across the road network.

4.1.2.2 *EPA Smart Growth Data*

Previous research has relied on manual inventories of curb-cut frequency, land use classification, and similar attributes to specify the BLOS model. This research considered the use of more readily available data attributes from the EPA's SmartGrowth Database (compiled in 2014, using the 2010 decennial census). This research focused on the use of three measures of land-use density and two measures of road-network density, each describing conditions within a US Census block-group.

Although some road links are contained entirely within a block-group, many links intersect several block-groups, and in other cases, parallel the dividing line within block-groups. While more advanced dasymetric mapping techniques might be useful in future research, a more replicable procedure was employed. Road lines were spatially joined to block-group polygon features using a geometric “intersects” predicate, creating a many-to-one condition. The attributes of the many intersecting block-groups joined to each link were then summarized by arithmetic mean.

4.1.2.3 *Hillslope Data*

Slope and grade estimates were derived from the United States Geological Survey’s National Elevation Dataset (NED). The topology of the road network dataset required careful examination, and in some cases required manual digitization – hillslope estimation required continuous, single-part line features. Once each included link’s topology was validated, the processing method *points along geometry* was used to create a single point feature every ten meters along the link. Next, these points were input to the *v.sample* tool from the Geographic Resources Analysis Support System (or GRASS); this tool appended attributes from the underlying raster cell (such as the NED) to the input points. Lastly, algebraic calculations derived distances, elevations, and estimates of gradient along links, as well as colloquial referents, such as vertical feet gained per mile.

4.1.3 Data Summary

The data were contingent upon participants’ selection of sample road sites. Aside from this sampling strategy, however, a similar workflow would be necessary for any future BLOS modeling. Generally, the TxDOT’s roadway inventory was used as a base map, pending small topological and attribute corrections. From there, EPA SmartGrowth attributes were joined to the road lines, and hillslopes were derived from the USGS NED. Once all data were joined to the roadway features, a variety of BLOS models were developed, and finally, model results were compared against participant evaluations of road links to determine the best performing model (Figure 4).

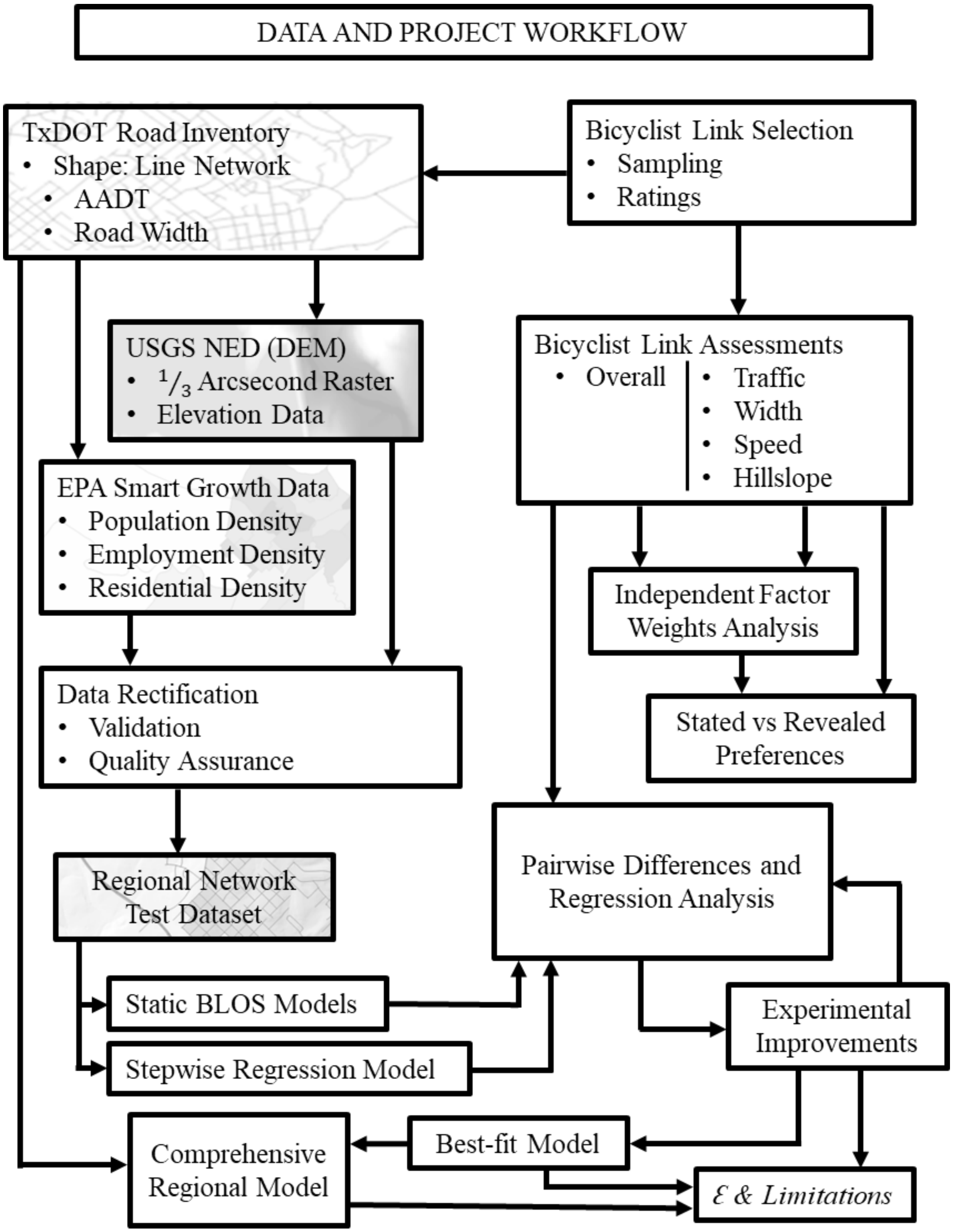

Figure 4: An overview of the data and project workflow.

4.2 Analysis and Improvement

4.2.1 Participant Analysis

The participant data were examined in several ways. First, demographic data regarding age, gender, and bicycling experience were evaluated with regard to existing surveys of experienced bicyclists; the aim was to assure a representative sample of bicyclists from the San Marcos area. Next, their individual ratings of road segments were summarized to assess potential bias and skew. Then, correlation tests were drawn between participants' ratings of width, traffic, and speed and the host of secondary data attributes that might best represent these ratings (such as average daily traffic, speed limits, and the various expressions of lane width); these correlations are presumed to assure the validity of participant responses. Cronbach's alpha and inter-item correlations were used to assure reasonable reliability of responses across participants and their sample of road sites. Finally, a series of regression models were developed to consider how each participant's ratings of width, traffic, and speed might predict a participant's overall rating of a link's bicycling suitability; this series of regression models represented each participants' revealed preference for weighting the three key factors and can be directly compared with their stated preferences during the third section of the interview.

4.2.2 Initial Modeling

Using the available data, each of the six BLOS models described in previous literature were replicated. Since each model outputs results using different numerical scales, the results from each model were rescaled to values of 1 through 5 (the underlying distribution of values remained unchanged). The rescaling procedure provided for direct comparison of model results with participants' Likert ratings and pseudo-standardized all regression coefficients (residual deviations, *RMSE,* are referential to the 1 to 5 scale).

To assess the performance of each model output, three statistical tests were considered. First, the pairwise differences were estimated, followed by a correlation and a regression analysis. The regression analyses were limited to an easily understood linear bivariate regression: model results were tested against the participants' overall averaged rating of road links to determine which model provided the best fit.

In addition to the six static BLOS models considered, a stepwise regression model was also considered. This regression utilized participants' overall ratings as a dependent variable while employing roadway attributes as independent predictors. Results from this model established the best possible fitting model (using dynamic coefficients) with which the general performance of the static parameter models could be compared.

4.2.3 Modeling Improvements

Having established the baseline performance, and selecting the ideal BLOS models, the next steps were to examine the potential improvement offered by the inclusion of additional parameters. Primarily, this effort centered around the inclusion of density measures to replace existing subjective land use classifications and the utilization of more precise referents for hillslope and gradient.

The first part of these procedures was to document the possible models and their associated fit resulting from the myriad variables and alternatives offered by measures of urban density, road network density, and hillslope. The second part of these procedures was to incorporate the most influential of these variables into the existing BLOS equations with the aim of improving the models' fit in regards to participants' ratings. Ultimately, these results provided a functional BLOS model that could, theoretically, be applied to a comprehensive regional road network.

4.2.4 Road Link Case Studies

With participant assessments and a finalized BLOS model in hand, thick descriptions were developed for a selection of key road links; these case studies were intended to provide additional insight into BLOS model results, and to highlight shortcomings in the present BLOS models by examining road-links for which the models produced large discrepancies from bicyclists' evaluations. The case studies presented were selected for a variety of reasons: because of the number of independent participants who gave mention to a particular link, because of the departure of the BLOS estimate from participant ratings, or because of other incongruous behavior by the model results.

4.2.5 Regional Model

To adhere to the overarching goals of this research, it was necessary to not only develop the idealized BLOS model, but to apply that model to the regional road network in the greater San Marcos region. This section of analysis included the careful documentation of challenges incurred and solutions implemented when applying a BLOS model to a large dataset of multiple thousand individual road links. The resulting product was a usable map of the estimated bicycling level of service in the region – a map that could provide planners with the foundations to develop a public-facing bicycle map, or to identify shortcomings in their regional bicycling network for future project development. This final model also sets the stage for a discussion about current transportation plans in the San Marcos region, as well as local bicyclists' experiences beyond the attributes of the model.

5 RESULTS

A regional model of BLOS was developed for the San Marcos region; the project was largely guided by experienced local bicyclists who provided a random sample of roads for evaluation as well as various assessments of serviceability for each road; participants' assessments were used to validate various arrangements and improvements to BLOS model equations. This chapter aims to describe the results of this process.

In addition to the final comprehensive regional BLOS model, there were several other useful results. Participant demographics and their roadway assessments are summarized, as are participants stated and revealed preferences towards road width, traffic volume, and speed limits. Six initial BLOS models are evaluated, along with a best-fit regression model; further, the use of urban density and hillslope are considered as potential avenues for model improvement. Finally, a single model is selected and applied to over 40,000 road links in the San Marcos' region; the results of this final model are further contextualized by a closer examination of several individual links.

5.1 Participant Demographics

The demographics of the bicycling community are not frequently evaluated. Generally, reports suggest that most US bicyclists are male, that the community encompasses all ages, and that most bicycling is for recreation (Kaplan 1975; Sener et al. 2009b); studies that emphasize commuting behaviors find significant differences from these general trends (Stinson and Bhat 2004; Winters et al. 2007). The interest in demographics for this research was primarily to assure a reasonable representation of the local bicycling community in San Marcos.

Descriptive demographic statistics resemble those from previous research. There were 16 participants in total. The participants ranged in age from 19 to 72 years of age; the median age was 44 and the average was 45. There were 12 males and 4 females; given the margin of error from the small sample size, the proportion of male cyclists in San Marcos could be as low as 58% (assuming 90% confidence) but is more likely between 70 and 80%. No respondents reported bicycling less than 1,000 miles per year. On average, San Marcos bicyclists rode 3,236 miles per year; this result mirrors a survey of the League of American Bicyclists: while the national average was just 2,332 miles, Texan bicyclists reported riding between 2,750 and 3,250 miles per year (Kaplan 1975).

5.2 Participant Ratings of Road Links

Participants were first asked to identify a road link with which they were familiar, then they were asked to provide ratings of that road link's overall bicycling suitability, the width of the roadway, its typical traffic volume, traffic speed, and finally, to assess each link's hilliness. Ratings were recorded on a Likert scale ranging from 1 to 5, where a 1 would represent the least suitable conditions for bicycling and a 5 would be the most suitable (Figure 5). In total, 260 individual link assessments were recorded, describing 133 unique links (54 links were selected by multiple participants).

Generally, participants selected both suitable and unsuitable road links for their assessments, with a very slight preference for higher rated roadways. This should be expected – we can assume bicyclists have more familiarity with roads they prefer riding and can recall those roads more readily, and this subsequently skews the sample of links towards those with higher ratings (Figure 6).

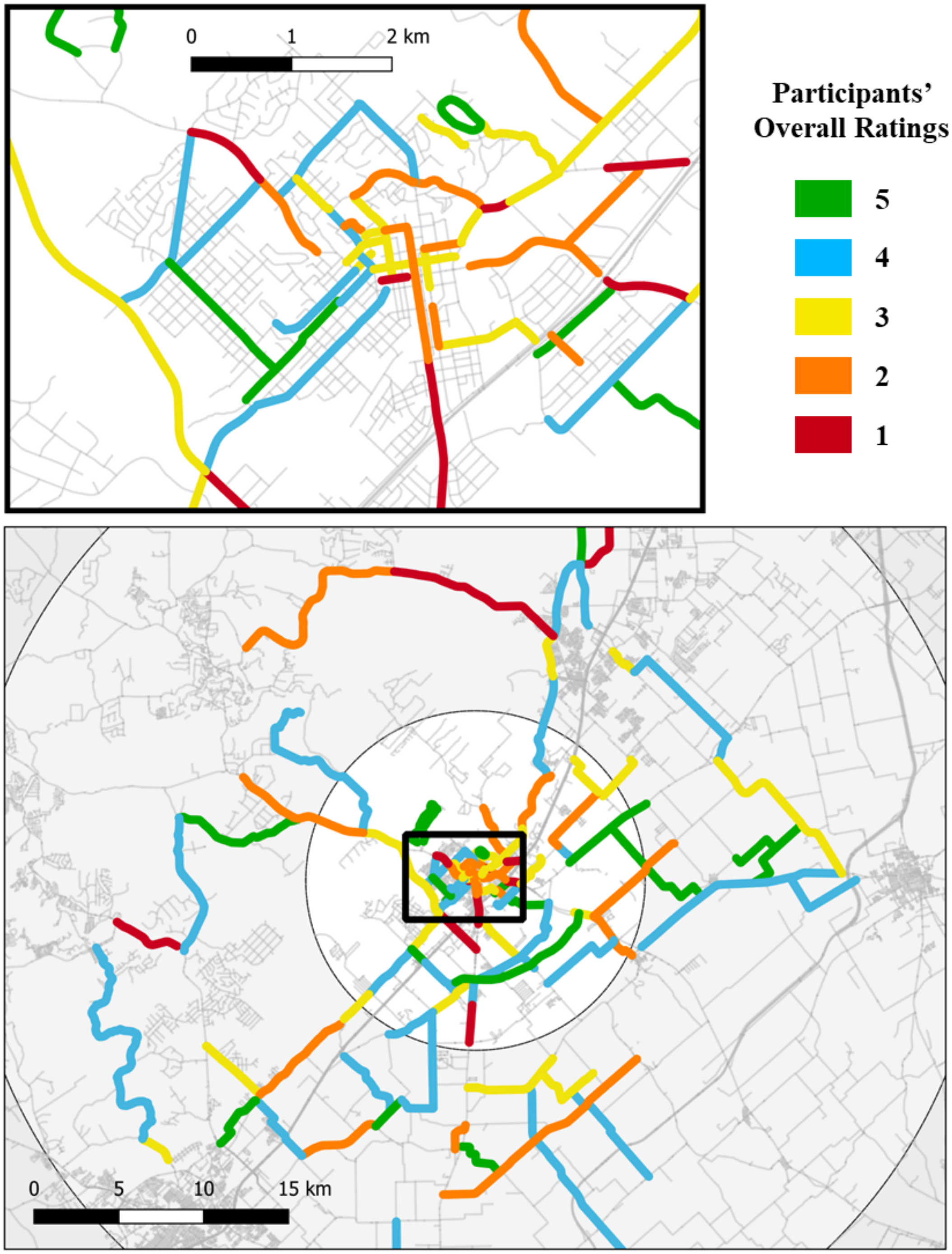

Figure 5: Map of participants' selected links and overall ratings. (Author's illustration.)

Rating	Med.	Mean	Sigma	Skew
Overall	3	3.225	1.227	0.183
Width	3	3.043	1.350	0.032
Traffic	3	3.031	1.275	0.024
Speed	3	3.143	1.108	0.129
Hills	4	3.744	1.169	-0.219

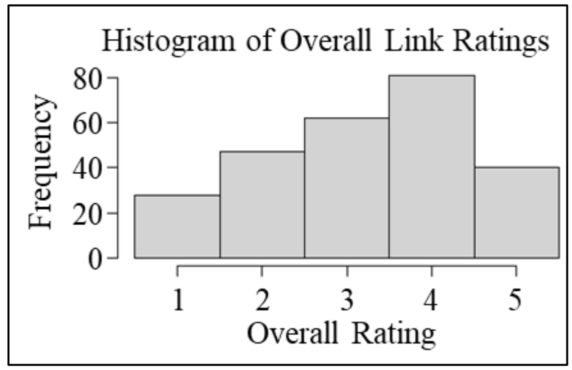

Figure 6: Summary statistics for 16 participants' ratings of 260 road-links.

5.2.1 Validity of Participant Ratings

Participant evaluations of road links were used to assess the performance of various BLOS models, therefore it was imperative to thoroughly examine the participant evaluations, as correlations, with empirical road attribute data. These results assure us that participants are indeed sensitive to differences in roadway conditions and highlight potential weaknesses in chosen model attributes (Table 2).

Traditional models have relied on various assemblies of width, speed, and traffic attributes. Of these three variables, participants' ratings were most strongly correlated in regards to traffic; the inclusion of heavy truck traffic counts slightly improved the correlation. The correlation between participant ratings and posted speed limits were also reasonably well correlated; the use of the HCM's 21 mph floor for all links offered significant improvement to the correlation. The width attribute returned a weak, but significant, correlation with participant ratings. The various assemblies of the width term confirm the importance of both lane and shoulder widths, but do not clarify whether it would be best to use the two attributes in sum, or to consider them as distinct, independent model factors.

Table 2: Correlations between participant ratings and road attributes

	Participant Ratings			
Secondary Data	**Width**	**Speed**	**Traffic**	**Hills**
LANE & SHOULDER WIDTH	0.191	--	--	--
LANE WIDTH	0.075	--	--	--
ROAD WIDTH	0.029	--	--	--
SHOULDER WIDTH	0.295	--	--	--
MAX SPEED	--	-0.330	--	--
MAX SPEED (Floor 21mph)	--	-0.387	--	--
CURRENT AADT	--	--	-0.594	--
AADT * (1 + %Heavy)	--	--	-0.595	--
AADT * (1 + %Heavy)^2	--	--	-0.596	--
Feet per Mile (2-ways)	--	--	--	-0.635
Sum Elevation Gain (2-ways)	--	--	--	-0.228
Max Elevation Gain (1-way)	--	--	--	-0.265
Max. Gradient (1-way, 30m)	--	--	--	-0.459
Max. Gradient (1-way, 50m)	--	--	--	-0.519
Modeled Terms				
HCM OUTSIDE LANE	-0.199	--	--	--
HCM SPEED TERM	--	-0.280	--	--
HCM TRAFFIC TERM	--	--	-0.620	--
BLOS 1: Width Classes	0.212	--	--	--
BLOS 1: Speed Classes	--	0.331	--	--
BLOS 1: Traffic Classes	--	--	0.442	--

Prior to this research, most BLOS models relied on simple (and often subjective) classifications of hill slope. The ability to derive precise hill slope estimates for road links could be highly beneficial to future research, as the correlations between participant ratings and empirical measures of hill slope demonstrate. The two most highly correlated hillslope referents are first, the elevation gained per mile (as feet per mile) along the road link in both directions and second, the maximum sustained gradient over 50 meters. While maximum gradient expresses the worst possible hillslope a bicyclist may encounter, this is often true only for one direction (uphill). Feet per mile may provide a more useful metric, both for modeling purposes and public communication.

In addition to correlations with secondary data attributes, the participant ratings are directly relatable to both the BLOS model proposed by Sorton and Walsh (1994) and with the HCM BLOS model – both model equations are broken down as factors of width, speed, and traffic. The former model relies on reclassification of interval data into ordinal values, while the latter includes more complex transformations and interactions of attribute values. The HCM traffic term outperforms all other traffic metrics, while the ordinal classification of width performs similar to empirical measures of lane width.

5.2.2 Reliability of Participant Ratings

In the previous section, correlation results established that participant ratings are reasonably valid in relation to empirical roadway attribute data; however, this does not address the reliability of participant's ratings. To evaluate the reliability of participant's Likert scale ratings, two measures of internal reliability – or consistency – were considered: Cronbach's alpha and the mean of inter-item correlations (Table 3). Reliability was assessed primarily across four key survey items: participants' rating of road links' overall bicycling suitability, as well as link width, traffic, and speed. Reliability was also reexamined with the inclusion of participants' hill slope ratings.

Cronbach's alpha was greater than the commonly cited threshold value of 0.7; this indicates reasonable reliability across the measures and is especially strong considering the limited number of items included in the assessment (p = 4). Mean inter-item correlations were also satisfactory, indicating reasonable cross-item reliability without the risk of overly high correlations (and the potential for multicollinearity).

The inclusion of hillslope decreased both alpha and the mean correlations. Although the values were still within generally accepted ranges, the decrease could indicate high variability and low reliability of participants' hill slope ratings.

Table 3: Cronbach's alpha and inter-item correlation for participants' ratings

Reliability Measure	Result	Items
Cronbach's alpha	0.785	4
--- w/ Hill Rating	0.721	5
Mean Inter-item Correlation	0.489	4
--- w/ Hill Rating	0.343	5

5.3 Participants' Overall and Factor Ratings in Regression

Although the survey instrument concludes by asking participants to rank the importance of width, traffic, and speed factors, it is also possible to estimate not just ranks, but precise factor weights using linear regression. The equation is described as:

$$\text{Overall Rating} = b_0 + b_1(\text{Speed}) + b_2(\text{Traffic}) + b_3(\text{Width}) + b_4(\text{Hills})$$

The resulting coefficients, $b_{0\text{-}4}$, can be considered as factor-weights for adjusting BLOS models; furthermore, the regression is also useful for highlighting roadways where predicted overall ratings diverge from participant ratings through a residuals analysis.

5.3.1 Regression Results

The initial regression specification returned significant results for all coefficients, including participants' hill ratings. However, since hillslope ratings exhibited significant skew and reduced inter-item reliability, the regression was also respecified without the hillslope parameter to emphasize the determination of coefficients for the remaining three factors: speed, traffic, and width (Table 4); both regressions have been standardized so that the coefficients can be interpreted as factor weights.

The most significant factors to predict participants' overall ratings were, in order, traffic volume, then roadway width, followed by roadway speeds. If hillslopes are considered, they would marginally out-rank speed as a factor in ratings. Although the regression models leave unaccounted variability in participants' overall ratings, the models do fit reasonably well.

Table 4: Two regression models using only participant response data

Factor	Coef	p-val
(Intercept)	-0.208	0.30
Rated Speed	0.121	0.03
Rated Traffic	0.536	0.00
Rated Width	0.300	0.00
Rated Hills	0.138	0.00
r-sq. = .6500; p-val = .0000		

Factor	Coef	p-val
(Intercept)	0.238	0.14
Rated Speed	0.137	0.02
Rated Traffic	0.530	0.00
Rated Width	0.312	0.00
r-sq. = .6330; p-val = .0000		

5.3.2 Residual Analysis of Participant Factor Regression

The residuals represent the divergence between participants' overall road ratings and the fitted regression estimate as determined by participants' evaluations of width, traffic, and speed. The residuals were normally distributed around a mean of approximately zero and were consistent even at the extrema of the model (Figure 7). This result provides additional confidence in participants' assessments of each individual factor, and in their assessment of each link's overall level of service.

For roadways where the participants' overall rating exceeded the regression model result, the disparity was typically caused by participants assigning a low rating to either width or traffic. This suggests that participants' overall ratings are dependent upon a complex interaction of width and traffic which cannot be adequately modeled by the linear regression.

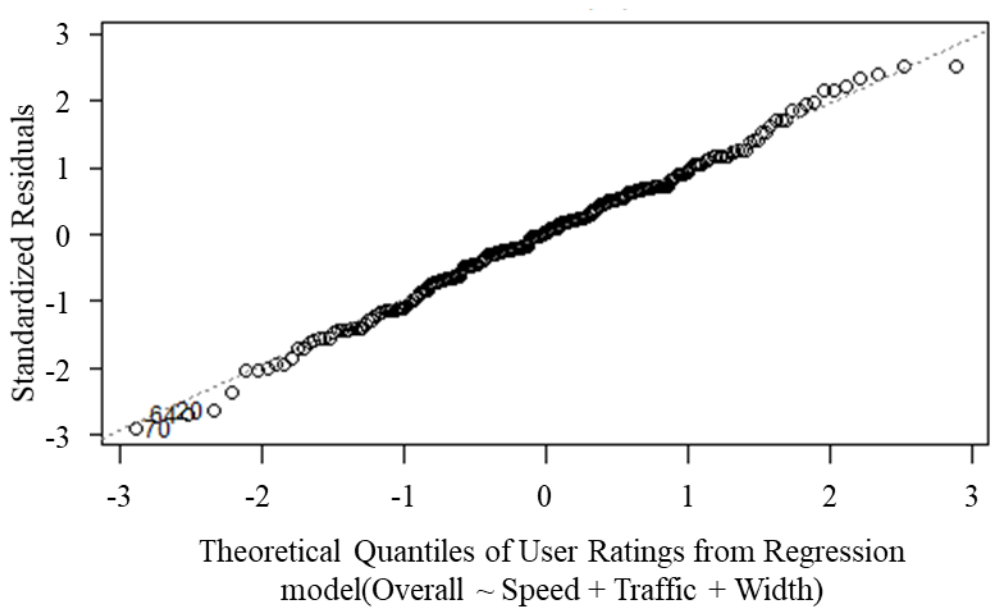

Figure 7: Normal Q-Q plot of predicted user ratings and residuals.

Roadways where the fitted rating exceeded the participants' overall ratings were often accompanied by unprovoked commentary. Largely, participants downgraded their overall ratings intentionally while considering unmeasured roadway attributes. Among the most popular of these comments included references to poor roadway surfaces, but participants also mentioned limited sight distances, distracted driving, and low water crossings as reasons for penalizing a roadway's overall score.

5.3.3 Stated vs Revealed Preferences for Factor Prioritization

Near the end of the survey, participants were asked to rank the three key factors – width, traffic, and speed – from most concerning to least concerning. The responses varied dramatically. A majority of respondents stated width was their first concern, however for others, width was the least concerning factor; the second most important factor by statement was traffic; if traffic was not the most concerning factor, then it was often second. Speed was generally a participants' second or third priority.

Linear regression coefficients of the three factors' contribution to overall ratings were determined for each participant, revealing everyone's personal weighting of speed, traffic, and width; the average *r-squared* value across participants was approximately .75, but ranged from as low as .21 to as high as .95 (Table 11, Appendix D). The regression coefficients were transformed into ranks (for each individual participant) for direct comparison with participants' stated preferences (Table 5).

The result of this analysis suggests that bicyclists' stated preferences for the prioritization of width, traffic, and speed are comparable to their revealed preferences. In both stated and revealed preferences, vehicle speed was the least concerning factor. The importance of width, however, might have been overstated, as revealed preferences suggest traffic volume is generally the most important factor, with width a close second.

Table 5: Participants' factor priorities, mean of rankings (1st priority to 3rd)

Factor	Stated	Revealed
Speed	2.25	2.44
Traffic	2.13	1.69
Width	1.63	1.88

5.4 Initial Bicycling Level-of-Service Models and Model Fit

Existing literature included many algebraic equations for deriving BLOS from empirical roadway attribute data. These equations were adapted to data available in the TxDOT roadway inventory, then processed to derive level-of-service estimates for each road-link assessed by participants (Appendix D). The estimates from each model were compared with participants' overall ratings using three statistical tests: pairwise differences, correlation, and bivariate regression (the sample had 131 links, as 2 links – interstate frontages – were withheld as outliers; results tables are shown in Appendix E).

5.4.1 BLOS from Model Equation #1

The first model considered relies on a rank-order classification of a road link's average daily traffic, posted speed limit, and effective curb lane width. Although some coding is required to provide the rankings, the equation is rather elegant in its simplicity.

Although this model exhibited only small pairwise differences from participants' overall ratings, neither the correlation nor regression were particularly strong. Attempts to weight the three factors – including using regression coefficients obtained from participant ratings – failed to encourage significant improvements in the model results. Ultimately, this model cannot compete with the more rigorous BLOS equations.

5.4.2 BLOS from Model Equations #2 and #3

Model equations 2 and 3 are relatively similar in form; the major distinction is that equation 2 anticipates metric inputs while equation 3 utilizes standard inputs. Another major difference was that in equation 2, the factors were additive, while in equation 3, the factors were originally published with division symbols – experimentation suggests this was only a misprint; when using addition of the factors, the two models perform nearly identically (*see* Landis 1994 and *see also* Epperson 1994).

In addition to the traffic, speed, and width, BLOS equations 2 and 3 require the number of traffic lanes, as well as a list of subjective modifiers (as a list of pavement and locational factors). First, both the standard and metric equations were processed without these extraneous locational factors. Second, 6 factors were modeled programmatically from roadway and density attributes (Table 6); though these additions reduced pairwise differences between the BLOS model estimates and participants' ratings, their inclusion only moderately improved the correlation and regression results.

Table 6: Programmatic application of locational factors.

	Logic	Modifier
Intense Land Use	if D1a > median(D1a)	+ .50
Freq. Curb Cuts	if D3b > median(D3b)	+ .50
Moderate Grade	if ft_mile > median(ft_mile)	+ .25
Severe Grade	if ft_mile > 3rd.Quartile(ft_mile)	+ .50
Angled Parking	if S_USE_O == 1	+ .75
Parallel Parking	if S_USE_O == 2	+ .50
Shoulder/Bike Lane	if LANE_WIDTH_O - LANE_WIDTH > 0	- .75

5.4.3 BLOS from Model Equation #4

Equation number 4 represents a transformation of Davis' original model, with the addition of interactions between speed and width, as well as a minor revision to the modifiers associated with the roadway's location. Despite the theoretical improvements offered by this equation, this model performed among the worst. Pairwise differences were greater than 1 and the predictive power of the regression was completely marginalized. The revised modifiers for locational factors had little effect – the revised list of BLOS modifiers increased pairwise differences and failed to improve either the correlation or regression results for equation 4.

5.4.4 BLOS from Model Equation #5

Model equation 5 requires more data than previous models, but also improved upon their results significantly. The fundamental basis of the equation remains width, traffic, and speed with added sensitivity for heavy truck traffic and pavement condition (although pavement condition is unreported in TxDOT roadway data and unused here). The equation includes just 2 additional modifiers, though they must be assigned non-zero values: land use intensity must be rated using values of 1 (non-commercial) to 15 (commercial), and curb cut frequency must be rated from 1 to 200 (as an approximation for number of curb cuts per mile). It is possible to operate the equation by setting placeholder constants (such as fixing land use at 15 and curb cuts frequency at 42); the use of constants establishes the model's basic performance.

Even when utilizing fixed values for land use intensity and curb cut frequency, BLOS equation 5 outperforms each of its predecessors in terms of regression fit and correlation. Unfortunately, it was the only model to exhibit significant positive pairwise differences, suggesting that the model tends to overestimate bicycling serviceability. There remains some potential for improvement by specifying dynamic values for land use and curb cuts, such as by using transformations of density to fit the expected input values.

5.4.5 BLOS from Model Equation #6

Equation 6, from the Highway Capacity Manual, requires careful attention to the preprocessing of variables because of its inclusion of logarithmic transformations. Some of these steps were simplified and still yielded satisfactory results; ultimately, the model is no more complex than equation 5 and appeared to perform about as well. The equation benefits from its inclusion of the percentage of heavy truck traffic, but it does not provide any account for locational factors, which might offer one path to improving the model. Though the regression result was slightly weaker than equation 5, the pairwise differences were reduced and negative – indicating mostly conservative model estimates; overall, this model was modestly successful, but nonetheless, may yet benefit from further revision – either through calibration of existing parameters or by inclusion of new parameters. Calibration of existing parameter coefficients may be difficult, as the coefficients already have precision to the ten-thousandths decimal place.

5.4.6 Dynamic Parameters: BLOS from Regression

Most recent research has eschewed the fixed parameter models documented thus far in favor of dynamic parameterizations, most commonly utilizing multivariate linear regression. Although this procedure makes experimental replication more difficult (as the parameter coefficients change with each new replication), it does provide strong evidence for determining the most influential roadway attributes for any possible BLOS model. Here, a stepwise regression procedure was used to determine the most likely predictors (from roadway attributes) for the current participants' overall link ratings.

In general, various combinations of predictors yielded models with *r-squared* values between .25 and .35, consistent with the fixed-parameter models outlined above. The most useful regression model consisted of five variables (Table 7); the inclusion of additional variables, such as heavy truck traffic, could increase the model's fit but produced poor overall performance and statistical artifacts (such as a negative coefficient for increasing lane width). The two most influential roadway attributes were average daily traffic and effective curb lane width, followed closely by population density, and finally, the maximum posted speed limit. The coefficients for width, traffic, and speed closely resemble participants' stated and revealed preferences, thus reiterating these factors' importance. Although perceptions of density were not measured during the participant survey, the usefulness of density for improving the fitted level-of-service estimates should not be overlooked. Finally, the inclusion of a binary flag representing the presence of a bike lane is crucial to the regression model – although this variable is not readily available in current roadway inventories, the role of dedicated facilities cannot be ignored in future modeling efforts (nor in future roadway redevelopment projects).

Table 7: Best fit BLOS regression model.

Variable	**Coefficient**	***p-val***
(Intercept)	5.81	0.0000
Max Speed	-.011	0.3077
A 10 mph increase in speed reduces effective BLOS by .11		
log(Lane AADT)	-.374	0.0000
A 10% increase in traffic volume reduces BLOS by .03		
Effective Lane Width	.079	0.0193
A 1-foot increase in outside lane width increases BLOS by .08		
Population Density	-.096	0.0000
A 1 person per acre increase in density decreases BLOS by .10		
Bike Lane	1.21	0.0000
The inclusion of a dedicated bike lane raises BLOS by 1.2		
Model Summary:	*Regression Model r-squared = .357* *Residual Standard Error = .964*	

5.4.7 Summary of Modeling Results

Of the six models assessed, only two are recommended for ongoing investigation: model equation 5 as revised by Landis (1994) and model equation 6 as presented in the Highway Capacity Manual (Huff and Ligget 2014). The correlation between participant ratings and the Landis model was an *r* of .48, while the HCM exhibited an *r* of .38. The regression results were closer to an *r-squared* of .2; this result suggests there are still unaccounted variations between bicyclist's overall ratings and the model results.

Finally, although both equation 5 and 6 produced similarly strong correlations and modest regression results, equation 6 minimized pairwise differences ($\bar{d}$ = -0.4), and remains conservative in doing so (often underestimating roadway ratings). Equation 5 unfortunately exhibited significant pairwise differences, of which most were positive: equation 5 overestimated the bicycling level of service from roadway attributes as compared with participants' ratings. The stepwise regression procedure suggests that there is yet room for improvement in each model: density and locational factors are important predictors that are absent from equation 6 and were subjectively inventoried for inclusion in previous applications of equation 5.

5.5 Improving the Best Fit Models

Initial experimentation with six published BLOS models suggested that there are several approaches with performance rivaling a best-fit regression (and utilizing static, replicable parameterizations). As presented, however, those models were often reliant on data attributes that are unavailable in many roadway inventories or alternatively, the models discounted factors which have strong theoretical basis for inclusion. The results thus far suggest adequate space for significant improvements to these existing models. While recalibrating the factors of width, traffic, and speed offers one avenue for improvement, the inclusion of additional parameters may yield more substantial success. For instance, the inclusion of density is suggested based upon theoretical grounds: more densified locations amplify cross-traffic generation and aggravate bicycle-vehicle interactions. Though empirical estimates (such as curb-cut frequency and access density) may be preferred here, density estimates from the US Census offer more complete data coverage with significantly less labor (an undeniable benefit for many planning organizations). The inclusion of a hillslope parameter represents another potential improvement, though obtaining this data is not as straightforward, nor is selecting the appropriate metric from the available referents (such as grade, elevation gained, or elevation gain per mile). Other possible improvements include the revision and calibration of additive location factors, an opportunity that is overlooked in equation six.

5.5.1 Density

Areal density is one possible attribute that might improve existing BLOS models. One challenge, however, is the selection of *which* densities offer the best improvement. BLOS equation 5, which includes both *land-use intensity* and *curb-cut frequency* as variables, provided an ideal framework to examine the inclusion of density through variable substitution. Density values were first transformed to match the expected input for equation 5 (using 1 to 15 for land use, and 1 through 200 for curb cut frequency).

Land use intensity was replaced with the use of population density, employment density, and residential density; meanwhile, curb cut frequency was substituted with both intersection density (intersections per lane-mile) and road-network density (lane-miles per square mile). These potential inputs provided 12 distinct models. There were significant and meaningful improvements to both pairwise differences and the correlation results. The addition of density values reduced pairwise differences by as much as 40% (from .77 down to .46) and strengthened the regression results by as much as 30% (from an initial *r-squared* of .23 up to an improved *r-squared* of .30).

Potential similarities in areal density raised the concern of multicollinearity. To that end, a correlation matrix of the various densities was produced; while some densities were highly correlated (residential density and road network density exhibited an *r* of .91), employment density was only moderately correlated with roadway densities (with road network density, $r = .54$, and with intersection density, $r = .52$). Nonetheless, the inclusion of road density (while retaining a fixed value for land use intensity) alleviated any concern of multicollinearity, significantly reduced pairwise differences, and strengthened the regression's fit ($\bar{d} = .46$, *r-squared* = .30; *see* Appendix E).

5.5.2 Hillslope

Five possible ways to report hillslope were initially considered: maximum gradient sustained over 50 meters, maximum gradient over 30 meters, maximum elevation gained in one direction, the sum of elevation gained traveling a link in both directions, and finally, the elevation gained per mile traveling both directions along a link. Correlations with participant responses, presented earlier, suggested that either feet per mile or the maximum gradient over 50 meters provided the best representation of participant's responses.

BLOS equation 5 was used to examine the potential for including hillslope: the pavement condition parameter remained in the equation but had held a fixed value (since the attribute was unavailable in the roadway inventory). Possible values for pavement conditions ranged from 1 to 5, with 5 representing better conditions; therefore, the hillslope referents were rescaled and inverted to match this range (the flattest route was given a five, while steeper hills received progressively lower ranks). A simpler method of applying penalties for hills with excessive maximum gradients was also considered: routes with gradients over 13% were assigned a 2, while all other routes received a 5.

The inclusion of hillslope in equation 5 produced only modest changes in the results, with no noticeable differences between the use of either maximum gradient or elevation gained per mile. The simplified penalty for excessively steep hills performed about as well. Regardless of whether hillslope measures are retained in the BLOS model equations, such measures remain valuable detail for bicyclists' route planning – hillslope information was among the most common requests by participants for inclusion on future bicycle map products.

5.5.3 Improvements to Equation 5

Although the task of preparing density data is not trivial, it is more reliable than other ratings of land use intensity or field counts of curb cuts. Although density's contribution to the model remains small, there remains strong theoretical basis for its inclusion: density anticipates more intense cross-traffic generation. Multicollinearity may be a concern when drawing upon census densities, as the combination of road-network density and population density were moderately correlated ($r = .77$); however, their inclusion produced satisfactory improvements. Three additional calibration coefficients (A1 through A3) remain unexplored in this research; investigating their sensitivity and effect on BLOS model results is a suggestion for future research.

5.5.4 Improvements to Equation 6

Model equation 6 relied on assumptions about the data inputs which often required additional data management. First, lane-width is expressed as the total effective width of the outside travel lane (inclusive of the bike lane and outside paved shoulder, except where this additional shoulder width is designated for on-street parking). Next, traffic counts should be expressed as hourly demand, and the minimum value of this hourly traffic count should be no less than 4 times the number of through-traffic lanes. Third, the speed limit attribute must be set to a minimum of 21 mph for all road links. Significant improvements came about from the inclusion of density and hillslope. Density inputs were transformed to values from 1 to 5 (5 representing higher density); hillslopes were treated similarly, but also inverted (so that 5 represents shallower gradients). High densities and steep hillslopes penalize the BLOS result. The resulting equation appears complex, but this is mostly a result of lengthy parameter coefficients.

The output of the revised Highway Capacity Manual BLOS model exhibited mean pairwise differences of just -0.19 and achieved an *r-squared* of .23. The r-squared is only slightly less than the best achieved across all models (the best fit here would be among the permutations of Equation 5, although those models collectively failed to minimize pairwise differences). The improvement in pairwise differences with the finalized model here is appreciable – the observed differences between these final model estimates and participant ratings were not statistically significant.

5.5.5 Summary of Improvements and Final Equations

The inclusion of additional parameters representing location density, hillslope, and the presence of a bike lane each provided subtle improvements to the overall fit of the BLOS models. The results suggest that these factors are necessary for developing the most representative model, but also reveal that there are yet significant and unresolvable uncertainties in the models.

The stepwise regression model suggested the importance of location density for estimating BLOS. The inclusion of residential density proved more influential than the use of road network or intersection density, however, it may be appropriate to include both residential and road network density (as was done in the revised BLOS Equation 5). At the least, where density is unavailable, a proxy for cross-traffic must be considered.

Hillslope was less effective at improving model results, but should not be discounted in future efforts to model BLOS. Recall the regression of participant's overall ratings as predicted by their individual factors of speed, width, traffic, and hills – the inclusion of hillslope was nearly as influential on the results as was participants' rating of the speed of traffic. How best to refer to hillslope, what gradients might be considered excessive, and how influential hills are to bicyclists' route choices remain questions largely unanswered, leaving intriguing paths for future research (*see* Appendix F).

The need for identifying bike lanes when modeling BLOS might seem intuitive, but raises a new host of challenges. State roadway inventories are yet to reliably include even basic bike lane attributes, while bike lanes themselves vary widely in design, separation, length, width, and continuity; in some cases, in San Marcos, a bike lane exists on one side of the roadway but not the other! (These are often logically placed, but difficult to model appropriately: bike lanes offered on uphill links with only lane "sharrows" for opposing downhill bicycle traffic.) Despite concerns about the availability of specific detail about individual bike lanes, the simplest binary flag, as used here, provided significant model improvements.

With the improvements suggested here, revised model equations 5 and 6 each excelled in particular ways. The revised version of Landis' (1994) equation (eq. 7, derived from previous model equation 5) exhibited the best fit in regard to participants' ratings, while the revised Highway Capacity Manual model (eq. 8, derived from previous equation 6) continued to minimize pairwise differences. Both equations require similar data inputs and similar preprocessing steps. For planning applications, the surest approach would be to produce BLOS models using both equations for the desired extent, and then to evaluate each model's performance with spot-checks. If the intention is to use these models to produce a public-facing bicycle map, then extra diligence should be warranted. Moving forward, BLOS equation 8 will be examined more thoroughly; this model was favored because of its minimization of pairwise differences and its relatively conservative estimates when compared with participant's evaluations.

$$BLOS = (AADT/NUM\_LANES) * (14/LANE\_WIDTH\_O)^{\wedge}2 * \\ A1 * (SPD\_MAX/30) * (1 + \%HV)^{\wedge}2 + \\ A2 * (1/FT\_MILE) + \\ A3 * (POP\_DENS * ROAD\_DENS) * .1 \quad (7)$$

Where,

FT_MILE =	*Inverted and rescaled from 1 to 5; steeper slopes get lower scores.*
POP_DENS =	*Rescaled from 1 to 15; high density gets high scores.*
ROAD DENS =	*Rescaled from 1 to 200; high density gets higher scores.*
	A1 = .01, A2 = .01, A3 = .024

$$BLOS = .760 + (-.005 * LANE\_WIDTH\_O^{\wedge}2) + \\ .507 * log(ADT\_HR/(4 * NUM\_LANES)) + \\ (0.199 * (1.1199 * \log(SPD\_MAX\_FLR - 20) + .8103) * \\ (1 + .1038 * TRK\_AADT\_P)^{\wedge}2) + \\ -0.5 * (1/POP\_DENS)) + 0.1 * (1/FT\_MILE)) + \\ -1 * BIKE\_LANE \quad (8)$$

Where,

FT_MILE =	*Inverted and rescaled from 1 to 5; steeper slopes get lower scores.*
POP_DENS =	*Rescaled from 1 to 5; high density gets high scores.*
ROAD_DENS =	*Rescaled from 1 to 5; high density gets higher scores.*
BIKE_LANE =	*Binary, 1 indicates dedicated bicycling facility.*

5.6 Comprehensive Regional Model

A BLOS model is a planning tool; in traditional applications, deployment of the model would require a laborious survey of road sites and meticulous calculation of results. The aim of this research has been, not only to compare and validate existing models, but also to consider the potential for applying such models across larger regional networks with geographic software. The final step of this research was to apply the revised BLOS equation to the comprehensive regional roadway inventory for San Marcos. The final product is a useful map of the region which highlights weaknesses in current roadway infrastructure, and by extension, the bicycling network (Figure 8).

The roadway inventory included over 40,000 links; although applying the BLOS model equation to the attribute data was straightforward, there were some spurious results from extreme attribute values. One initial step was to remove interstate highways (some with as many as 8 to 10 lanes) from the model's consideration – this was accomplished by eliminating any links with a federal functional classification of 1. Next, shoulder and lane widths were both limited to 16 feet; more extreme values were not observed in the participants' roads sample, but several hundred links in the regional dataset included lane and shoulder widths as high as 40 feet. Even with these preprocessing steps, some extreme BLOS values were still produced; outliers were constrained using a floor and ceiling on the model results using bounds of the lower and upper ventile, respectively. The final values were classified into five equal intervals. The model is not a perfect reflection of participants' ratings, but it is, at the least, highly conservative in its estimates as the model's residuals illustrate (Figure 9).

On one hand, the map reveals miles of highly serviceable roads for bicyclists' use. The reality, however, is that the majority of these ideal roads are constrained to residential neighborhoods; as soon as one examines the interconnections between neighborhoods, serviceable links become fewer and farther between. It is likely that traversals of the San Marcos region would force bicyclists into stressful situations, navigating narrow roadways with large volumes of high-speed traffic.

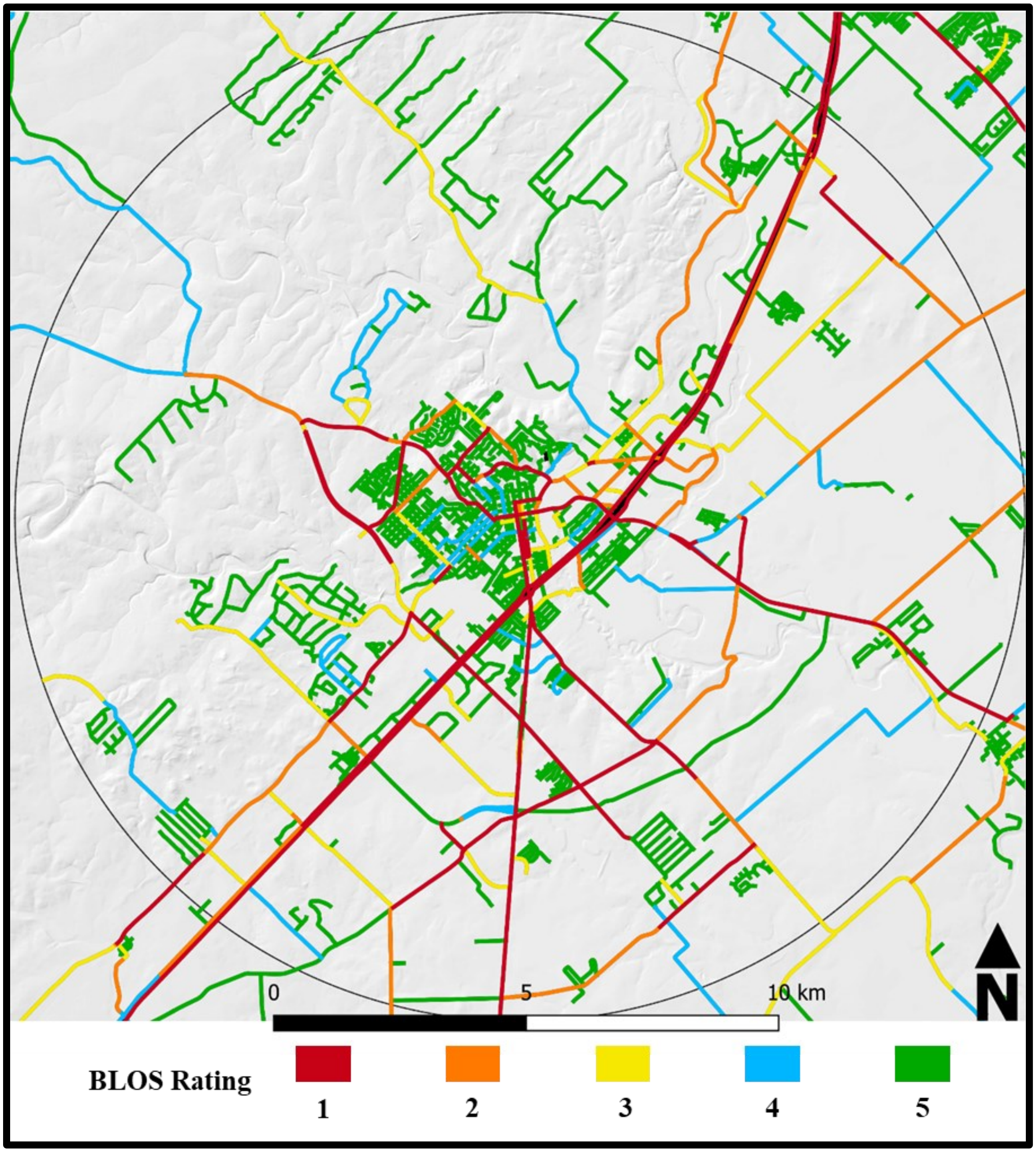

Figure 8: Map of the comprehensive regional BLOS model. (Author’s illustration.)

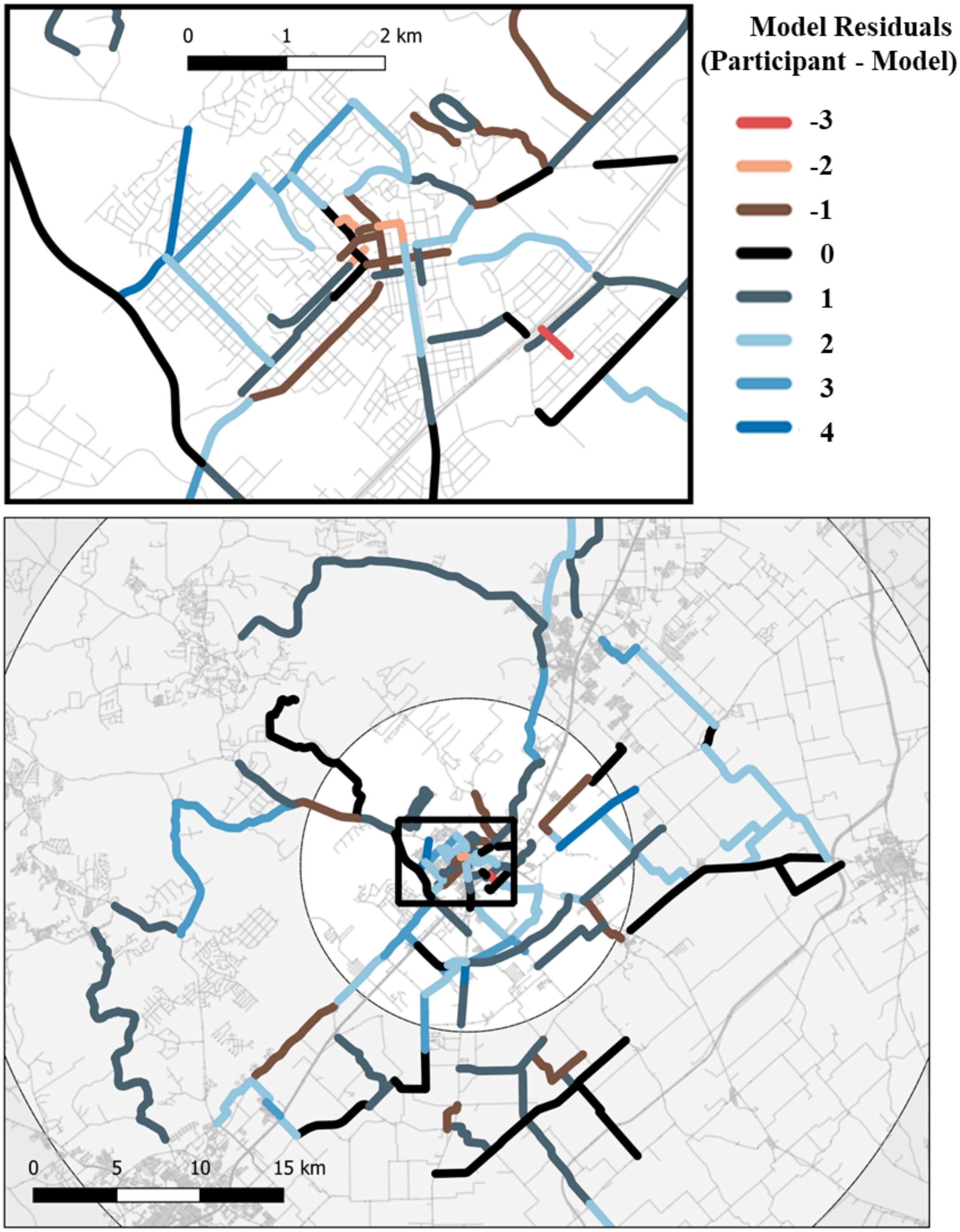

Figure 9: Map of comprehensive model residuals. Determined as participant's overall rating minus model estimates of BLOS. (Author's illustration.)

5.7 Contextual Analysis

A model's performance cannot be evaluated without adequate context, nor can a model be improved without fully understanding its shortcomings. To provide this context and comprehension, several roadway sites were selected for more intensive analysis. The selection represents an intentional cross-section of results, to include links popular with participants, and to highlight both success and failure of the BLOS model.

5.7.1 Link #16: West San Antonio Street

Link 16 was chosen for analysis because 7 unique participants highlighted it for their assessments. The consensus suggests the road is above average or even great for bicycling. The BLOS models frequently report this road as the best among the entire survey dataset, a result of low daily traffic counts, wide lanes (16 feet), and a low speed limit (30 mph, which was recently further reduced to 25 mph). It is important to highlight that on-street parking is permitted, but rarely occupied along the street; in terms of the data, TxDOT does not assign any shoulder width nor parking designation here. Furthermore, there are no bike lanes or shoulder striping – the entire roadway is understood as a shared, not separated, space.

5.7.2 Link #105: I-35 Frontage

Link 105 represented one of the lowest possible scores derived from the BLOS model (and matched by a participant rating of 1). Such a road is notable for several reasons. This type of roadway – interstate frontage – is often exempted from this kind of research under the wrongful assumption that bicyclists would never use such infrastructure. In practice however, this roadway represents a significant connection between downtown San Marcos and neighborhoods to the east of the freeway; although there are some suggested alternatives, they each suffer their own failings too (such as the unprotected left turn required to access Cape Road further south on Highway 123). Aside from recognizing the utility of these roadways for bicyclists, it also provides an extreme data point to test the BLOS model against; an outside lane of just 10 feet on a road carrying 21,000 cars per day at signed speeds of 45 mph (with speeds likely higher).

5.7.3 Link #21: Highway 80

After the I-35 frontage, the next worst model result among surveyed links was the stretch of Highway 80 east of the San Marcos River. This link connects city hall, the city library, the university's athletic complex, and the city's primary grocer, HEB. The low modeled rating is a product of significant traffic volume (30,000 cars per day) and successfully recalls participant's impressions (whom gave a rating between 1 and 2). Piecemeal multi-user sidewalks have appeared around various intersections along the link, providing at least some alternatives for pedestrians and cautious bicyclists – these efforts suggest the city is aware of the urgent need for improvements along this link.

5.7.4 Link #74: Aquarena Springs

The worst rated link – a 1 according to three participants – was Aquarena Springs between Sessom Drive and Charles Austin; this link is a key arterial for university traffic (pedestrians, bicyclists, and cars) as it connects major residential developments on the east side of town and also acts as the major highway feeder from Interstate 35 to the campus. The conflict is, in part, because of a 4-lane bridge over the San Marcos River, where narrow sidewalks offer limited space for pedestrians and dictate that bicyclists should occupy the traffic lanes. Despite being only 800 feet of roadway, participants commented that drivers treat the stretch "like a racetrack" and that it is a hindrance to bicycle commuting towards campus. The BLOS model rating was a lowly 2.

5.7.5 Link #26: Hunter Road

Another participant favorite (n = 5) was Hunter Road between "new" Ranch Road 12 and Dixon Street – the short link was only recently reconstructed with green-painted bicycle lanes including vertical delineators, or bollards. Before reconstruction, the link was still a two-way street with wide paved shoulders; nonetheless, in the original inventory data, the shoulder was marked as unpaved which reduced the derived effective outside lane width (expectedly diminishing the modeled BLOS estimate). Even after correcting the road attributes, the modeled result was significantly less than participant ratings, likely resulting from moderate speeds (40 mph) with high traffic counts (16,000 vehicles per day). Participant comments suggested that the bike lane, with separation from traffic, significantly improved their ratings; still, a common critique was that dirt and debris are allowed to accumulate in the bicycle lane and rarely cleared.

5.7.6 Link #17: Hutchinson Road

The section of Hutchinson Road between Comanche Street and C.M. Allen Parkway represents a key crosstown connection; despite crossing downtown traffic, the roadway retains relatively consistent width, speed, and traffic volume. This link is notable because the modeled BLOS largely exceeds participants ratings. This could be because, as with other links, the roadway inventory (particularly traffic count) could be out of date, or alternatively, the overestimation could suggest that cross-traffic interactions are poorly represented (despite the inclusion of the road density term). Another factor that may reduce participant ratings is that on-street parking is permissible along certain segments along the roadway, but this parking is not recognized by the roadway inventory data (despite the presence of attributes and values to do so).

5.7.7 Link #107: Centerpoint Road

Centerpoint Road, just south of Old Bastrop Road, reemphasizes the uncertainties remaining in the model. This link held among the highest residual differences; the BLOS model rating was a 2 (like Aquarena Springs), but the participant rating was above average, at a 4. The model results come from narrow travel lanes (9 feet) and a high speed limit (45 mph); though traffic volumes are moderate (1,855 cars per day), the road is relatively quiet many hours of the day, which may influence participants' rating.

5.7.8 Link #101: McCarty Road

McCarty Road, on San Marcos' southern margin, was modeled with a BLOS rating of 5, and reasonably reflects participants' averaged rating of 4.3 ($n = 3$). This result demonstrates that high traffic roads can also accommodate bicyclists: McCarty moves over 10,000 cars per day at a designated speed of 45 mph. The road has long been 4 travel lanes with paved shoulders; more recently, the shoulders have been designated as a bicycle lane (the road width is the same). There is still room for improvement, such as by installing physical separators between the bike lane and traffic lanes.

Table 8: Summary of selected links' attributes.

Link ID	Name	Avg. Rating	*n*	Model BLOS	Outside Lane Width	Traffic	# Lanes	Speed (mph)	Bike Lane	Pop. Density
16	W San Antonio	4.29	7	5	16	374	2	30	0	4.91
105	I-35 Frontage	1.00	1	1	10	20926	2	45	0	4.00
21	Highway 80	1.67	3	1	12	30249	4	40	0	5.68
74	Aquarena Springs	1.00	3	2	15	11874	4	30	0	3.34
26	Hunter Rd	4.20	5	3	15	16468	2	40	1	0.75
17	Hutchinson St	3.00	3	5	12	265	2	30	0	6.27
107	Centerpoint Rd	4.00	1	2	9	1855	2	45	0	0.33
101	McCarty Rd	4.33	3	5	18	10559	4	45	1	0.33

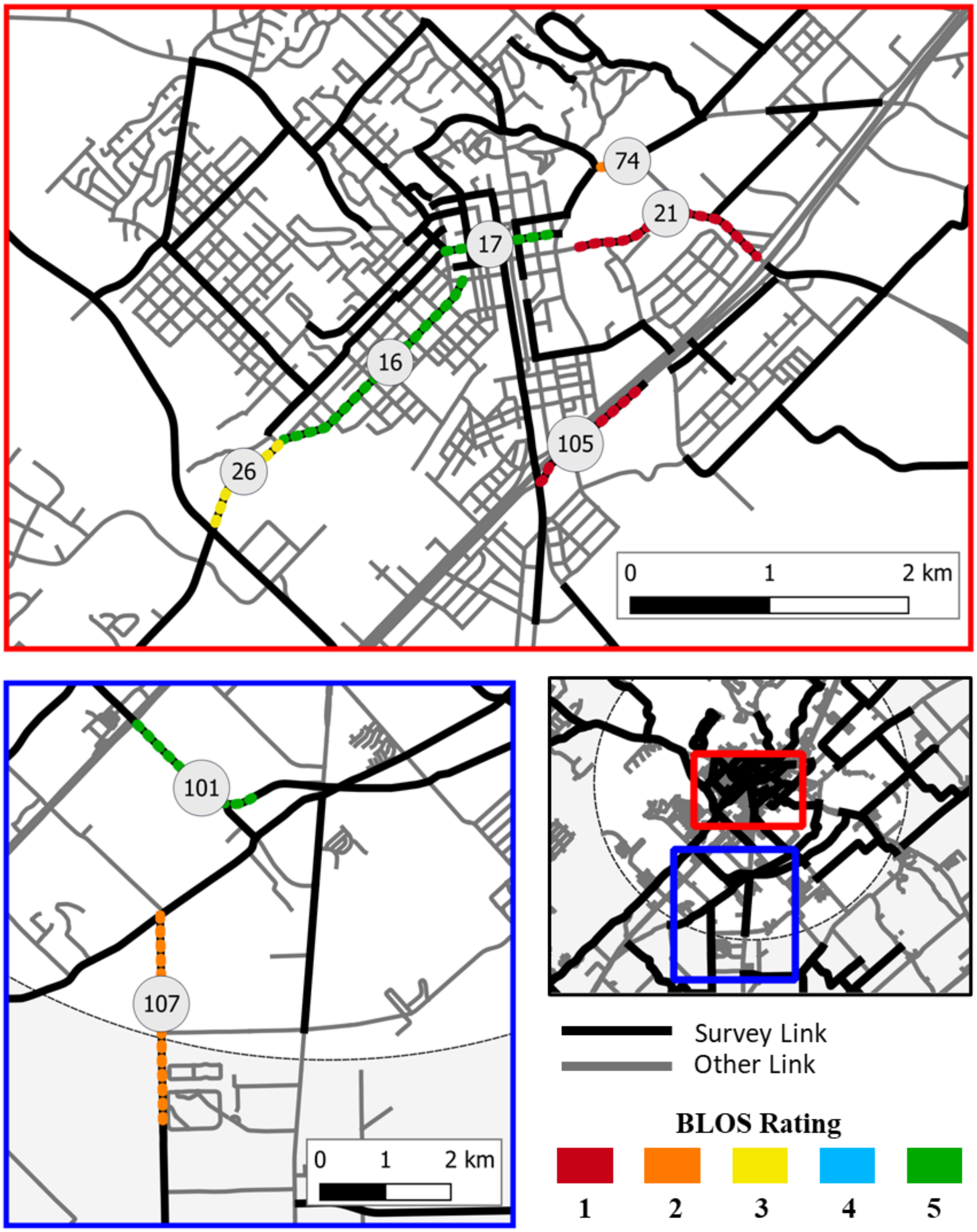

Figure 10: Map of links from contextual analysis. (Author's illustration.)

5.8 Summary of Results

The results of this research confirm and expand on the state-of-the-art in BLOS modeling. Analysis of the participant survey data validates bicyclists' sensitivity to factors of road width, traffic volume, and posted speed limits. The combination of these and other roadway attributes are shown to be useful for approximating overall user evaluations. The models do unfortunately leave significant unaccounted for variations in bicyclists' evaluations. Although these variations demand continued scrutiny, the modeling results are shown to be reasonably useful, both for the analysis of individual road links, and equally useful for the generation of a regional network model. While this research suggests some potential improvements to the existing BLOS models, these results are far from definitive. Rather, the results shared here offer a foundation for the continued exploration of BLOS modeling – first, by rethinking strategies for the validation of model results, and second, by considering the potential utilization of ever more powerful geographic information systems.

6 DISCUSSION

Though it is possible to utilize a BLOS model to assess a regional road network, there is still room for improvements to the modeling strategy. However, as the participants in this research often shared, there is also a need to improve the bicycling experience more generally, that is, if bicycling is to be recognized as a viable mode of alternative (and sustainable) transportation. In many respects, our current roadway infrastructure is more than adequate, but merely predominated by automobiles which pose a grave threat to bicyclists, pedestrians, and other drivers alike. Discussion of the study methods, objectives, and results should further clarify this point.

6.1 Methods

6.1.1 Participant Survey

Road site ratings provided by local bicyclists gave undeniable weight and credibility to the modeling results, despite the unaccounted variations. The chosen research design limited the sample pool, which may have increased the difficulty of the recruitment procedure. Nonetheless, while the participant pool was small, all participants came to the table eager and willing to share their experiences – they cared very deeply about the subject matter and felt strongly about advocating for better conditions for alternative transportation within their community.

The snowball sampling design did expand the reach of the survey, though most participants were personal connections from local bicycling groups – the San Marcos Cycling Friends, MoveSM, Sustainable San Marcos, and the Texas State Cycling Club. These personal connections were cultivated over thousands of miles ridden together; without these shared experiences, this research design may not have been as effective.

6.1.2 Modeling

Modeling is a popular and practical application of modern GIS systems. In many cases, however, a model is employed unquestioningly. The modeling results demonstrate the importance of testing (and retesting) our models from time to time. Although these results suggest that more recent models do outperform older and simpler models, it must be remembered that this success is observed only in the context of San Marcos when comparing the results with the perspectives of mostly experienced (not novice) bicyclists. How these results might transfer to larger or smaller cities, to other sample populations, or to other cultures should be carefully considered when choosing a BLOS model.

6.1.2.1 *Data*

Data remains the most important aspect of a model's design. In this research, for instance, *percentage of peak parking demand* – specified in the Highway Capacity Manual – was ignored because that data is not recorded by any transportation agency with jurisdiction over the study area. Likewise, speed limits, road widths, and traffic counts were not always available or up to date. Future efforts to recreate a regional model of BLOS must keep the availability of data in mind, and modelers must be prepared to adapt modeling equations to the data they have.

6.1.2.2 *Static versus Dynamic Models*

Recent literature has focused largely on dynamic regression models, where parameters are permitted to vary widely in pursuit of the best-fit model. As the results show, a regression model does provide the highest r-squared value, and the least pairwise differences (of course, zero); however, this should not be considered the most practical model! The regression model limits the ability to incorporate variable interactions as well as prevents adjustment and analyses of individual parameters. Static parameterization, although complex and potentially frustrating, enhances our ability to communicate about the modeling terms, their meanings, potential interactions, and ultimately, our rationalization for the inclusion or exclusion of specific terms.

6.2 Variables

The methods aimed to address four research questions; the first stated: which variables (describing the road network) are most essential for representing bicyclists' assessments of roads, and how should those variables be mathematically arranged?

6.2.1 The Big Three

Perhaps as expected, outside lane width, traffic volume, and traffic speeds were among the most influential variables across all models. Beyond these initial three factors, the results also suggest the inclusion of density attributes (which are presumed proxies for modeling cross traffic) and furthermore, the results do not suggest the exclusion of hillslope as a factor for bicycling serviceability. Although these factors may seem obvious, they warrant scrutiny for forward-looking transportation planning efforts.

Speed was nearly excluded from the stepwise regression BLOS model. In fact, it was the least important factor according to both participants' stated and revealed preferences. Considering an application of the regression coefficient, the difference between a 35 mph road and a 65 mph road is less than 0.5 points (using the 5 point classification system). While automobiles have made great strides in protecting occupants from increasingly higher travel speeds, these protections rarely extend to vulnerable road users; in many cases, safer vehicles are heavier and taller than their forebears (White 2004). It might be argued that, for bicyclists, the fear of being struck at higher speeds is not so considerably different from being struck at more modest travel speeds (Richter et al. 2006; Tefft 2013). In the San Marcos region, speed limits are rarely set below 30 mph, with little hope of those limits being reduced – furthermore, traffic often ignores posted limits, and traffic enforcement is lackadaisical.

Width and traffic volume compounds an unfortunate realization – rather than any fanciful measure of the bicycling experience or infrastructure serviceability, the BLOS models merely restate a presupposition of bicycling safety. Decreases in road width and increases in traffic volume both contribute significantly to the bicyclist's risk of collision with fast moving automobiles. Optimistically, it bears reiterating that the interaction between width and traffic is never fully realized in any model: where wide shoulders are provided, high traffic volumes are almost entirely inconsequential and where traffic is relatively light, narrower lanes without special provision are insignificant.

The mention of lane width naturally leads to the discussion of bike lanes, which were more influential than initially hypothesized. Generally speaking, the provision of a bike lane implies a guaranteed minimum width for bicycling; in the San Marcos region, these bike lanes often include vertical separators, which may enhance the perception of separation, and thus safety. The results of this research, however, cannot distinguish between the various bike lane designs, a point which demands future investigations (Veillette et al. 2019). Furthermore, while a bike lane may be *generally* preferable to none, there are still concerns that bike lanes are not adequately maintained and are often littered with gravel and road debris; occasionally, these obstructions are severe enough to force bicyclists into the vehicular travel lanes.

6.2.2 Hillslope and Density

The importance of hillslope and density for bicycle service was well established prior to this research, but their inclusion in previous models was often treated rather subjectively. The major improvements – beyond proving the validity of including these two factors – was in the method of their inclusion: a data-driven and empirically grounded approach that is readily replicated by professional GIS technicians.

Occasionally bicyclists may seek out steep hills for training purposes, but more generally, such hills will always represent an impediment to easy transportation (if not in terms of effort expended, then in terms of travel time at the very least). However, it remains difficult to specify appropriate thresholds for hillslopes; for instance, North LBJ (an average grade of 7% with a maximum ramp of 11% grade) was provided as an example of a hill that was too steep by some participants, and yet remarked as the perfect hill by others. New Ranch Road 12 – a 4 lane highway – was also suggested as a perfect hill by multiple participants; that stretch of road maintains a much more reasonable grade: averaging around 3% with a maximum ramp of 8%. If the modeling results are to be considered revealed preferences, a penalty for very steep hills (over 13%) provided the best overall BLOS estimates. In any case, the inclusion of hillslope had a marginal effect on the model results, reiterating once again that bicyclists appear to relay their experience and ratings in terms of safety rather than alternative measures of service.

Following this line of thinking would then rationalize the significance of density in the model: less densified areas decrease traffic conflicts, and therefore provide an increased sense of bicycling safety. Various density measures are likely to be highly correlated in any context (as they were across San Marcos), thus the use of density in a

BLOS model must be carefully considered. Ultimately, the performance differences between the different density measures were marginal, and the inclusion of any one density measure significantly improved model results. On theoretical grounds, the use of road network density should offer the best proxy for the unmeasurable concern regarding cross-traffic; furthermore, this measure could be derived for any areal unit (including an artificial grid) overlaid upon the road network.

The possibility of deriving hillslope and density measures is a relatively recent capability enabled by the widespread adoption of geographical information systems. There are numerous methods by which these measures may be defined and equally numerous data sources to explore. Between the two, however, density is far more accessible to most transportation planning agencies and far more vital for accurately modeling BLOS. The derivation of hillslope for linear segments needs far more attention, and far more simplification to encourage its widespread inclusion in modeling efforts – this point is applicable to more than just bicyclists, and might be useful for hiking trails, as well as search and rescue efforts in that domain.

6.3 Models

After an initial assessment of each modeling variable individually, the second research question asked: which BLOS modeling strategy produces the most representative map of bicyclists' reported experiences in the local study area?

The earlier models (BLOS equations 1 through 4) were useful starting points, and relied on seemingly useful arrangements of key variables. Unfortunately, the results from these models were not nearly as robust as contemporary revisions. Both equations 5 and 6 performed similarly in terms of model fit, but did vary somewhat according to pairwise differences. Equation 6, adapted from the Highway Capacity Manual, provided reasonable correlation with participant ratings while minimizing pairwise differences.

Equation 5 benefits from the inclusion of local land use and curb cut frequencies, variables which are assumed to indicate increased risk to bicyclists in the form of increased generation of cross-traffic. The study results show that density measures can be used in place of both the subjective classification of land use and for the laborious task of counting curb cuts. The inclusion of density improved the model's correlation as much as 5%, and as such, inclusion of these factors should not be ignored.

While Equation 6 performed well in initial tests, it would be logical to argue that the inclusion of density (or similar metrics to account for cross-traffic) could further improve the model. Adding these measures to the equation significantly reduced (the already marginal) pairwise differences between the model results and participant ratings, and increased the overall strength of the correlations. Although the revised model is not considered definitive, it does represent a step forward in modeling BLOS.

6.4 The Bicycling Experience

Reducing participant's interviews to a mere model could be construed as overly reductionist; to better recognize their voice and expertise, the third research question posed: how do local bicyclists' experience and perceive their regional road networks level of service, beyond those attributes reflected in the BLOS model? Acknowledging participants' commentary throughout the survey helped to identify key issues, including the definition of a link, the recognition of informal routing barriers, the importance of a heuristic measure of roadway quality, and the lesser role of each link when contrasted with the value of a safe network of routes for bicycling.

6.4.1 Links

The definition of a link presented the first of challenges. The purest definition of a link is a single unit between nodes; governments' road data itself rarely reflects this topological ideal out of necessity (for instance, posted speed limits often vary mid-link). For the participants, a link neither fit the ideal definition nor the data from TxDOT. The definition of a link in this research ultimately fell to an on-the-spot agreement between the interviewer and each participant: that there were no major changes in lane count, traffic volume, roadway width, or traffic speeds along rated segments. The different conceptions of a link introduced a data management challenge that must be, and was, subjectively resolved – this is doubly true when considering that currently, dedicated bicycling infrastructure is often stored in database schema wholly detached from its counterpart roadway data.

The definition of a link during participant interviews also reiterated the importance of the holistic bicycling network. When first asked to provide their own link for discussion, many participants offered short routes, best represented as a collection of links. In some cases, the process of subdividing short routes into agreeable links led participants into making clear relative comparisons of each link – if not in the overall rating, then at least as a nudge in either traffic, width, or speed. The argument, then, is that bicyclists' initial BLOS heuristics are more readily available as an assembly of potential route choices – if this is convincing, then it reaffirms that poorly serviced (and potentially dangerous) links work to undermine small improvements made elsewhere in the network.

6.4.2 Barriers

This line of thinking extends itself further – that major barriers in the bicycle network, while often plainly obvious, are often severely underserved. Rivers, railroads, and freeways each impose linear boundaries on the fabric of any city; vehicular traffic is often granted dedicated bridges and flyways and given first consideration at any highway junction. In San Marcos, car traffic is afforded two different railyard overpasses, neither of which accommodates bicycles or pedestrians. The latest trend in interstate engineering – the diverging diamond – keeps cars moving at 30-45 mph as they enter and exit the city; bicyclists are expected to either take the lane, or to act as a pedestrian (which is possibly worst of all: crossing 5-8 lanes of high speed, yield-only interstate traffic at a trot). There is no surprise that bicyclists have been involved in collisions at every possible link under Interstate 35 (and along each adjacent arterial, Figure 11).

6.4.3 Quality

Beyond the heuristics of route-choice and the fabric of the city, bicyclists occasionally granted reference to pavement quality. Rough chipseal or fractured concrete led to expressions of dissatisfaction with certain links. In some cases, rough pavement was attributed to fresh chipseal, which according to highway standards, would represent

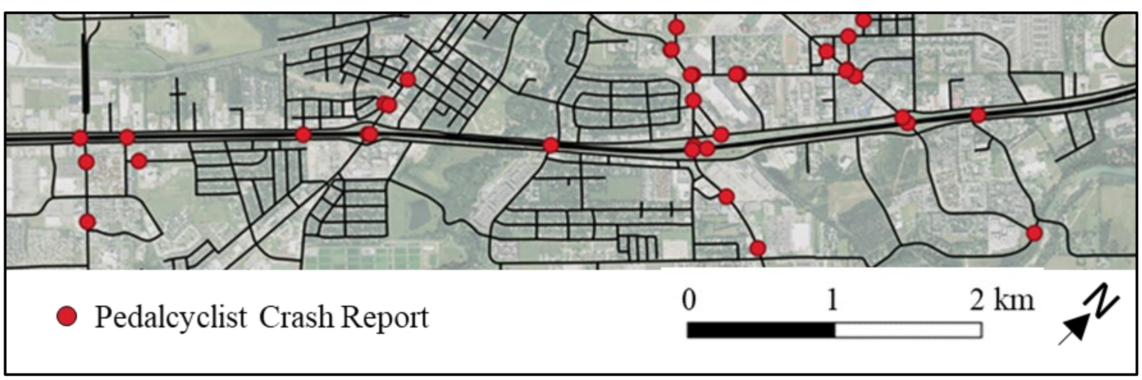

Figure 11: Bicyclist involved crash reports along the I-35 corridor, 2010-2020. (Author's illustration.)

high-quality pavement conditions (and confound BLOS models). In any case, pavement condition data was not available in San Marcos, and the term was either removed or held constant in models. The majority of roads deemed rough by participants were rated as either a BLOS of 2 or 3; in contrast, some participants assigned high ratings to unpaved gravel roads (their focus, of course, was the minimal traffic on these routes, rather than roughness of the surface). Addressing these challenges in a future BLOS model will require first, better data describing surface qualities, and second, a better understanding of the tradeoffs in preferences regarding surface quality, road width, and traffic.

The concerns about quality naturally prompt a discussion of bike lanes. While the presence of bicycle lanes typically increased participants' rating of a link, their commentary suggests there is more to learn about the different designs and maintenance needs of bike lanes. Participants appreciated the dedicated width and especially the vertical delineators which visibly separate them from vehicular traffic; however, many commented that sand, glass, and debris would tend to collect in bicycle lanes – in some cases, forcing them to ride in vehicle lanes. In San Marcos' city center, some bicycle lanes are constructed on the lefthand side of one-way arterials; this layout frustrated some participants: before and after these leftward lanes, traffic expects a bicyclist to ride "as far right as practicable", creating unnecessary conflicts. These same lanes also create conflicts with downtown parking – the bike lanes tend to occupy the "door zone", an issue that has been well documented. Finally, although participants may perceive bicycle lanes to provide higher levels of service, this may be a false sense of security. Research shows that unbuffered bike lanes reduce the space given by passing vehicles: white paint does not adequately protect a bicyclist from negligent (or malicious) driving.

A discussion about the quality of bicyclist infrastructure would not be complete without mentioning automotive drivers. Many participants' comments about traffic extended beyond simple traffic counts, towards an unquantifiable sense of drivers' behavior. Some comments were positive, such as 'drivers are more aware of us' along popular bicycling routes, but most comments were more distressed – 'they drive too fast for conditions', 'drivers are distracted', 'it's a racetrack', 'turning traffic rarely yields'. From this perspective, the solution to increasing BLOS is rather simple: slow the cars down and get drivers to pay more attention. Participants mentioned two solutions, which are both strongly accepted in academic literature. First, reduce the speed limit. Second, enforce the laws on the books – that's it.

6.4.4 Network

Multiple participants remarked upon the lack of connectivity between existing highly-rated bicycling infrastructures, or across the region as a whole. In most cases, improvements for bicyclists are done as a matter of convenience and expediency under the umbrella of larger roadway (automotive) improvement projects; this has the predictable consequence of leaving the bicycle route network in pieces with little concern about missed connections, dead-end routes, or problems with the last-mile. Within the city, there is the potential that connectivity in the bicycle network can be greatly improved with only small improvements to road widths, lane paints, and route signage. Improving network level of service across the region, however, will require more generous projects and lane allotments – this is not a call for inter-city bicycle lanes, but rather small reductions in rural highway travel speeds, fully paved shoulders, and appropriate consideration for bicyclists where guardrails and bridges are required.

While intra- and inter-city bicycling improvements are feasible within the current transportation design paradigm, the last-mile may thwart planners' best efforts. At the moment, the convention of suburban-styled big-box grocery and retail at the heart of the city implies that bicyclists and pedestrians will inevitably cross paths with motor-vehicles in the parking lot. The lack of conflicts – reportable crashes – at these locations is likely a reflection of current societal conditions: there are no conflicts because few bicyclists dare to go. Among the transportation cyclists interviewed, each had strategies for overcoming these conflicts – getting off and walking, taking to the sidewalks, and ultimately, waiting for cars. The right of the driver to cruise by the front entrance of the grocer is more important than either pedestrians' or bicyclists' accessibility. Even as cities rapidly update their roadway infrastructure, the planning and zoning of private developments will likely lag for decades to come, preferencing customers in cars over any possible alternative. The fullest understanding of urban accessibility – and its improvement – should carefully consider the challenges that lay beyond the road.

6.5 Planning and Infrastructure

The fourth and final research question sought to frame the model results within the local planning context by asking: does current regional transportation planning address infrastructural inadequacies highlighted by the proposed BLOS model?

The comprehensive regional model highlights both roadways with high levels of service and those roads with subpar service for bicyclists. In particular, roads with low levels of service (scores of 1 or 2) are often major arterials – in many cases, there are no alternatives for bicyclists to consider. The model results suggest that these roadways could be improved easily by either increasing lane width or reducing travel speeds. More complex interventions could seek to reduce daily vehicle traffic, or to add dedicated bicycle facilities with adequate vehicular separation.

6.5.1 Anticipating Bicyclists

In San Marcos, there are a handful of completed and planned projects that suggest the city as a whole is anticipating bicyclists in future roadway designs. The bicycle lane on Hunter Road (Link #26) was added in 2020 and was received favorably by local survey participants. A similar project along Guadalupe Road promises enhanced BLOS in the city's core: the road underwent a road diet (from 3 travel lanes to 2) and added a wide bike lane. The specifics of that project expose the current challenge of qualifying bicycle infrastructure, as the two-way bike lane and adjacent on-street parking have both drawn some criticism from local bicyclists; because construction was delayed during 2021, final opinions on the overall project design were generally reserved.

Two other projects remain in the planning phase, but promise significant improvements, not just for link-wise level-of-service, but for holistic connectivity of the bicycle network in the city (specifically, connecting campus to popular student residences). Sessom Drive borders the north side of the Texas State University campus, and is currently a 4-lane arterial with 35 mph speed limits (limits which are frequently exceeded); the road is also a significant hill, meaning major speed differentials between car and bicycle traffic in the uphill direction. Planning documents show the street is to be restriped, narrowing from its current 4 lanes down to just 2 traffic lanes, with buffered bicycle lanes on either side – the buffers will include vertical delineators near intersections and conflict points. This project will almost certainly enhance BLOS along the northern perimeter of the campus.

The second significant restriping project promises improvements along Old Ranch Road 12, improving BLOS along a major connection between several large apartment complexes and the campus (almost reaching the Sessom Drive improvement project). Old Ranch Road 12 currently varies from 35 to 50 mph with very high traffic volumes in just 2 travel lanes. Although there are paved shoulders in some places, the pavement quality and width vary dramatically; worse, the shoulders disappear right as the hillslope begins. The restriping project maintains the two travel lanes, but expands the shoulders and designates that space specifically for use by bicyclists.

6.5.2 Challenges and Missed Opportunities

The major challenge ahead is that these sorts of improvements are only passable when roadway reconstruction is already planned (i.e., because pavement quality is impeding the flow of vehicular traffic). This means that city-wide improvements to the bicycle network remain a long-term concern; it also means that the success of current projects will be used to evaluate the viability of future projects – of course, improvements along individual road links should not be expected to provide dramatic shifts in modal choice or bicyclists' route selections.

Rather than focus on improvements to individual road links, planners must recognize the importance of improving network connectivity. The stretch of Aquarena Springs between Sessom Drive and Charles Austin Drive is an exemplar: the link is less than two-tenths of a mile, but was rated a BLOS of 1 by three participants. To the west of this link is the university campus, with its extensive pedestrian and bicycle facilities; to the east is Post Road with dedicated bicycle lanes. Plenty of students brave this stretch of road for their bicycle commute – but no doubt, others have put off the idea of bicycle commuting solely to avoid the real and perceived danger of those two-tenths of a mile.

Aside from obvious oversights as on Aquarena Springs, bicycles are often overlooked in the process of making improvements. As a minor example, consider the three-way intersection of Dixon Road, Hunter Road, and West San Antonio Street (all three routes were rated favorably by participants): the space between these three bicycle routes is a constricted two-lane space with moderate traffic volumes and no facilities indicating bicyclists may need to make a left turn (across traffic). A more egregious example concerns an unmapped paved connector between Riverside Drive and River Road (Fairchild 2021). This connector offered a low-traffic route *under* Interstate 35, but has been blockaded – potentially permanently – for highway improvements; bicyclists and pedestrians are now forced to navigate the divergent-diamond interchanges either at State Highway 123 or 80 (an extra mile even by the shortest route, the frontage road).

6.5.3 Planning in San Marcos

Overall, San Marcos is taking steps to include bicycles within its long-term transportation plans (and living up to the promise of providing for sustainable alternative transportation). These projects, however, are limited in extent and reliant on justifications (such as increasing bicycle use) that may not fully materialize without better consideration of the complete bicycling network. Furthermore, infrastructural improvements for bicycles are often contingent on improvements to the vehicular network – making improvements for bicycles contingent on the maintenance of vehicular spaces all but assures overlooked opportunities and missed connections. Only time will tell if San Marcos maintains its current momentum and meaningfully improves BLOS.

7 CONCLUSIONS

7.1 Narrative

De Nobis Ipsis Silemus [we are silent concerning ourselves] – Kant

Quantitative modeling – such as BLOS – provides an unshakeable sense of authority: numbers don't lie. Yet, as much recent literature suggests, it would be foolish to ignore the positionality and subjectivities of the primary investigator (Holmes 2021; Mruck and Breuer 2003). This research was the product of significant personal interest (as well as the usual tenacity and discipline required of a dissertation).

I am a bicyclist. In the four years I have lived in the San Marcos region, I have pedaled over 30,000 miles – 10,000 miles in just the last year. My experience, I argue, was essential to the development of this BLOS model, rather than biasing it. The research methods – especially the reliance on participants for sampling and rating roads – were intentionally structured to distance my own subjectivity from the results. More importantly, the scope of the research – a regional BLOS model – would not have been feasible without my familiarity of the roads within the study area. Finally, the survey recruitment strategy and long-form interviews were likely dependent on the strong personal ties I have developed with local bicyclists in the study area.

Although I forwardly advocate for bicyclists – for better infrastructure, for better protection under the law, and for better treatment by drivers – I fully understand the limitations of the bicycle as a primary means of transportation. A common critique is that bicycling advocacy is 'ableist' (Surico 2020); often this argument arises from vocal able-bodied individuals and not those who are truly disabled. Good roads and safe streets are important for all manner of assisted-mobility devices (Fucoloro 2011; Galatan 2019).

The bigger story here is not about increasing bicycle level of service, but rather that society is owed a reconciliation between drivers and *literally everyone else*. We should no longer accept 35,000 deaths per year as the status quo – it is unacceptable. Some might point to automated and self-driving cars as one potential solution, but even if one such vehicle was to make it to market by the end of this decade, how long before the entire manually-operated fleet could feasibly be replaced?

And if self-driving cars become more prevalent in the next decades, we must still grapple with the natural resource costs of such vehicles. Electrification is happening at a record pace, sure; not surprisingly, we are also setting a record pace on the extraction of lithium and other rare earth minerals. Will there be enough of these precious metals to satisfy our demand for automobility? Even if there is, allowing the design of our cities to revolve centrally around such vehicles only assures that the poor, young, old, and disabled are left with limited accessibility.

The bigger story is that, in order to promote bicycling – and walking, and public transit – we will need to take a harsh stance on the use of the personal motor car. Encouraging alternative, sustainable modes of transportation necessarily implies reducing speed limits, eliminating excess travel lanes, dedicating road space to bicycles (as well as assistive e-bikes and scooters), and enforcing traffic regulations. The bigger story is that we need to address the *normal automobile driver as a traffic problem* (Forbes 1939).

My story, like many others, is that I see dangerous driving daily. I, like other bicyclists, am afraid that one day I will be struck by a vehicle and seriously injured or killed. I am strong, but not *fearless* (Heine 2013). I am afraid that, like cyclists before me, I will be in a bag and my assailant on the couch – that is motivation for this research.

7.2 Limitations

In addition to the limitations foreshadowed by existing literature, the research process revealed some additional considerations for BLOS modeling. The initial limitations remarked that the study must be considered within the context of mid-sized US cities (and more specifically, within the context of southern sunbelt cities which largely grew after the popularization of automobility). Furthermore, the participant sampling design intentionally highlights the perspectives of experienced cyclists, and may differ from those of infrequent or non-cyclists (whom are often the target audience for mode-shifting advocacy). Finally, San Marcos does not enjoy any extensive dedicated bicycling infrastructure, which means that the results may underappreciate the demand and preference for facilities such as bike lanes.

Beyond the above, anticipated limitations, data quality was the most distressing factor in BLOS model experimentation. The roadway inventory database was woefully out-of-date and entire attribute columns were empty – even basic attributes, such as the speed limit, were often missing. One reason for the low quality of the road network data may have been jurisdictional: the records regarding state-managed roadways were much more complete and accurate than roadways within city and county jurisdictions. In any case, low quality and erroneous data significantly changes the reported BLOS (and many other transportation models); improving road network data must be an ongoing priority among transportation planners at every level, from administrators to users.

Even where perfect data may exist, the presumed relationship between database road links, the ideal theory of a road link, and the bicyclists' conception of a link may not always be compatible. In theory, an ideal link is a road between two nodes; in a roadway inventory, a link may cross nodes, or not connect to any nodes at all (usually these links are divided along practical or functional geometries, such as to capture changes in road names, pavement width, or traffic volumes). To the bicyclist, neither of these 'links' are relevant; bicyclists respond more readily to *routes*, understanding the series of links needed to get from one point to another in the network. In this way, the selection of a route involves the selection of a collection of links, most likely with a heuristic system for selecting the best collection of links (and balancing the best and worst of the links in the set). This implies a need to further evaluate BLOS in terms of bicyclist route choice.

7.3 Future Research

The BLOS model is a useful tool for assessing and visualizing large road network datasets (such as in the San Marcos region), but the results deserve continued attention. Currently, the BLOS model only accounts for a third of the variations in participants' ratings, implying many unaccounted factors – these unobserved factors may be the most inhibiting to increasing bicycle ridership. Participants commented about sand and glass debris, signage (or lack thereof), the sense of other bicyclists frequenting the route, and the aesthetics along certain roads. While the variation in road debris accumulation and cleanup may be beyond modeling efforts, careful consideration of route signage and bicycle ridership are reasonable experiments in the near term; similarly, investigating route popularity and assessing roadway aesthetics might suggest clues for increasing bicycle use, and the requisite data are likely to be increasingly available with the popularization of social media and fitness-tracking applications.

Among the most difficult of participants' comments to codify was their common reference to a range of driver behaviors. In some places, this reference was positive: "drivers are aware of us" or "drivers are cautious"; other links however, were colored with anecdotes of drivers traveling too fast for conditions, bicyclists "being buzzed" (or passed too closely) despite empty passing lanes, or bicyclists having their right-of-way violated by turning traffic. Future modeling may seek to incorporate a larger emphasis on daily traffic patterns or include observed vehicular travel speed to better represent actual driver behaviors within the road network. In lieu of these more data intensive operations, continued investigation of urban density could offer a useful proxy for BLOS modeling and may also lead to larger questions about ideal urban densities.

Another important avenue for future research concerns hillslope. Measuring hillslope is perhaps foremost, a generalized issue for geographic information systems: how can the process be simplified and how can its results be standardized? If these improvements are within reach, then how do we consider demarcating mild, moderate, and steep slopes (for bicyclists, pedestrians, and others)? Such a question is not as simple as it may seem, for there are clear differences between a maximum gradient struck over a short length, a sustained moderate gradient over a longer length, and mild gradients that are sustained over many miles (such as found in many mountainous areas).

One final suggestion for future research regards the cartography of bicycle maps. Unlike road maps, there is little standardization of such maps, even for fundamentals such as preferred roads and routes. An ideal bike map would include the various roads and routes within the area, and a sort of ranking system to compare and contrast route options (such as BLOS), but it should also share details about particular routes' hills and traffic. Beyond such fundamentals, participants largely suggested the inclusion of water fountains, bicycle shops, and similar points of interest. While these mentions may seem obvious, making such a detailed map both pleasant and easy to read is no easy task.

7.4 Final Remarks

Bicycle level of service is an underappreciated metric – encouraging its adoption and use could drastically improve planning for bicycles. At present, the model is largely contingent upon measures of vehicular traffic (speed and volume) and road width (as a measure of separation from the aforementioned traffic). Nonetheless, the model is shown to be easily adapted (thanks to the use of well-defined and fixed parameter coefficients); this implies that as traffic and infrastructure changes in the future, the BLOS model can be updated to match. In the future, BLOS could be incorporated in route choice models, include measures of roadway aesthetics, and ultimately, be made available to end users.

Beyond the model, we might recognize the growing utility of geographic information systems. Two decades ago, a BLOS model was constrained to a handful of links penciled onto a clipboard; today, the same model can be cast across an entire network dataset (with some limitations). In the future, this technology may become more user friendly and more readily deployed; this anticipated commodification will be significant for organizations with tight budgets and limited technical expertise. It is vital that we continue to simplify and refine our models so that they may be replicated (and adapted) to suit the planning needs of various regions and organizations.

Finally – and most generally – we should recognize the importance of the bicycle as a mode of transportation and recreation. The bicycle provides for personal and societal wellness in a way that few other industrial-era machines can match. 'Twould be foolish to ignore the bicycle while championing the electric car or the diesel bus; 'twould be foolish to decry the bicycle as ableist in an era of obesity. As attributed to H.G. Wells, and still ringing true for millions today, "every time I see an adult on a bicycle I no longer despair for the future of the human race".

APPENDIX SECTION

APPENDIX A: Tools of the Trade

- **Salsa Vaya Touring Bicycle**
  - 700c wheels with 32c-45c tires
  - Downtube friction shifters
  - 3x9 drivetrain (22/36/46 x 11-36 :: 17" – 112" gear-inches)
  - Disc brakes with 160mm rotors

Figure 12: Salsa Vaya

- **Garmin Edge 130 Consumer Bicycling GPS**
  - Supports GPS, GLONASS, and Galileo constellations

Figure 13: Edge 130

- **Custom Built Personal Computer**
  - AMD Ryzen 5 2600 Six-Core Processor (3.40 Ghz)
  - Gigabyte B450 AORUS M Motherboard
  - 32.0 GB DDR4 3200MHz RAM
  - 256 GB SDD system drive, 1 TB HDD storage drive
  - NVIDIA GeForce GT710 2GB GDRR3 Display Adapter
  - Windows 10
    - MS Office 2016 (Word, Excel, Powerpoint)
    - QGIS 3.17, with GRASS
    - R processing language with R-Studio GUI

APPENDIX B: Survey Instrument

Part 1: Demographics (maximum 10 minutes)
Name:
Survey Consent:
Age:
Gender:
San Marcos Resident:
Hays County Resident:
Self-identified Cyclist:
Primary Discipline(s):
Estimated Miles per Year:

Part 2: Road-Sites for BLOS Analysis (maximum 40 minutes)
Preselected Site #1 (expected *High BLOS*)
Location: *York Creek Road, approximately at Soechting Ln*
Location Description: *halfway along York Creek, between Old Bastrop & Francis Harris*
Overall BLOS (perceived): 1 2 3 4 5
Road Width (perceived): 1 2 3 4 5
Traffic (perceived): 1 2 3 4 5
Speed Limit (perceived): 1 2 3 4 5
Hill Slope (perceived): 1 2 3 4 5

Preselected Site #2 (expected *Moderate BLOS*)
Location: *Hunter Road, northeast from McCarty Ln*
Location Description: *aka FM2439, approximately near Sienna Pointe Apartments*
Overall BLOS (perceived): 1 2 3 4 5
Road Width (perceived): 1 2 3 4 5
Traffic (perceived): 1 2 3 4 5
Speed Limit (perceived): 1 2 3 4 5
Hill Slope (perceived): 1 2 3 4 5

Preselected Site #3 (expected *Low BLOS*)
Location: *East Hopkins, between North Guadalupe Street and North LBJ*
Location Description: *East Hopkins, north of historic county courthouse and city square.*
Overall BLOS (perceived): 1 2 3 4 5
Road Width (perceived): 1 2 3 4 5
Traffic (perceived): 1 2 3 4 5
Speed Limit (perceived): 1 2 3 4 5
Hill Slope (perceived): 1 2 3 4 5

Participant Selected Site # _____
Location: ____
Location Description: ____
Overall BLOS (perceived): 1 2 3 4 5
Road Width (perceived): 1 2 3 4 5
Traffic (perceived): 1 2 3 4 5
Speed Limit (perceived): 1 2 3 4 5
Hill Slope (perceived): 1 2 3 4 5
Any Comments: ____

Participant Selected Site # _____
Location: ____
Location Description: ____
Overall BLOS (perceived): 1 2 3 4 5
Road Width (perceived): 1 2 3 4 5
Traffic (perceived): 1 2 3 4 5
Speed Limit (perceived): 1 2 3 4 5
Hill Slope (perceived): 1 2 3 4 5
Any Comments: ____

Participant Selected Site # _____
Location: ____
Location Description: ____
Overall BLOS (perceived): 1 2 3 4 5
Road Width (perceived): 1 2 3 4 5
Traffic (perceived): 1 2 3 4 5
Speed Limit (perceived): 1 2 3 4 5
Hill Slope (perceived): 1 2 3 4 5
Any Comments: ____

Participant Selected Site # _____
Location: ____
Location Description: ____
Overall BLOS (perceived): 1 2 3 4 5
Road Width (perceived): 1 2 3 4 5
Traffic (perceived): 1 2 3 4 5
Speed Limit (perceived): 1 2 3 4 5
Hill Slope (perceived): 1 2 3 4 5
Any Comments: ____

Part 3: Open Discussion (maximum 10 minutes)

a. What do you believe is the most important information to include for mapping bicycle routes?

b. As a cyclist, can you describe a local hill that is 'too steep' (or at least, too steep for novice bicyclists, or one that you might avoid)?

c. Is there a 'perfect hill' locally, a hill that gains elevation but at a lower effort?

d. Is it possible to rank the importance of 1) road width, 2) traffic, and 3) speed limits, in regards to bicycling comfort?

APPENDIX C: Data Review

To provide additional context for the model inputs, the following review is intended to summarize key attributes of secondary and primary datasets.

Secondary Data – All Links in Region

Table 9: Secondary data summary, all links

Variable	
OBJECTID	n = 42,775 road links

Variable	**Min**	**Median**	**Mean**	**Max**	**St. Dev.**	**Count NAs**
SPD_MAX	0	0	21.7	85	28.4	
SPD_MAX_FLR	21	21	34.6	85	18.8	
LANE_WIDTH	4	10	10.5	36	1.9	21
LANE_WIDTH_O	8	10	12.3	53	5.3	
NUM_LANES	1	2	2.3	14	0.9	1
S_WID_O	0	0	2.0	40	4.4	251
ADT_CUR	1	405	11956.7	271276	30430.1	1
TRK_AADT_P	0	3	5.0	56	5.2	1
D1A_mean	0	1	1.4	31	1.8	
D1B_mean	0	2	3.3	72	4.0	
D1C_mean	0	0	2.5	203	13.2	
D3a_mean	1	9	10.4	44	7.9	
D3b_mean	0	23	39.6	813	49.6	

Variable	
S_USE_O	*Classes* -- 1:Parking, 2:Parking, 3:Bicycle, 5:Emergency …
S_TYPE_O	*Classes* -- 1:Paved, 2:Concrete, 3:Unpaved, 5:Earth …
SRF_TYPE	*Classes* -- 1:Concrete, … 7:Composite, … 13:Gravel

Sampled Links – Secondary and Primary Data

Table 10: Secondary and primary data summary, sampled links

Variable					
Link_ID	n = 131 road links				

Variable	Min	Median	Mean	Max	St. Dev.
SPD_MAX	0	40	39	65	15
SPD_MAX_FLR	21	40	40	65	12
LANE_WIDTH	8	10	10	16	2
LANE_WIDTH_O	8	10	12	24	4
NUM_LANES	2	2	2	4	1
S_WID_O	0	0	2	15	3
ADT_CUR	17	2706	7112	40149	8977
TRK_AADT_P	0	3	4	22	3
D1A_mean	0.02	0.24	0.94	3.96	1.12
D1B_mean	0.04	0.66	3.05	16.50	4.28
D1C_mean	0.00	0.32	2.99	49.91	8.33
D3a_mean	1.14	4.02	8.43	23.58	7.47
D3b_mean	0.66	8.60	33.87	163.99	41.15
Rate_Overall	1	3	3.21	5	1.18
Rate_Width	1	3	2.81	5	1.24
Rate_Traffic	1	3	3.13	5	1.26
Rate_Speed	1	3	3.15	5	1.04
Rate_Hill	1	4	3.63	5	1.14
ft_mile	0	54	62	238	45
max.grad.50m	0.00	0.07	0.08	0.34	0.05
BLOS_5	1.0	4.3	4.0	5.0	1.1
BLOS_6	1.0	2.7	2.8	5.0	1.1
BLOS_REGR	1.5	3.2	3.2	4.9	0.7
BLOS_5_Rev	1.0	4.2	3.8	5.0	1.1
BLOS_6_Rev	1.0	3.0	3.0	5.0	1.2

Table 11: Individualized regression results (Participants' Overall Level of Service Ratings as follows their own ratings of Speed, Traffic, and Width)

	Individual Regression Coef.				Model Fit		*Revealed Priorities*			Stated Priorities		
ID	*b0*	Speed	Traffic	Width	r_sq	p_val	*S*	*T*	*W*	S	T	W
1	-3.26	0.55	0.40	0.98	0.95	0.00	2	3	1	2	3	1
2	-0.24	0.08	0.94	0.20	0.67	0.81	3	1	2	2	1	3
3	-0.39	0.26	0.89	0.01	0.84	0.24	2	1	3	2	1	3
4	0.05	0.86	-0.13	0.29	0.68	0.04	1	3	2	2	3	1
5	1.45	0.20	-0.02	0.38	0.21	0.37	2	3	1	3	2	1
6	-2.05	0.27	0.93	0.47	0.61	0.42	3	1	2	2	1	3
7	0.24	0.18	0.41	0.40	0.91	0.06	3	1	2	3	2	1
8	-0.72	0.13	0.98	0.60	0.89	0.67	3	1	2	2	3	1
9	0.00	0.06	0.67	0.28	0.87	0.82	3	1	2	3	2	1
10	0.93	-0.05	0.20	0.65	0.80	0.82	3	2	1	3	2	1
11	1.34	0.16	0.42	0.17	0.75	0.36	3	1	2	1	2	3
12	0.33	-0.19	0.65	0.38	0.78	0.50	3	1	2	3	1	2
13	0.07	1.50	-0.86	0.06	0.88	0.06	1	3	2	1	3	2
14	-1.63	0.21	0.73	0.47	0.75	0.36	3	1	2	2	3	1
15	1.09	1.59	-1.43	0.59	0.88	0.03	1	3	2	2	3	1
16	-0.41	0.19	0.66	0.42	0.92	0.29	3	1	2	3	2	1
					Average Prioritization of Factors:		2.44	1.69	1.88	2.25	2.13	1.63

APPENDIX D: Data Management

Road Digitization

A duplicate copy of the TxDOT road-line features were created to enable advanced digitization of the features while preserving the original topology for potential future analysis (Figure 14). The newly edited feature shapefile was reprojected to UTM Zone 14, and was also trimmed of extraneous attribute columns for faster processing: the attribute table includes only two fields, *OBJECTID*, a reference to the link in the TxDOT feature set, and *Link_ID,* the reference used for tracking road-links as selected by interview participants.

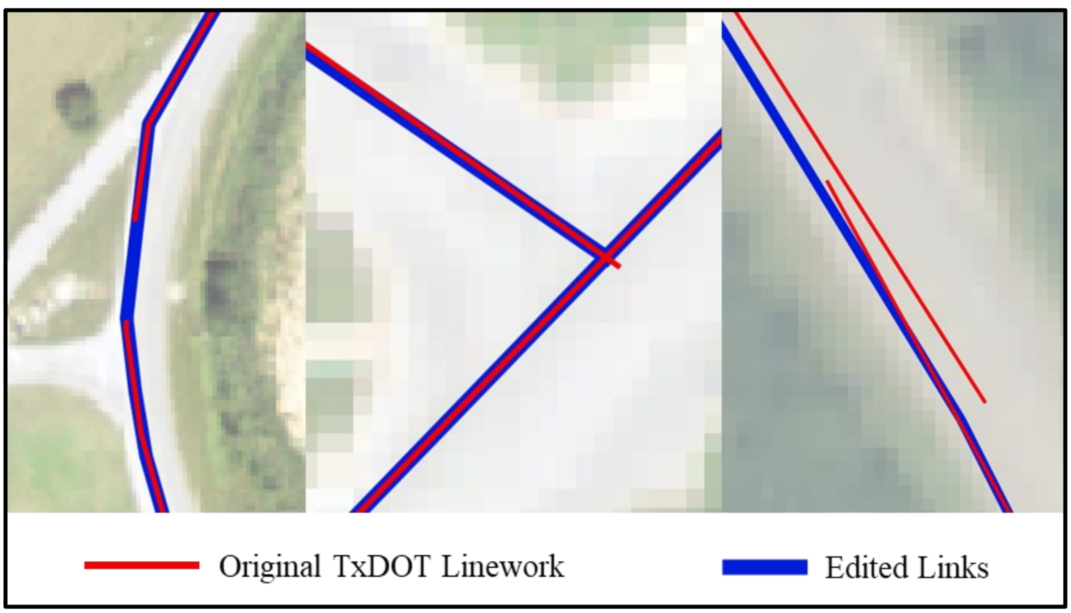

Figure 14: Advanced digitization of road network topology

Road Attributes

The R programming language was used for many processing steps, as it is often faster than processing within a GIS environment. Several road attributes were updated or derived from initial attributes. One key attribute was road speed, stored as *SPD_MAX.*

For many road links, the speed was either 0 or null. Links described by participants were manually updated to reflect real world conditions. For all remaining links with missing values, the following code was used to update the field to a floor of 21 mph, as suggested in the Highway Capacity Manual.

```
## Speed Limit, HCM FLOOR

   yt <- roads_data

   ## Update Null values to zero
   yt$SPD_MAX[is.na(yt$SPD_MAX)] <- 0

   ## If SPD_MAX < 21, update to 21, else leave as is
   yt$SPD_MAX_FLR <- ifelse( yt$SPD_MAX < 21, 21, yt$SPD_MAX)
```

For many BLOS models, the lane-width value is intended to include the outside shoulder width as well, assuming that the outside shoulder is paved and not designated as parking. The following codes were used to create a new attribute field to represent the effective outside lane width, *LANE_WIDTH_O*.

```
## Outside Lane Width -- Check that Outside Shoulder is paved,
##     and not designated parking.

   yt <- roads_data

   ## Seed new field with existing Lane Width attributes
    yt$LANE_WIDTH_O <- yt$LANE_WIDTH

   ## Iterate across every road link
   for (i in seq_along(yt$LANE_WIDTH_O)) {

     ## If shoulder width is not null, proceed:
     if (!is.na(yt[i, "S_WID_O"])) {

     ## Assume Shoulder Width extends Lane Width
      yt[i,"LANE_WIDTH_O"] <- yt[i,"LANE_WIDTH"]+yt[i,"S_WID_O"]

      ## Undo change if Shoulder is unpaved, or parking:
           if (yt[i, "S_USE_O"] == 1 || yt[i, "S_USE_O"] == 2) {
            yt[i,"LANE_WIDTH_O"] <- yt[i,"LANE_WIDTH"]}
           if (yt[i, "S_TYPE_O"] > 2 || yt[i, "S_TYPE_O"] < 1) {
            yt[i,"LANE_WIDTH_O"] <- yt[i,"LANE_WIDTH"]}
      }}
```

Density Attributes from the EPA Smart Growth Database

The QGIS Graphical Processing Modeler was utilized to manage the many-to-one join between census block groups and intersecting road-links (Figure 15). The output of the model is a line-feature shapefile mirroring the input line features; the new feature shapefile includes summary means from selected block-group attributes: land-use densities D1a, D1b, and D1c, as well as road-network densities D3a and D3b.

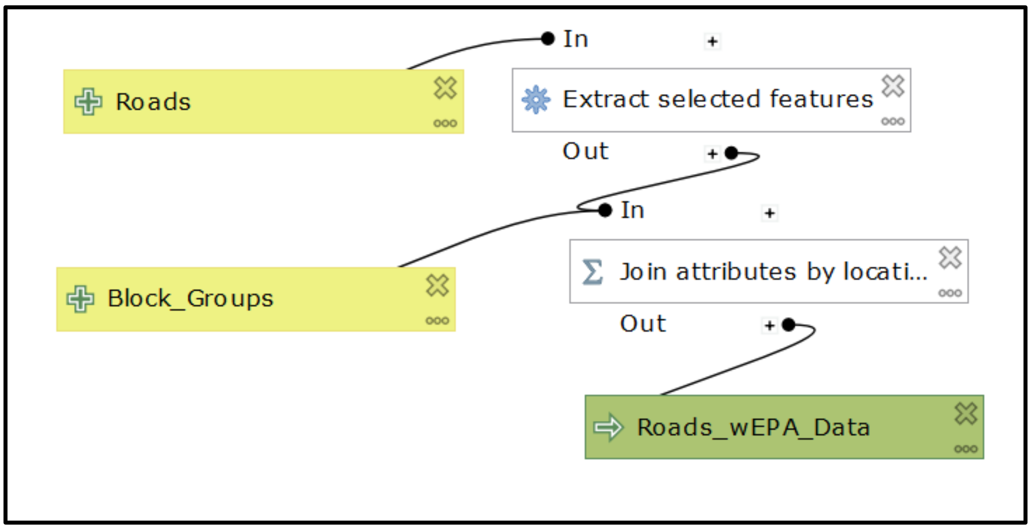

Figure 15: Process model for summarizing block groups attributes for road links.

Hillslope Derivation

Estimation of hillslopes along links was complicated in two ways. First, issues with the road network topology prevented implementation of the *points along geometry* on all links, limiting the application of hillslope derivation to manually-adjusted links. Second, even where meaningful chainages of the network were determined, hillslope estimates were occasionally exaggerated, either because of challenging terrain (roads along cliffsides), link-DEM mismatches (where road lines did not align with bridges), or where the National Elevation Dataset did not capture new road infrastructure (bridges, water crossings, and road-grade cuts). Despite these challenges, a GIS process model was useful for deriving hillslope estimates from topologically corrected links, from which the output could be manually validated (Figure 16). The output is a shapefile of point features, each with a unique *Point_ID*, a reference to its source *Link_ID*, x and y coordinates (in UTM), and a z coordinate sampled from the input DEM (the USGS NED was used for this research).

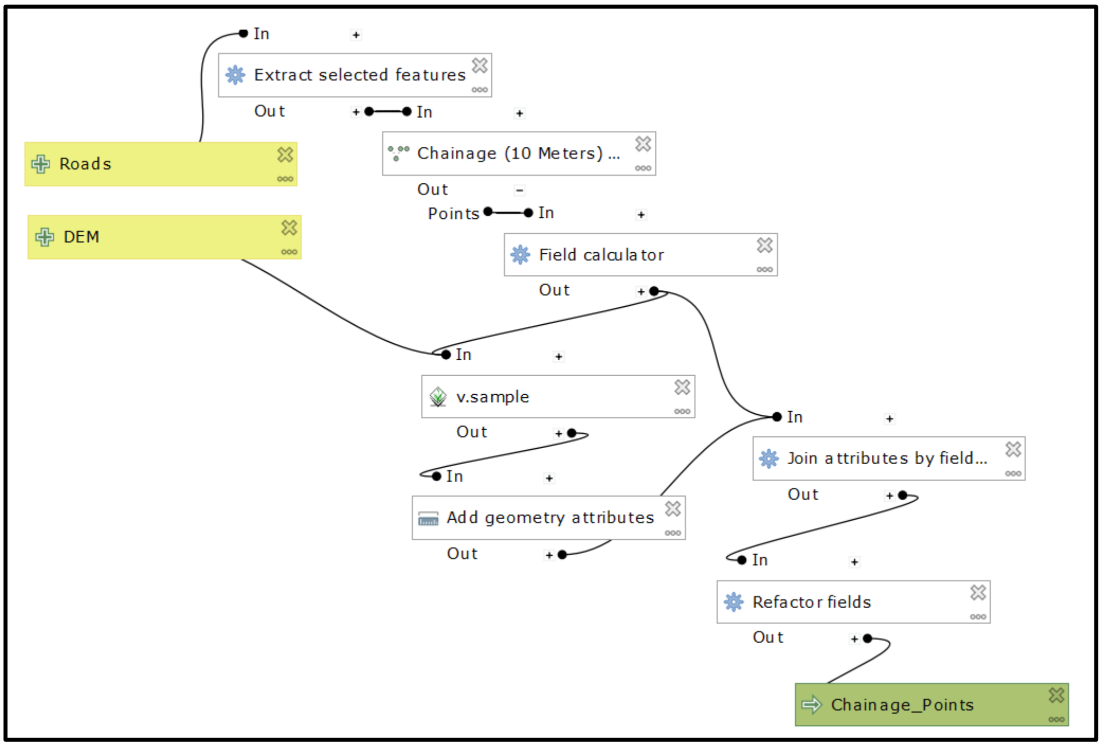

Figure 16: Process model for deriving points (and elevations) along road links.

The output data within *Chainage_Points* were further processed using an R script file. The script tracks the distance along links as reported by the GIS and as derived from x and y coordinates. For each increment, the process records the change in elevation, and then concludes by summarizing elevation gain and gradient using multiple referents. The process operates on each link both forward and backward to consider differences in the directionality of hillslopes along links.

```
###### Hillslope Preparation ######
library(tidyverse)
library(dplyr)
options(scipen=999)

####### BASIC DISTANCE/TRIGONOMETRY ATTRIBUTES #######
###### distance, elevation change, and gradient. #####

## FORWARD ALONG LINKS
df <- Chainage_Points
df <- data.frame(pid=df$Point_ID, Link_ID=df$Link_ID,
distance=df$distance,
                   x=df$xcoord, y=df$ycoord, z=df$rast_val)
seed_dist = 0
point_count = seed_dist

df$dist <- seed_dist; df$dist.cum <- seed_dist
df$dist_3d <- seed_dist; df$dist_3d.cum <- seed_dist
df$vert_change <- seed_dist; df$gradient <- seed_dist
df$gradient_30m <- seed_dist; df$gradient_50m <- seed_dist

for (i in seq_along(df$distance)) {
  if (df[i, "distance"] == 0) {
  ## New Link; reset metrics
     df[i,7:13]  <-seed_dist
     point_count <-seed_dist
  }

  if (point_count > 0) {
  ## Safe to compare current row to previous row
    df[i, "dist"] <- sqrt((df[i, "x"]-df[i-1, "x"])^2 +
                          (df[i, "y"]-df[i-1, "y"])^2)
    df[i, "dist.cum"] <- df[i, "dist"] + df[i-1, "dist.cum"]

    df[i, "dist_3d"] <- sqrt((df[i, "x"]-df[i-1, "x"])^2 +
                             (df[i, "y"]-df[i-1, "y"])^2 +
                             (df[i, "z"]-df[i-1, "z"])^2)

    df[i, "dist_3d.cum"] <- df[i, "dist_3d"] + df[i-1, "dist_3d.cum"]
    df[i, "vert_change"] <- df[i,"z"]-df[i-1,"z"]
    df[i, "gradient"]    <- df[i,"vert_change"] / df[i,"dist"]
  }

  if (point_count > 3) {  ## Safe to reference third previous row
    df[i, "gradient_30m"] <- (df[i,"z"]-df[i-3,"z"]) /
(df[i,"dist.cum"]-df[i-3,"dist.cum"])
  }

  if (point_count > 5) {  ## Safe to reference fifth previous row
    df[i, "gradient_50m"] <- (df[i,"z"]-df[i-5,"z"]) /
(df[i,"dist.cum"]-df[i-5,"dist.cum"])
  }

  point_count = point_count+1}
forward <- df

## PARSE LINKS IN REVERSE
df <- na.omit(Chainage_Points)
df <- data.frame(pid=df$Point_ID, Link_ID=df$Link_ID,
distance=df$distance,
                   x=df$xcoord, y=df$ycoord, z=df$rast_val)

seed_dist = 0
point_count <- seed_dist
```

```
df$dist <- seed_dist; df$dist.cum <- seed_dist
df$dist_3d <- seed_dist; df$dist_3d.cum <- seed_dist
df$vert_change <- seed_dist; df$gradient <- seed_dist
df$gradient_30m <- seed_dist; df$gradient_50m <- seed_dist

df$pid <- rev(df$pid)
df <- df[order(df$pid),]

for (i in seq_along(df$pid)) {
if (i > 1){
  if (df[i-1, "distance"] == 0) {
  ## New Link; reset metrics
      df[i, 7:13]  <- seed_dist
      point_count  <- seed_dist
    }
  else { ## Safe to compare current row to previous row


   df[i,"dist"]<-sqrt((df[i,"x"]-df[i-1,"x"])^2 + (df[i,"y"] -
                                                   df[i-1,"y"])^2)
   df[i, "dist.cum"] <- df[i, "dist"] + df[i-1, "dist.cum"]

   df[i, "dist_3d"] <- sqrt((df[i, "x"]-df[i-1, "x"])^2+
                            (df[i, "y"]-df[i-1, "y"])^2+
                            (df[i, "z"]-df[i-1, "z"])^2)

   df[i, "dist_3d.cum"] <- df[i, "dist_3d"] + df[i-1, "dist_3d.cum"]
   df[i, "vert_change"] <- df[i,"z"]-df[i-1,"z"]
   df[i, "gradient"]    <- df[i,"vert_change"]/df[i,"dist"]

   if (point_count > 3) {## Safe to reference third previous row
       df[i, "gradient_30m"] <- df[i, "z"] - df[i-3, "z"] /
                                (df[i,"dist.cum"]-df[i-3,"dist.cum"]
   }
   if (point_count > 5) {## Safe to reference fifth previous row
     df[i, "gradient_50m"] <- (df[i,"z"]-df[i-5,"z"]) /
                              (df[i,"dist.cum"]-df[i-5,"dist.cum"])
   }
        point_count = point_count+1}
}}
backward <- df

## SUMMARY RESULTS (Forward/Backward/Generalized)

for_result <- data.frame(Link_ID = unique(forward$Link_ID))
for (i in seq_along(for_result$Link_ID)){
  ## Subset unique link and attributes from large "df"
  temp <- forward[forward$Link_ID == for_result[i, "Link_ID"],]
  for_result[i, "dist_GIS"]          <- round(max(temp$distance),0)
  for_result[i, "dist_h"]            <- round(max(temp$dist.cum), 0)
  for_result[i, "dist_3d"]           <- round(max(temp$dist_3d.cum),0)
  for_result[i, "elev.gain"]         <-

round(sum(temp$vert_change[temp$vert_change>0]),0)
  for_result[i, "elev.change.abs"]   <- round(sum(temp$vert_change),0)
  for_result[i, "grad.avg.up"]       <-
                    round(mean(temp$gradient[temp$gradient>0]),3)
  for_result[i, "grad.trig.up"]      <- round(for_result[i,
                    "elev.gain"]/for_result[i, "dist_h"],3)
  for_result[i, "grad.max"]          <-
                                round(max(temp$gradient_30m),3)
  for_result[i, "grad.max.30m.abs"]  <-
```

```
                                round(max(abs(temp$gradient_30m)),3)
  for_result[i, "grad.max.50m.abs"]   <-
                    round(max(abs(temp$gradient_50m)),3)
  for_result[i, "ft_mile"]            <- round(for_result[i,
         "elev.gain"*3.28084]/(for_result[i,"dist_3d"]*0.000621371),0)
}

bac_result <- data.frame(Link_ID = unique(backward$Link_ID))
for (i in seq_along(bac_result$Link_ID)){
  ## Subset unique link and attributes from large "df"
  temp <- backward[backward$Link_ID == bac_result[i, "Link_ID"],]
  bac_result[i, "dist_GIS"]           <- round(max(temp$distance),0)
  bac_result[i, "dist_h"]             <- round(max(temp$dist.cum), 0)
  bac_result[i, "dist_3d"]            <- round(max(temp$dist_3d.cum),0)
  bac_result[i, "elev.gain"]          <-
                   round(sum(temp$vert_change[temp$vert_change>0]),0)
  bac_result[i, "elev.change.abs"]    <- round(sum(temp$vert_change),0)
  bac_result[i, "grad.avg.up"]        <-
    round(mean(temp$gradient[temp$gradient>0]),3)
  bac_result[i, "grad.trig.up"]       <- round(bac_result[i,
                   "elev.gain"] / bac_result[i, "dist_h"],3)
  bac_result[i, "grad.max"]           <-
                   round(max(temp$gradient_30m),3)
  bac_result[i, "grad.max.30m.abs"]   <-
                   round(max(abs(temp$gradient_30m)),3)
  bac_result[i, "grad.max.50m.abs"]   <-
                   round(max(abs(temp$gradient_50m)),3)
  bac_result[i, "ft_mile"]            <- round(bac_result[i,
         "elev.gain"*3.28084]/(bac_result[i,"dist_3d"]*0.000621371),0)
}

## Generalized Hillslope Result
gh <- left_join(for_result, bac_result,
        by="Link_ID", suffix=c(".for", ".bac"))

gh <- data.frame(
  Link_ID          = gh$Link_ID,            dist_GIS = gh$dist_GIS.for,
  dist_h           = gh$dist_h.for,          dist_3d = gh$dist_3d.for,
  elev.gain.for    = gh$elev.gain.for, elev.gain.bac =
gh$elev.gain.bac,
  elev.gain.max    = gh$elev.gain.for, elev.gain.abs =
    abs(gh$elev.change.abs.for),
  elev.gain.sum    = gh$elev.gain.for + gh$elev.gain.bac,
  grad.avg.up.for  = gh$grad.avg.up.for,
  grad.avg.up.bac  = gh$grad.avg.up.bac,
  grad.max.30m.abs = gh$grad.max.30m.abs.for,
  grad.max.50m.abs = gh$grad.max.50m.abs.for,
  ft_mile.for      = gh$ft_mile.for,     ft_mile.bac = gh$ft_mile.bac,
  ft_mile          = (gh$ft_mile.bac + gh$ft_mile.for) / 2)

  for (i in seq_along(gh$Link_ID)){
  gh[i,"elev.gain.max"] <-
max(c(gh[i,"elev.gain.for"],gh[i,"elev.gain.bac"]))  }
```

Complete Model Coding

```
options(scipen=999)
library(dplyr)
library(ltm)

##########################
#### RAW DATA SOURCES ####
##########################

## DEMOGRAPHICS (Table of User Demographic Data, no Roadways)
de <- Survey_Res_Demographic

## SURVEY DATA RESULTS (User Sample and Ratings of Roadways)
dt <- Survey_Res_Location

## ROAD EMPIRICS (Topologically Corrected, Roadway Attributes)
yt <- Roads_Empirics

## HILLSLOPE (Previously Derived from Chainage/Hillslope Scripts)
gh <- gh

#########################################
#### OUTLIERS AND SAMPLE MANAGEMENT #####
#########################################

## REMOVE OUTLIERS
outliers <- c(105, 129, 991)
dt <- dt[!(dt$Link_ID %in% outliers),]
yt <- yt[!(yt$Link_ID %in% outliers),]

## RANDOMIZE SAMPLE
##dt <- dt[!dt$Participant_ID == round(runif(1,min(dt$Participant_ID),
   max(dt$Participant_ID))),]

############################
#### DERIVED ATTRIBUTES ####
############################

## Outside Lane Width -- Check that Shoulder is paved and not parking
    yt$LANE_WIDTH_O <- yt$LANE_WIDTH
      for (i in seq_along(yt$LANE_WIDTH_O)) {
        if (!is.na(yt[i, "S_WID_O"])) {
          yt[i,"LANE_WIDTH_O"] <- yt[i,"LANE_WIDTH"]+yt[i,"S_WID_O"]

          if (yt[i, "S_USE_O"] == 1 || yt[i, "S_USE_O"] == 2) {
           yt[i,"LANE_WIDTH_O"] <- yt[i,"LANE_WIDTH"]}
          if (yt[i, "S_TYPE_O"] > 2 || yt[i, "S_TYPE_O"] < 1) {
           yt[i,"LANE_WIDTH_O"] <- yt[i,"LANE_WIDTH"]}
      }}

## Effective Shoulder Width:
    yt$SHOULDER_WIDTH <- yt$LANE_WIDTH_O - yt$LANE_WIDTH

## Speed Limit, HCM FLOOR
    yt$SPD_MAX[is.na(yt$SPD_MAX)] <- 0
    yt$SPD_MAX_FLR <- ifelse( yt$SPD_MAX < 21,   21, yt$SPD_MAX)

## ADT HOURLY, HCM FLOOR
    yt$ADT_HR <- ifelse((yt$ADT_CUR/24) <
      (4*yt$NUM_LANES),4*yt$NUM_LANES, yt$ADT_CUR/24)
```

```
## PREPARTIONS FOR BLOS 1 (RANKING REAL ATTRIBUTES)
   #RANK ADT (--- per lane, per hour!---)
  b <- c(0,150,250,350,450,99999);   a <- c("5","4","3","2","1")
  yt$BLOS_1_rankADT <- cut((yt$ADT_HR/yt$NUM_LANES),breaks=b,labels=a)
  ## BODGE TO GET NUMERIC RANKS:
       yt$BLOS_1_rankADT <- (as.numeric(yt$BLOS_1_rankADT)-6)*(-1)

  #RANK WIDTH
    yt$BLOS_1_rankWidth <-  yt$S_WID_O + yt$LANE_WIDTH_O
     b <- c(0, 11.5, 12.5, 13.5, 14.5,1000)
     a <- c("1", "2", "3", "4", "5")
    yt$BLOS_1_rankWidth <- cut(yt$BLOS_1_rankWidth,breaks=b,labels=a)
    yt$BLOS_1_rankWidth <- as.numeric(yt$BLOS_1_rankWidth)

   #RANK SPEED
     b <- c(-1,26,31,36,46,100)
     a <- c("5","4","3","2","1")
    yt$BLOS_1_rankSpeed <- cut(yt$SPD_MAX, breaks=b, labels=a)
     ## BODGE TO GET NUMERIC RANKS:
    yt$BLOS_1_rankSpeed <- (as.numeric(yt$BLOS_1_rankSpeed)-6)*(-1)

## END DERIVING ATTRIBUTES

##############################
#### DEMOGRAPHICS RESULTS ####
##############################

## COMPARE WITH KAPLAN(1975):2,332 Mi/Year; higher in Texas (2751-3250)
  mean(de$Miles_Year)

#####################################
#### PARTICIPANT ROADWAY RATINGS ####
#####################################

##### SUMMARY AND SKEW #####
    Quick_Summary <- function(VectorToSummary){
      s <- double()
      s[1] <- median(VectorToSummary)
      s[2] <- mean(VectorToSummary)
      s[3] <- sd(VectorToSummary)
      return(s)}

  Results_1a <- data.frame(Rating=character(), Median=integer(),
   Mean=double(), Sigma=double())
  Results_1a[1, "Rating"] <- "Overall"
    Results_1a[1, 2:4] <- Quick_Summary(dt$Rate_Overall)
  Results_1a[2, "Rating"] <- "Width"
    Results_1a[2, 2:4] <- Quick_Summary(dt$Rate_Width)
  Results_1a[3, "Rating"] <- "Traffic"
    Results_1a[3, 2:4] <- Quick_Summary(dt$Rate_Traffic)
  Results_1a[4, "Rating"] <- "Speed"
    Results_1a[4, 2:4] <- Quick_Summary(dt$Rate_Speed)
  Results_1a[5, "Rating"] <- "Hills"
    Results_1a[5, 2:4] <- Quick_Summary(dt$Rate_Hill)
  Results_1a$Skew <- (Results_1a$Mean - Results_1a$Median) /
   Results_1a$Sigma
    Results_1a[,2:5] <- round(Results_1a[,2:5],3)

##### VALIDITY: CORRELATION BETWEEN PERCEPTIONS AND EMPIRICS #####
    xt <- ''
    xt <- inner_join(dt, yt, by="Link_ID")
    xt <- inner_join(xt, gh, by="Link_ID")
```

```
Results_1b <-data.frame(Parameter=character(),
  Rate_Overall=double(),Rate_Width=double(),Rate_Speed=double(),
  Rate_Traffic=double(),Rate_Hills=double())

Results_1b[1, "Parameter"]    = "LANE & SHOULDER WIDTH"
Results_1b[1, "Rate_Overall"] = cor(xt$Rate_Overall, xt$LANE_WIDTH_O)
Results_1b[1, "Rate_Width"]   = cor(xt$Rate_Width,   xt$LANE_WIDTH_O)
Results_1b[2, "Parameter"]    = "LANE WIDTH"
Results_1b[2, "Rate_Overall"] = cor(xt$Rate_Overall, xt$LANE_WIDTH)
Results_1b[2, "Rate_Width"]   = cor(xt$Rate_Width,   xt$LANE_WIDTH)
Results_1b[3, "Parameter"]    = "ROAD WIDTH"
Results_1b[3, "Rate_Overall"] =
  cor(xt$Rate_Overall,xt$LANE_WIDTH*xt$NUM_LANES)
Results_1b[3, "Rate_Width"]   =
  cor(xt$Rate_Width,xt$LANE_WIDTH*xt$NUM_LANES)
Results_1b[4, "Parameter"]    = "SHOULDER WIDTH"
Results_1b[4, "Rate_Overall"] = cor(xt$Rate_Overall, xt$S_WID_O)
Results_1b[4, "Rate_Width"]   = cor(xt$Rate_Width,   xt$S_WID_O)
Results_1b[5, "Parameter"]    = "MAX SPEED"
Results_1b[5, "Rate_Overall"] = cor(xt$Rate_Overall, xt$SPD_MAX)
Results_1b[5, "Rate_Speed"]   = cor(xt$Rate_Speed,   xt$SPD_MAX)
Results_1b[6, "Parameter"]    = "MAX SPEED (Floor 21mph)"
Results_1b[6, "Rate_Overall"] = cor(xt$Rate_Overall, xt$SPD_MAX_FLR)
Results_1b[6, "Rate_Speed"]   = cor(xt$Rate_Speed,   xt$SPD_MAX_FLR)
Results_1b[7, "Parameter"]    = "CURRENT AADT"
Results_1b[7, "Rate_Overall"] = cor(xt$Rate_Overall, xt$ADT_CUR)
Results_1b[7, "Rate_Traffic"] = cor(xt$Rate_Traffic, xt$ADT_CUR)
Results_1b[8, "Parameter"]    = "AADT * (1 + %Heavy)"
Results_1b[8, "Rate_Overall"] =
  cor(xt$Rate_Overall,xt$ADT_CUR*(1+(.01*xt$TRK_AADT_P)))
Results_1b[8, "Rate_Traffic"] =
  cor(xt$Rate_Traffic,xt$ADT_CUR*(1+(.01*xt$TRK_AADT_P)))
Results_1b[9, "Parameter"]    = "AADT * (1 + %Heavy)^2"
Results_1b[9, "Rate_Overall"] =
  cor(xt$Rate_Overall,xt$ADT_CUR*((1+(.01*xt$TRK_AADT_P))^2))
Results_1b[9, "Rate_Traffic"] =
  cor(xt$Rate_Traffic,xt$ADT_CUR*((1+(.01*xt$TRK_AADT_P))^2))
Results_1b[10, "Parameter"]    = "Feet per Mile (2-ways)"
Results_1b[10, "Rate_Overall"] = cor(xt$Rate_Overall, xt$ft_mile)
Results_1b[10, "Rate_Hills"]   = cor(xt$Rate_Hill,    xt$ft_mile)
Results_1b[11, "Parameter"]    = "Sum Elevation Gain (2-ways)"
Results_1b[11, "Rate_Overall"] = cor(xt$Rate_Overall,xt$elev.gain.sum)
Results_1b[11, "Rate_Hills"]   = cor(xt$Rate_Hill,xt$elev.gain.sum)
Results_1b[12, "Parameter"]    = "Max Elevation Gain (1-way)"
Results_1b[12, "Rate_Overall"] = cor(xt$Rate_Overall,xt$elev.gain.max)
Results_1b[12, "Rate_Hills"]   = cor(xt$Rate_Hill,xt$elev.gain.max)
Results_1b[13, "Parameter"]    = "Max. Gradient (1-way, 30m)"
Results_1b[13, "Rate_Overall"] =
  cor(xt$Rate_Overall,xt$grad.max.30m.abs)
Results_1b[13, "Rate_Hills"]   = cor(xt$Rate_Hill,xt$grad.max.30m.abs)
Results_1b[14, "Parameter"]    = "Max. Gradient (1-way, 50m)"
Results_1b[14, "Rate_Overall"] =
  cor(xt$Rate_Overall,xt$grad.max.50m.abs)
Results_1b[14, "Rate_Hills"]   = cor(xt$Rate_Hill,xt$grad.max.50m.abs)
Results_1b[15, "Parameter"]    = "Bike Lane (binary)"
Results_1b[15, "Rate_Overall"] = cor(xt$Rate_Overall, xt$BIKE_LANE)
Results_1b[16, "Parameter"]    = "  HCM OUTSIDE LANE"
Results_1b[16, "Rate_Overall"] = cor(xt$Rate_Overall,
  -.005*xt$LANE_WIDTH_O^2)
Results_1b[16, "Rate_Width"]   = cor(xt$Rate_Width,
  -.005*xt$LANE_WIDTH_O^2)
Results_1b[17, "Parameter"]    = "  HCM SPEED TERM"
Results_1b[17, "Rate_Overall"] = cor(xt$Rate_Overall,
```

```
   0.199*( 1.1199*log(xt$SPD_MAX_FLR-
   20)+.8103)*(1+.1038*(.01*xt$TRK_AADT_P))^2)
 Results_1b[17, "Rate_Speed"]   = cor(xt$Rate_Speed,
   0.199*(1.1199*log(xt$SPD_MAX_FLR-20) +
   .8103)*(1+.1038*(.01*xt$TRK_AADT_P))^2)
 Results_1b[18, "Parameter"]  = "  HCM TRAFFIC TERM"
 Results_1b[18, "Rate_Overall"] = cor(xt$Rate_Overall,
   .507*log(xt$ADT_HR/(4*xt$NUM_LANES)))
 Results_1b[18, "Rate_Traffic"] = cor(xt$Rate_Traffic,
   .507*log(xt$ADT_HR/(4*xt$NUM_LANES)))
 Results_1b[19, "Parameter"]    = "  BLOS 1: Width Classes"
 Results_1b[19, "Rate_Overall"] = cor(xt$Rate_Overall,
   xt$BLOS_1_rankWidth)
 Results_1b[19, "Rate_Width"]   = cor(xt$Rate_Width,
   xt$BLOS_1_rankWidth)
 Results_1b[20, "Parameter"]    = "  BLOS 1: Speed Classes"
 Results_1b[20, "Rate_Overall"] =
   cor(xt$Rate_Overall,xt$BLOS_1_rankSpeed)
 Results_1b[20, "Rate_Speed"]   =
   cor(xt$Rate_Speed,xt$BLOS_1_rankSpeed)
 Results_1b[21, "Parameter"]  = "  BLOS 1: Traffic Classes"
 Results_1b[21, "Rate_Overall"]=cor(xt$Rate_Overall,xt$BLOS_1_rankADT)
 Results_1b[21, "Rate_Traffic"]=cor(xt$Rate_Traffic,xt$BLOS_1_rankADT)

    Results_1b[,2:6] <- round(Results_1b[,2:6],3)
    Results_1b[is.na(Results_1b)] <- '--'

##### Internal Consistency / Participant Reliability #####
 Results_1c <- data.frame(Measure=character(),
     Result=numeric(), Items=double())

    # w/o Hills
    Results_1c[1, "Measure"] = "Cronbach's Alpha"
     temp <- data.frame(Overall=dt$Rate_Overall, Width=dt$Rate_Width,
              Traffic=dt$Rate_Traffic, Speed=dt$Rate_Speed)
    Results_1c[1, "Result"]  = cronbach.alpha(temp)$alpha
    Results_1c[1, "Items"]   = cronbach.alpha(temp)$p

     temp <- cor(temp)
     temp <- ifelse(temp==1, NA, temp)
    Results_1c[3, "Measure"] = "Mean Inter-item Correlation"
    Results_1c[3, "Result"] = mean(temp, na.rm=TRUE)
    Results_1c[3, "Items"]  = 4

    # w/ Hills
    temp <- data.frame(Overall=dt$Rate_Overall, Width=dt$Rate_Width,
      Traffic=dt$Rate_Traffic, Speed=dt$Rate_Speed, Hills=dt$Rate_Hill)

    Results_1c[2, "Measure"] = "--- w/ Hill Rating"
    Results_1c[2, "Result"]  = cronbach.alpha(temp)$alpha
    Results_1c[2, "Items"]  = cronbach.alpha(temp)$p

      temp <- cor(temp)
      temp <- ifelse(temp==1, NA, temp)
    Results_1c[4, "Measure"] = "--- w/ Hill Rating"
    Results_1c[4, "Result"] = mean(temp, na.rm=TRUE)
    Results_1c[4, "Items"]  = 5

##### REGRESSION: USER EVALUATIONS VS OVERALL EVAL #####
 Results_1d <- data.frame(Factor=character(),
     Coef=double(), p_val=double())

 model <- lm(Rate_Overall ~ Rate_Speed+Rate_Traffic+
       Rate_Width+Rate_Hill, data=dt)
```

```
 Results_1d[8,    "Factor"] = "User Eval. Regression, w/ Hill Rating"
 Results_1d[9:13, "Factor"] = names(model$coefficients)
 Results_1d[9:13, "Coef"]   = model$coefficients[1:5]
 Results_1d[9:13, "p_val"]  = summary(model)$coefficients[,4]
 Results_1d[14, "Factor"] = "   Model r-sq"
 Results_1d[14, "Coef"] = summary(model)$r.squared

 model <- lm(Rate_Overall~Rate_Speed+Rate_Traffic+Rate_Width,data=dt)
 Results_1d[1,   "Factor"] = "User Eval. Regression"
 Results_1d[2:5, "Factor"] = names(model$coefficients)
 Results_1d[2:5, "Coef"]   = model$coefficients[1:4]
 Results_1d[2:5, "p_val"]  = summary(model)$coefficients[,4]
 Results_1d[6, "Factor"] = "   Model r-sq"
 Results_1d[6, "Coef"] = summary(model)$r.squared

 Results_1d$Coef <- round(Results_1d$Coef, 3)
 Results_1d$p_val<- round(Results_1d$p_val, 4)

 ## Return Std-Residuals Plot
    jpeg(file="Results_1d.jpg", width=600, height=350)
     plot(model, which = 2)
    dev.off()

##### Stated vs Revealed Preference #####
  # Ranking of Factor Importance: Traffic, Speed, Width
    pt <- data.frame(p = unique(dt$Participant_ID))
    temp <- 0

  ## Regression for individual participants ratings/sample sites
    for (i in seq_along(pt$p)) {
      temp <- dt[dt$Participant_ID == pt[i, "p"],]
      model <- lm(temp$Rate_Overall ~
   temp$Rate_Speed+temp$Rate_Traffic+temp$Rate_Width)

      pt[i, "b0"] <- model$coefficients[1]
      pt[i, "b1"] <- model$coefficients[2]
      pt[i, "b2"] <- model$coefficients[3]
      pt[i, "b3"] <- model$coefficients[4]
      pt[i, "r_sq"] <- round(summary(model)$r.squared,4)
      pt[i, "p_val"]  <- round(summary(model)$coefficients[2,4],4)
    }

    ## find ranks from Regression
    temp_matrix <- t(apply(-pt[,3:5], 1, rank))
    pt$rS = 0; pt$rT = 0; pt$rW = 0
    pt[,8:10] <- temp_matrix

    ## join with stated preference (from demographics table)
    temp <- data.frame(p = de$Index, sS = de$Rank_Speed,
                       sT = de$Rank_Traffic, sW = de$Rank_Width)
    pt <- left_join(pt, temp, by="p")

    Results_1e <- data.frame(Factor = c("Speed", "Traffic", "Width"),
        Stated   = c(mean(pt$sS), mean(pt$sT), mean(pt$sW)),
        Revealed = c(mean(pt$rS), mean(pt$rT), mean(pt$rW)))

#############################################
############ INITIAL BLOS MODELS ############
#############################################

##### PRE-PROCESSING #####
## Aggregate many-to-one (Multiple Responses, One Site)
dt$k = dt$Link_ID
k    <- length(unique(dt$k))
```

```
da <- data.frame(Link_ID = unique(dt$k))
for (i in 1:k) {
  da[i, "Rate_Overall"] <- round(mean(dt$Rate_Overall[dt$k == da[i,
   "Link_ID"]]),2)
  da[i, "Rate_Width"]   <- round(mean(dt$Rate_Width[dt$k == da[i,
   "Link_ID"]]),2)
  da[i, "Rate_Traffic"] <- round(mean(dt$Rate_Traffic[dt$k == da[i,
   "Link_ID"]]),2)
  da[i, "Rate_Speed"]   <- round(mean(dt$Rate_Speed[dt$k == da[i,
   "Link_ID"]]),2)
  da[i, "Rate_Hill"]    <- round(mean(dt$Rate_Hill[dt$k == da[i,
   "Link_ID"]]),2)
  da[i, "n"]            <- length(dt$Rate_Overall[dt$k == da[i,
   "Link_ID"]])
  da[i, "Var_Overall"]  <- round(var(dt$Rate_Overall[dt$k == da[i,
   "Link_ID"]]),2)
}

fa <- left_join(yt, da, by="Link_ID")


## Create new working data table -- "Final Build"

fb <- fa
fb <- fb[!is.na(fb$Link_ID),]

## Select key hillslope attributes
 join_hills <- data.frame(Link_ID = gh$Link_ID, ft_mile = gh$ft_mile,
                          max.grad.50m = gh$grad.max.50m.abs)
fb <- left_join(fb, join_hills, by="Link_ID")
 remove(join_hills)

##### Write Processing Functions #####

 # FUNCTION: Rescale values
 rescale <- function(value_toScale, invert=FALSE, min=1, max=5) {
   if (invert){value_toScale <- value_toScale * -1}
     a <- max(value_toScale)-min(value_toScale)
     b = max-min
   value_toScale <- (((value_toScale-min(value_toScale))*b) / a) + min
   return(value_toScale)
 }

 # FUNCTION: Multiple Models Summary -- Paired t, Correlation,
   Regression
 model_summary <- function(x, y){
   s <- vector()
   model <- t.test(x, y, paired=TRUE)
   s[1]  <- 0
   s[2]  <- round(model$estimate,4)
   s[3]  <- round(model$p.value,4)

   model <- cor.test(x, y)
   s[4]  <- round(model$estimate,4)
   s[5]  <- round(model$p.value,4)

   model <- lm(y~x)
   s[6]  <- round(summary(model)$r.squared,4)
   s[7]  <- round(summary(model)$coefficients[2,4],4)

   s[8]  <- sigma(model)

   return(s) }
```

```
 # FUNCTION: BLOS Equation 5
 run_BLOS5 <- function(LU = 15, CCF = 42, PC = 4){
   ## LU:  'Land Use Intensity'   1-15
   ## CCF: 'Curb Cut Frequency'   1-200
   ## PC:  'Pavement Conditions'  1-5
   BLOS_5_coefficientA1 = .01;
   BLOS_5_coefficientA2 = .01;
   BLOS_5_coefficientA3 = .024

   BLOS5_Vector <- ((fb$ADT_CUR/fb$NUM_LANES)*(14/fb$LANE_WIDTH_O)^2 *
                   (BLOS_5_coefficientA1*(fb$SPD_MAX/30)*

   ((1+(fb$TRK_AADT_P*1/100))^2)+(BLOS_5_coefficientA2*(1/PC))) +
                   BLOS_5_coefficientA3*LU*CCF ) * .1
   BLOS5_Vector <- rescale(BLOS5_Vector, TRUE)
   return(BLOS5_Vector)
 }

#########################################
######## INITIAL MODEL EQUATIONS ########
#########################################

##### MODELS: BLOS EQ. 1 #####
 ## SUM RANKS and Normalize
 fb$BLOS_1 <- as.integer(fb$BLOS_1_rankADT) +
   as.integer(fb$BLOS_1_rankSpeed) + as.integer(fb$BLOS_1_rankWidth)
 fb$BLOS_1 <- rescale(fb$BLOS_1)

##### MODELS: WEIGHTED BLOS EQ. 1 #####
 ## Weights for BLOS 1 with Coefficients from User Eval. Regression

 a =  Results_1d[2, "Coef"]
 b1 = Results_1d[3, "Coef"]
 b2 = Results_1d[4, "Coef"]
 b3 = Results_1d[5, "Coef"]

 fb$BLOS_1_w <- a + (b1*fb$BLOS_1_rankSpeed) + (b2*fb$BLOS_1_rankADT) +
   (b3*fb$BLOS_1_rankWidth)
 fb$BLOS_1_w <- rescale(fb$BLOS_1_w)

##### MODELS: BLOS EQ. 2 #####
 ## Inputs require conversion to metric: meters, km/h
 ## Initial Model: Fixed Parameters for "missing" data
 fb$BLOS_2 <- (fb$ADT_CUR/(fb$NUM_LANES*2500)) +
   ((fb$SPD_MAX*1.60934)/56) +
   (4.25 - (fb$LANE_WIDTH*0.3048))*1.635 + 0

 ## Revised w/ Location Factors (Programmatic Derivations)
 fb$BLOS_2_wLF <- fb$BLOS_2
 fb$BLOS_2_wLF <- ifelse(fb$D1A_mean > median(fb$D1A_mean),
   fb$BLOS_2_wLF+.50, fb$BLOS_2_wLF)
 fb$BLOS_2_wLF <- ifelse(fb$D3b_mean > median(fb$D3b_mean),
   fb$BLOS_2_wLF+.50, fb$BLOS_2_wLF)
 fb$BLOS_2_wLF <- ifelse(fb$ft_mile  > median(fb$ft_mile) ,
   fb$BLOS_2_wLF+.25, fb$BLOS_2_wLF)
 fb$BLOS_2_wLF <- ifelse(fb$ft_mile  > quantile(fb$ft_mile)[4],
   fb$BLOS_2_wLF+.25, fb$BLOS_2_wLF)
 fb$BLOS_2_wLF <- ifelse(fb$LANE_WIDTH_O-fb$LANE_WIDTH > 0,
   fb$BLOS_2_wLF-.75, fb$BLOS_2_wLF)
 fb$BLOS_2_wLF <- ifelse(fb$S_USE_O==1,
   fb$BLOS_2_wLF+.75, fb$BLOS_2_wLF)
 fb$BLOS_2_wLF <- ifelse(fb$S_USE_O==2,
   fb$BLOS_2_wLF+.50, fb$BLOS_2_wLF)
```

```
 ## Revised w/ Location Factor Weights from Epperson (1994)
 fb$BLOS_2_wLF_2 <- fb$BLOS_2
 fb$BLOS_2_wLF_2 <- ifelse(fb$D1A_mean > median(fb$D1A_mean),
   fb$BLOS_2_wLF_2+.25, fb$BLOS_2_wLF_2)
 fb$BLOS_2_wLF_2 <- ifelse(fb$D3b_mean > median(fb$D3b_mean),
   fb$BLOS_2_wLF_2+.25, fb$BLOS_2_wLF_2)
 fb$BLOS_2_wLF_2 <- ifelse(fb$ft_mile  > median(fb$ft_mile) ,
   fb$BLOS_2_wLF_2+.20, fb$BLOS_2_wLF_2)
 fb$BLOS_2_wLF_2 <- ifelse(fb$ft_mile  > quantile(fb$ft_mile)[4],
   fb$BLOS_2_wLF_2+.30, fb$BLOS_2_wLF_2)
 fb$BLOS_2_wLF_2 <- ifelse(fb$LANE_WIDTH_O-fb$LANE_WIDTH > 0,
   fb$BLOS_2_wLF_2-.75, fb$BLOS_2_wLF_2)
 fb$BLOS_2_wLF_2 <- ifelse(fb$S_USE_O==1,
   fb$BLOS_2_wLF_2+.75, fb$BLOS_2_wLF_2)
 fb$BLOS_2_wLF_2 <- ifelse(fb$S_USE_O==2,
   fb$BLOS_2_wLF_2+.25, fb$BLOS_2_wLF_2)

 ## Normalize all BLOS 2 variants
 fb$BLOS_2 <- rescale(fb$BLOS_2, TRUE)
 fb$BLOS_2_wLF <- rescale(fb$BLOS_2_wLF, TRUE)
 fb$BLOS_2_wLF_2 <- rescale(fb$BLOS_2_wLF_2, TRUE)

##### MODELS: BLOS EQ. 3 #####
 ## Initial Model: Fixed Parameters for "missing" data
 fb$BLOS_3 <- (fb$ADT_CUR/(fb$NUM_LANES*2500)) +
   (fb$SPD_MAX/35) + (14 - fb$LANE_WIDTH)/2 + 0

 ## Revised w/ Location Factors
 fb$BLOS_3_wLF <- fb$BLOS_3
 fb$BLOS_3_wLF <- ifelse(fb$D1A_mean > median(fb$D1A_mean),
   fb$BLOS_3_wLF+.50, fb$BLOS_3_wLF)
 fb$BLOS_3_wLF <- ifelse(fb$D3b_mean > median(fb$D3b_mean),
   fb$BLOS_3_wLF+.50, fb$BLOS_3_wLF)
 fb$BLOS_3_wLF <- ifelse(fb$ft_mile  > median(fb$ft_mile) ,
   fb$BLOS_3_wLF+.25, fb$BLOS_3_wLF)
 fb$BLOS_3_wLF <- ifelse(fb$ft_mile  > quantile(fb$ft_mile)[4],
   fb$BLOS_3_wLF+.25, fb$BLOS_3_wLF)
 fb$BLOS_3_wLF <- ifelse(fb$LANE_WIDTH_O-fb$LANE_WIDTH > 0,
   fb$BLOS_3_wLF-.75, fb$BLOS_3_wLF)
 fb$BLOS_3_wLF <- ifelse(fb$S_USE_O==1,
   fb$BLOS_3_wLF+.75, fb$BLOS_3_wLF)
 fb$BLOS_3_wLF <- ifelse(fb$S_USE_O==2,
   fb$BLOS_3_wLF+.50, fb$BLOS_3_wLF)

 fb$BLOS_3 <- rescale(fb$BLOS_3, TRUE)
 fb$BLOS_3_wLF <- rescale(fb$BLOS_3_wLF, TRUE)

##### MODELS: BLOS EQ. 4 #####
 ## Inputs require conversion to metric: meters, km/h
 ## Initial Model: Fixed Parameters for "missing" data
 fb$BLOS_4 <- (fb$ADT_CUR/(fb$NUM_LANES*3100)) +
   ((fb$SPD_MAX*1.60934)/48) +
   ((fb$SPD_MAX*1.60934)/48) * (4.25 - (fb$LANE_WIDTH_O*0.3048))*1.635

 ## Revised w/ Location Factors
 fb$BLOS_4_wLF <- fb$BLOS_4
 fb$BLOS_4_wLF <- ifelse(fb$D1A_mean > median(fb$D1A_mean),
   fb$BLOS_4_wLF+.25, fb$BLOS_4_wLF)
 fb$BLOS_4_wLF <- ifelse(fb$D3a_mean > median(fb$D3a_mean),
   fb$BLOS_4_wLF+.25, fb$BLOS_4_wLF)
 fb$BLOS_4_wLF <- ifelse(fb$ft_mile  > median(fb$ft_mile) ,
   fb$BLOS_4_wLF+.20, fb$BLOS_4_wLF)
```

```
 fb$BLOS_4_wLF <- ifelse(fb$ft_mile  > quantile(fb$ft_mile)[4],
   fb$BLOS_4_wLF+.30, fb$BLOS_4_wLF)
 fb$BLOS_4_wLF <- ifelse(fb$LANE_WIDTH_O-fb$LANE_WIDTH > 0,
   fb$BLOS_4_wLF-.75, fb$BLOS_4_wLF)
 fb$BLOS_4_wLF <- ifelse(fb$S_USE_O==1,
   fb$BLOS_4_wLF+.75, fb$BLOS_4_wLF)
 fb$BLOS_4_wLF <- ifelse(fb$S_USE_O==2,
   fb$BLOS_4_wLF+.25, fb$BLOS_4_wLF)

 fb$BLOS_4 <- rescale(fb$BLOS_4, TRUE)
 fb$BLOS_4_wLF <- rescale(fb$BLOS_4_wLF, TRUE)

##### MODELS: BLOS EQ. 5 #####
 fb$BLOS_5 <- run_BLOS5()

##### MODELS: BLOS EQ. 6 #####
 fb$BLOS_6 <- .760 + (-.005*fb$LANE_WIDTH_O^2) +
   (.507*log(fb$ADT_HR/(4*fb$NUM_LANES))) +
   (0.199*( 1.1199*log(fb$SPD_MAX_FLR-20) +
   .8103)*(1+.1038*(fb$TRK_AADT_P*.01))^2) +
   (7.066 / 4^2)
 fb$BLOS_6 <- rescale(fb$BLOS_6, TRUE)

 ##### PRINT MODELS SUMMARIES
 Results_2a <- data.frame(model=as.character(),source= as.character(),
   paired_diff=as.double(),paired_p=as.double(),
   r=as.double(),r_p=as.double(),
   r_sq=as.double(),r_sq_p=as.double(),sigma=as.double())

   ## MODELS LIST (model='name', source='source column')
   Results_2a[1:11, "model"]  <- c("BLOS 1", "BLOS 1, weighted",
     "BLOS 2", "BLOS 3", "BLOS 2 w/LF", "BLOS 2 w/LF 2", "BLOS 3 w/LF",
     "BLOS 4", "BLOS 4 w/LF", "BLOS 5", "BLOS 6")
   Results_2a[1:11, "source"] <- c("BLOS_1", "BLOS_1_w", "BLOS_2",
     "BLOS_3", "BLOS_2_wLF", "BLOS_2_wLF_2", "BLOS_3_wLF", "BLOS_4",
     "BLOS_4_wLF",  "BLOS_5", "BLOS_6")

   ## Models' Summary Results
   for (i in seq_along(Results_2a$source)) {
     y <- fb$Rate_Overall
     x <- fb[,Results_2a[i,"source"]]
     s <- model_summary(x, y)
     Results_2a[i, 3:9] <- s[2:8]
   }

 #####################################
 ##### STEPWISE REGRESSION MODEL #####
 #####################################

 ## REGRESSION OF REAL ROADWAY ATTRIBUTES TO USER RATINGS
 mod <- fb$Rate_Overall ~ fb$SPD_MAX_FLR + log(fb$ADT_CUR) +
   fb$LANE_WIDTH_O + fb$TRK_AADT_P +
   log(fb$D1B_mean) +     log(fb$D3a_mean) +  fb$ft_mile + fb$BIKE_LANE

 step(lm(mod), direction="both")

 Results_3a <- data.frame()
 Results_3a[1, "Model"] = "fb$Rate_Overall ~ fb$SPD_MAX_FLR +
   log(fb$ADT_CUR) + fb$LANE_WIDTH_O + fb$TRK_AADT_P +
      fb$D1B_mean + fb$D3a_mean + fb$ft_mile"

 Results_3a[2, "Model"] = "fb$Rate_Overall ~ fb$SPD_MAX_FLR +
   log(fb$ADT_CUR) + fb$LANE_WIDTH_O + log(fb$D3a_mean)"
```

```
Results_3a[3, "Model"] = "fb$Rate_Overall ~ fb$SPD_MAX_FLR +
  log(fb$ADT_CUR) +  fb$LANE_WIDTH_O"

Results_3a[4, "Model"] = "fb$Rate_Overall ~ fb$SPD_MAX_FLR +
  log(fb$ADT_CUR) + fb$LANE_WIDTH_O + fb$D3a_mean + fb$BIKE_LANE"

Results_3a[5, "Model"] = "fb$Rate_Overall ~ fb$SPD_MAX_FLR +
  log(fb$ADT_CUR) + fb$LANE_WIDTH_O + fb$D1B_mean + fb$BIKE_LANE"

## SUMMARY OF BEST FIT REGRESSIONS
for (i in seq_along(Results_3a$Model)){
  Results_3a[i, "r_sq"] <- summary(lm(as.formula(Results_3a[i,
  "Model"])))$r.squared
}

## MODEL SPECIFIC RESULTS
model <- lm(Results_3a[1, "Model"])
Results_3b_1 <- data.frame(
  Variable = variable.names(model),
  Coefficient = summary(model)$coefficients[,1],
  p_val     = round(summary(model)$coefficients[,4],4)
)

model <- lm(Results_3a[2, "Model"])
Results_3b_2 <- data.frame(
  Variable = variable.names(model),
  Coefficient = summary(model)$coefficients[,1],
  p_val     = round(summary(model)$coefficients[,4],4)
)

model <- lm(Results_3a[3, "Model"])
Results_3b_3 <- data.frame(
  Variable = variable.names(model),
  Coefficient = summary(model)$coefficients[,1],
  p_val     = round(summary(model)$coefficients[,4],4)
)

model <- lm(Results_3a[4, "Model"])
Results_3b_4 <- data.frame(
  Variable = variable.names(model),
  Coefficient = summary(model)$coefficients[,1],
  p_val     = round(summary(model)$coefficients[,4],4)
)

model <- lm(Results_3a[5, "Model"])
Results_3b_5 <- data.frame(
  Variable = variable.names(model),
  Coefficient = summary(model)$coefficients[,1],
  p_val     = round(summary(model)$coefficients[,4],4)
)

## SAVE "BEST" REGRESSION RESULTS
fb$BLOS_REGR <- model$fitted.values

####################################
######## Density Experiments ########
####################################

# Checking for Modeled Fit Against:
y = fb$Rate_Overall

## Blank Results Table for all Permutations
Results_4a <- data.frame(Model=as.character(),
    Paired_diff=as.double(), Paired_p=as.double(),
```

```
    r=as.double(), r_p=as.double(),
    r_sq=as.double(), r_sq_p=as.double(), sigma=as.double())

## BASELINE FIXED PARAMETERS
Results_4a[1,1] <- "Fixed LU = 15, Fixed CCF = 42"
Results_4a[1,2:8] <- model_summary(run_BLOS5(LU=15, CCF=42, PC=4),
  y)[2:8]

## DENSITY FOR LU
Results_4a[2,1] <- "LU = Emp. Density, Fixed CCF = 42"
Results_4a[2,2:8] <-  model_summary(run_BLOS5(
  LU=rescale(fb$D1C_mean, FALSE, 1, 15), CCF=42), y)[2:8]

Results_4a[3,1] <- "LU = Pop. Density, Fixed CCF = 42"
Results_4a[3,2:8] <- model_summary(run_BLOS5(
  LU=rescale(fb$D1B_mean, FALSE, 1, 15), CCF=42), y)[2:8]

Results_4a[4,1] <- "LU = Res. Density, Fixed CCF = 42"
Results_4a[4,2:8] <- model_summary(run_BLOS5(
  LU=rescale(fb$D1A_mean, FALSE, 1, 15), CCF=42), y)[2:8]

## DENSITY FOR CCF
Results_4a[6,1] <- "Fixed Density, CCF = Intersection Density"
Results_4a[6,2:8] <- model_summary(run_BLOS5(
  LU=15, CCF=rescale(fb$D3b_mean, FALSE, 1, 200)), y)[2:8]

Results_4a[7,1] <- "Fixed Density, CCF = Road Density"
Results_4a[7,2:8] <- model_summary(run_BLOS5(
  LU=15, CCF=rescale(fb$D3a_mean, FALSE, 1, 200)), y)[2:8]


## DENSITY FOR LU AND FOR CCF
Results_4a[9,1] <- "LU = Pop. Density, CCF = Road Density"
Results_4a[9,2:8] <- model_summary(run_BLOS5(
  LU=rescale(fb$D1B_mean, FALSE, 1, 15),
  CCF=rescale(fb$D3a_mean, FALSE, 1, 200)), y)[2:8]

Results_4a[10,1] <- "LU = Res. Density, CCF = Road Density"
Results_4a[10,2:8] <- model_summary(run_BLOS5(
  LU=rescale(fb$D1A_mean, FALSE, 1, 15),
  CCF=rescale(fb$D3a_mean, FALSE, 1, 200)), y)[2:8]

Results_4a[11,1] <- "LU = Emp. Density, CCF = Road Density"
Results_4a[11,2:8] <- model_summary(run_BLOS5(
  LU=rescale(fb$D1C_mean, FALSE, 1, 15),
  CCF=rescale(fb$D3a_mean, FALSE, 1, 200)), y)[2:8]

Results_4a[13,1] <- "LU = Pop. Density, CCF = Intersection Density"
Results_4a[13,2:8] <- model_summary(run_BLOS5(
  LU=rescale(fb$D1B_mean, FALSE, 1, 15),
  CCF=rescale(fb$D3b_mean, FALSE, 1, 200)), y)[2:8]

Results_4a[14,1] <- "LU = Res. Density, CCF = Intersection Density"
Results_4a[14,2:8] <- model_summary(run_BLOS5(
  LU=rescale(fb$D1A_mean, FALSE, 1, 15),
  CCF=rescale(fb$D3b_mean, FALSE, 1, 200)), y)[2:8]

Results_4a[15,1] <- "LU = Emp. Density, CCF = Intersection Density"
Results_4a[15,2:8] <- model_summary(run_BLOS5(
  LU=rescale(fb$D1C_mean, FALSE, 1, 15),
  CCF=rescale(fb$D3b_mean, FALSE, 1, 200)), y)[2:8]
```

```
## PRESENT THREE_D TABLE
Results_4a$Paired_p <- as.double(format(Results_4a$Paired_p,
  nsmall=4))
Results_4a$sigma <- round(Results_4a$sigma, 3)
## CORRELATION MATRIX FOR DENSITIES, ROADWAY ATTRIBUTES
Results_4a_2 <- round(cor(fb[,
  c("SPD_MAX", "LANE_WIDTH_O", "ADT_CUR", "D1A_mean", "D1B_mean",
  "D1C_mean", "D3a_mean", "D3b_mean")]), 4)

#######################################
######## Hillslope Experiments ########
#######################################

# Checking for Modeled Fit Against:
y = fb$Rate_Overall

## Blank Results Table for all Permutations
Results_4b <- data.frame(Model=as.character(),
  Paired_diff=as.double(), Paired_p=as.double(),
  r=as.double(), r_p=as.double(),
  r_sq=as.double(), r_sq_p=as.double(), sigma=as.double())

## BASELINE FIXED PARAMETERS
Results_4b[1,1] <- "Fixed LU = 15, Fixed CCF = 42"
Results_4b[1,2:8] <- model_summary(run_BLOS5(LU=15, CCF=42, PC=4),
  y)[2:8]

Results_4b[2,1] <- "Fixed Density, CCF = Road Density"
Results_4b[2,2:8] <- model_summary(run_BLOS5(
  LU=15, CCF=rescale(fb$D3a_mean, FALSE, 1, 200)), y)[2:8]

## Add Hillslope (substitution for Pavement Condition)
Results_4b[4,1] <- "LU = Emp. Density, CCF = Road Density, PaveCon =
  ft_mile"
Results_4b[4,2:8] <- model_summary(run_BLOS5(
  LU =rescale(fb$D1C_mean, FALSE, 1, 15),
  CCF=rescale(fb$D3a_mean, FALSE, 1, 200),
  PC =rescale(fb$ft_mile,  TRUE,  1, 5)), y)[2:8]

## Add Hillslope (substitution for Pavement Condition)
Results_4b[5,1] <- "LU = Fixed Density, CCF = Road Density, PaveCon =
  ft_mile"
Results_4b[5,2:8] <- model_summary(run_BLOS5(
  CCF=rescale(fb$D3a_mean, FALSE, 1, 200),
  PC =rescale(fb$ft_mile,  TRUE,  1, 5)), y)[2:8]

## Add Hillslope (substitution for Pavement Condition)
Results_4b[6,1] <- "LU = Emp. Density, CCF = Road Density, PaveCon =
  max grade (50)"
Results_4b[6,2:8] <- model_summary(run_BLOS5(
  LU =rescale(fb$D1C_mean, FALSE, 1, 15),
  CCF=rescale(fb$D3a_mean, FALSE, 1, 200),
  PC =rescale(fb$max.grad.50m,  TRUE,  1, 5)), y)[2:8]

## Add Hillslope (substitution for Pavement Condition)
Results_4b[7,1] <- "LU = Fixed Density, CCF = Road Density, PaveCon =
  max grade (50)"
Results_4b[7,2:8] <- model_summary(run_BLOS5(
  CCF=rescale(fb$D3a_mean, FALSE, 1, 200),
  PC =rescale(fb$max.grad.50m,  TRUE,  1, 5)), y)[2:8]

## Add Hillslope (substitution for Pavement Condition)
Results_4b[8,1] <- "LU = Fixed Density, CCF = Road Density, PaveCon =
  13% Hill Penalty"
```

```
Results_4b[8,2:8] <- model_summary(run_BLOS5(
  CCF=rescale(fb$D3a_mean, FALSE, 1, 200),
  PC = ifelse(fb$max.grad.50m > .13, 2, 5) ), y)[2:8]

## Add Hillslope (substitution for Pavement Condition)
Results_4b[9,1] <- "LU = Fixed Density, CCF = Road Density, PaveCon =
  8% Hill Penalty"
Results_4b[9,2:8] <- model_summary(run_BLOS5(
  CCF=rescale(fb$D3a_mean, FALSE, 1, 200),
  PC = ifelse(fb$max.grad.50m > .08, 2, 5) ), y)[2:8]

Results_4b$sigma <- round(Results_4b$sigma, 3)

#################################
##### REVISED BLOS MODELS #######
#################################

## Baselines (Eq. 5 and 6)
fb$BLOS_6 <- .760 + (-.005*fb$LANE_WIDTH_O^2) +
  (.507*log(fb$ADT_HR/(4*fb$NUM_LANES))) +
  (0.199*( 1.1199*log(fb$SPD_MAX_FLR-20) +
  .8103)*(1+.1038*(fb$TRK_AADT_P*.01))^2) +
  (7.066 / 4^2)

## Revised Eq. 5 (Density for LU/CCF, add Hillslope)
fb$BLOS_5_Rev <- run_BLOS5(
  LU =rescale(fb$D1B_mean, FALSE, 1, 15),
  CCF=rescale(fb$D3a_mean, FALSE, 1, 200),
  PC =rescale(fb$ft_mile,  TRUE,  1, 5))

## Revised Eq. 6 (HCM)
fb$BLOS_6_Rev <- .760 + (-.005*fb$LANE_WIDTH_O^2) +
  (.507*log(fb$ADT_HR/(4*fb$NUM_LANES))) +
  (0.199*( 1.1199*log(fb$SPD_MAX_FLR-20) +
  .8103)*(1+.1038*(fb$TRK_AADT_P*.01))^2) +
  -0.5*(1/rescale(fb$D1B_mean)) +
  0.1*(1/rescale(fb$ft_mile, TRUE, 1, 5)) +
  -1 * fb$BIKE_LANE

## w/ Location Factors
fb$BLOS_6_wLF <- fb$BLOS_6
fb$BLOS_6_wLF <- ifelse(fb$D1A_mean > median(fb$D1A_mean),
  fb$BLOS_6_wLF+.25, fb$BLOS_6_wLF)
fb$BLOS_6_wLF <- ifelse(fb$D3a_mean > median(fb$D3a_mean),
  fb$BLOS_6_wLF+.25, fb$BLOS_6_wLF)
fb$BLOS_6_wLF <- ifelse(fb$ft_mile  > median(fb$ft_mile) ,
  fb$BLOS_6_wLF+.20, fb$BLOS_6_wLF)
fb$BLOS_6_wLF <- ifelse(fb$ft_mile  > quantile(fb$ft_mile)[4],
  fb$BLOS_6_wLF+.30, fb$BLOS_6_wLF)
fb$BLOS_6_wLF <- ifelse(fb$LANE_WIDTH_O-fb$LANE_WIDTH > 0,
  fb$BLOS_6_wLF-.50, fb$BLOS_6_wLF)
fb$BLOS_6_wLF <- ifelse(fb$BIKE_LANE == 1,
  fb$BLOS_6_wLF-.50, fb$BLOS_6_wLF)
fb$BLOS_6_wLF <- ifelse(fb$S_USE_O==1,
  fb$BLOS_6_wLF+.50, fb$BLOS_6_wLF)
fb$BLOS_6_wLF <- ifelse(fb$S_USE_O==2,
  fb$BLOS_6_wLF+.25, fb$BLOS_6_wLF)

fb$BLOS_6 <- rescale(fb$BLOS_6, TRUE)
fb$BLOS_6_wLF <- rescale(fb$BLOS_6_wLF, TRUE)
fb$BLOS_6_Rev <- rescale(fb$BLOS_6_Rev, TRUE)
```

```
##### PRINT REVISED MODELS SUMMARIES #####
Results_5a <- data.frame(model  = as.character(), source=
  as.character(),paired_diff=as.double(), paired_p=as.double(),
  r=as.double(), r_p=as.double(),
  r_sq=as.double(), r_sq_p=as.double(), sigma=as.double())

## MODELS LIST (model='name', source='source column')
Results_5a[1:6, "model"]  <- c("REGRESSION", "BLOS 5", "BLOS 5 Rev.",
  "BLOS 6",       "BLOS 6 w/ LF",     "BLOS 6 Rev.")
Results_5a[1:6, "source"] <- c("BLOS_REGR", "BLOS_5", "BLOS_5_Rev",
  "BLOS_6", "BLOS_6_wLF", "BLOS_6_Rev")

## Models' Summary Results
for (i in seq_along(Results_5a$source)) {
  y <- fb$Rate_Overall
  x <- fb[,Results_5a[i,"source"]]
  Results_5a[i, 3:9] <- model_summary(x,y)[2:8]
}

##################################
######  CONTEXTUAL ANALYSIS ######
##################################

Results_6a <- data.frame(Link_ID  = as.integer(), Name =
  as.character(),BLOS=as.double(),Rating=as.double(),
  Width_O=as.numeric(),Traffic=as.numeric(),Speed=as.numeric(),
  Bike_Lane=as.numeric(),Max_Grad=as.double(), Ft_Mi=as.double(),
  Res_Dens=as.double())

Results_6a[1:8,"Link_ID"] <- c(16, 26, 17, 33, 34, 107, 21, 74)
for (i in seq_along(Results_6a$Link_ID)){
  temp <- fb[fb$Link_ID == Results_6a[i, "Link_ID"],  ]
  Results_6a[i, "Name"]      <- temp$STE_NAM
  Results_6a[i, "BLOS"]      <- temp$BLOS_6_Rev
  Results_6a[i, "Rating"]    <- temp$Rate_Overall
  Results_6a[i, "Width_O"]   <- temp$LANE_WIDTH_O
  Results_6a[i, "Traffic"]   <- temp$ADT_CUR
  Results_6a[i, "Speed"]     <- temp$SPD_MAX_FLR
  Results_6a[i, "Bike_Lane"] <- temp$BIKE_LANE
  Results_6a[i, "Max_Grad"]  <- temp$max.grad.50m
  Results_6a[i, "Ft_Mi"]     <- temp$ft_mile
  Results_6a[i, "Res_Dens"]  <- temp$D1B_mean
 }

## HILLS CONTEXT
temp <- c(130,33,991,19,30,24,35,39,8)
gb <- fb[(fb$Link_ID %in% temp),]
gb <- gb[order(gb$ft_mile),]
  gb$Perfect <- c(1,1,1,1,1,0,1,0)
  gb$Steep   <- c(0,0,0,0,0,1,1,1)

#Fulton Ranch
gb[9, "STE_NAM"] <- "FULTON RANCH RD"
gb[9, "ft_mile"]      <-   363
gb[9, "max.grad.50m"] <- .209
gb[9, "Perfect"] <-    0
gb[9, "Steep"]   <-    1
gb[9, "Rate_Hill"]    <-    "~1"
gb[4, "STE_NAM"] <- "WONDER WORLD / 12"
gb <- gb[order(gb$ft_mile),];  Results_A3 <- gb[,c("Link_ID","STE_NAM",
  "ft_mile","max.grad.50m","Perfect","Steep","Rate_Hill")]
```

```
###################################
#### FINAL COMPREHENSIVE MODEL ####
###################################

## ROAD EMPIRICS (Topologically Corrected, Roadway Attributes)
rt <- Roads_Empirics
  ## Remove "Functional Interstate Hwys"
  rt <- rt[!(rt$F_SYSTEM == 1),]

  ## Check and fill null values where needed
  rt[is.na(rt)] <- 0

  ## Limit Shoulder Width
  length(rt$S_WID_O[rt$S_WID_O>16])
  rt$S_WID_O <- ifelse(rt$S_WID_O > 16, 16, rt$S_WID_O)

  ## Limit Lane Width
  length(rt$LANE_WIDTH[rt$LANE_WIDTH>16])
  rt$LANE_WIDTH <- ifelse(rt$LANE_WIDTH >= 16, 16, rt$LANE_WIDTH)

#### DERIVED ATTRIBUTES ####

## Outside Lane Width -- Checks that Shoulder is paved, not parking.
rt$LANE_WIDTH_O <- rt$LANE_WIDTH
for (i in seq_along(rt$LANE_WIDTH_O)) {
  if (!is.na(rt[i, "S_WID_O"])) {
    rt[i,"LANE_WIDTH_O"] <- rt[i,"LANE_WIDTH"]+rt[i,"S_WID_O"]
    if (rt[i, "S_USE_O"] == 1 || rt[i, "S_USE_O"] == 2) {
      rt[i,"LANE_WIDTH_O"] <- rt[i,"LANE_WIDTH"]}
    if (rt[i, "S_TYPE_O"] > 2 || rt[i, "S_TYPE_O"] < 1) {
      rt[i,"LANE_WIDTH_O"] <- rt[i,"LANE_WIDTH"]}
  }}

## Effective Shoulder Width:
rt$SHOULDER_WIDTH <- rt$LANE_WIDTH_O - rt$LANE_WIDTH

## Speed Limit, HCM FLOOR
rt$SPD_MAX[is.na(rt$SPD_MAX)] <- 0
rt$SPD_MAX_FLR <- ifelse( rt$SPD_MAX < 21,   21, rt$SPD_MAX)

## ADT HOURLY, HCM FLOOR
rt$ADT_HR <- ifelse((rt$ADT_CUR/24)<(4*rt$NUM_LANES), 4*rt$NUM_LANES,
   rt$ADT_CUR/24)

####### CALCULATE BLOS #####
rt$BLOS_6_Rev <- .760 + (-.005*rt$LANE_WIDTH_O^2) +
  (.507*log(rt$ADT_HR/(4*rt$NUM_LANES))) +
  (0.199*( 1.1199*log(rt$SPD_MAX_FLR-20) +
   .8103)*(1+.1038*(rt$TRK_AADT_P*.01))^2) +
  .1*(1/rescale(rt$D1B_mean,TRUE,1,5)) +
  -1 * rt$BIKE_LANE

## REMOVE ANY INCALCULABLE ROWS
rt <- rt[!is.na(rt$BLOS_6_Rev),]

## RECAST OUTLIERS AS MINIMA/MAXIMA
    t_min <- quantile(rt$BLOS_6_Rev, probs=.05)
    t_max <- quantile(rt$BLOS_6_Rev, probs=.95)
  rt$BLOS_6_Rev[rt$BLOS_6_Rev <= t_min] <- t_min
  rt$BLOS_6_Rev[rt$BLOS_6_Rev >= t_max] <- t_max

rt$BLOS_6_Rev <- rescale(rt$BLOS_6_Rev, TRUE, 0.001, 5)
```

APPENDIX E: Results Tables

Table 12: Results from the initial six BLOS models

Model	$\bar{d}$	p-val	r	p-val	r^2	p-val	RMSE
BLOS 1	-0.44	0.00	-0.02	0.80	0.00	0.80	1.18
BLOS 1, weighted	-0.24	0.04	0.13	0.14	0.02	0.14	1.17
BLOS 2	-0.55	0.00	0.14	0.11	0.02	0.11	1.17
BLOS 3	-0.55	0.00	0.14	0.11	0.02	0.11	1.17
BLOS 2 w/LF	-0.51	0.00	0.24	0.01	0.06	0.01	1.15
BLOS 2 w/LF 2	-0.67	0.00	0.19	0.03	0.03	0.03	1.16
BLOS 3 w/LF	-0.51	0.00	0.23	0.01	0.06	0.01	1.15
BLOS 4	-1.06	0.00	0.05	0.54	0.00	0.54	1.18
BLOS 4 w/LF	-1.17	0.00	0.06	0.47	0.00	0.47	1.18
BLOS 5	0.77	0.00	0.48	0.00	0.23	0.00	1.04
BLOS 6	-0.42	0.00	0.38	0.00	0.14	0.00	1.10

Table 13: Results from density experimentation.

Model	$\bar{d}$	p-val	r^2	p-val	RMSE
Fixed LU = 15, Fixed CCF = 42	0.77	0.00	0.23	0.00	1.04
LU = Emp. Density, Fixed CCF = 42	0.76	0.00	0.24	0.00	1.03
LU = Pop. Density, Fixed CCF = 42	0.73	0.00	0.25	0.00	1.03
LU = Res. Density, Fixed CCF = 42	0.71	0.00	0.25	0.00	1.03
Fixed Density, CCF = Intersection Density	0.59	0.00	0.27	0.00	1.01
Fixed Density, CCF = Road Density	0.46	0.00	0.30	0.00	0.99
LU = Pop. Density, CCF = Road Density	0.59	0.00	0.29	0.00	0.99
LU = Res. Density, CCF = Road Density	0.57	0.00	0.28	0.00	1.00
LU = Emp. Density, CCF = Road Density	0.68	0.00	0.25	0.00	1.02
LU = Pop. Density, CCF = Intersection Density	0.64	0.00	0.27	0.00	1.01
LU = Res. Density, CCF = Intersection Density	0.62	0.00	0.26	0.00	1.02
LU = Emp. Density, CCF = Intersection Density	0.71	0.00	0.25	0.00	1.03
Fixed LU = 15, Fixed CCF = 42	0.77	0.00	0.23	0.00	1.04

Table 14: Correlation matrix of urban densities.

	SPD_MAX	LANE_WIDTH_O	ADT_CUR	D1A_mean	D1B_mean	D1C_mean	D3a_mean
SPD_MAX							
LANE_WIDTH_O	0.42						
ADT_CUR	0.25	0.56					
D1A_mean	-0.51	0.03	0.28				
D1B_mean	-0.67	-0.05	0.08	0.67			
D1C_mean	-0.27	0.00	0.13	0.63	0.43		
D3a_mean	-0.59	0.02	0.27	0.91	0.77	0.54	
D3b_mean	-0.58	-0.01	0.17	0.82	0.77	0.52	0.95

Table 15: Results from including hillslope in BLOS models.

Model	$\bar{d}$	**p-val**	r^2	**p-val**	**RMSE**
Fixed LU = 15, Fixed CCF = 42	0.77	0.00	0.23	0.00	1.04
Fixed LU = 15, CCF = Road Density	0.46	0.00	0.30	0.00	0.99
LU = Emp. Density, CCF = Road Density, PaveCon = ft_mile	0.67	0.00	0.26	0.00	1.02
LU = Fixed Density, CCF = Road Density, PaveCon = ft_mile	0.50	0.00	0.30	0.00	0.99
LU = Emp. Density, CCF = Road Density, PaveCon = max grade (50)	0.67	0.00	0.25	0.00	1.02
LU = Fixed Density, CCF = Road Density, PaveCon = max grade (50)	0.46	0.00	0.29	0.00	1
LU = Fixed Density, CCF = Road Density, PaveCon = 13% Hill Penalty	0.44	0.00	0.30	0.00	0.99
LU = Fixed Density, CCF = Road Density, PaveCon = 8% Hill Penalty	0.44	0.00	0.29	0.00	1

Table 16: Results from revised BLOS models.

Model	$\bar{d}$	**p-val**	r	**p-val**	r^2	**p-val**	**RMSE**
REGRESSION	0.00	1.00	0.60	0.00	0.36	0.00	0.95
BLOS 5	0.77	0.00	0.48	0.00	0.23	0.00	1.04
BLOS 5 Rev.	0.57	0.00	0.54	0.00	0.29	0.00	1.00
BLOS 6	-0.42	0.00	0.38	0.00	0.14	0.00	1.10
BLOS 6 w/ LF	-0.18	0.09	0.42	0.00	0.18	0.00	1.07
BLOS 6 Rev.	-0.19	0.07	0.48	0.00	0.23	0.00	1.04

APPENDIX F: Hill Slope Results

Although hillslope derivations demand careful attention to topology, the resulting measures were reliable for the majority of sampled links. Hill slope preferences unsurprisingly varied significantly across participants (Table 17); nonetheless, their evaluations may be useful for classification of extreme hill slopes – slopes that might warrant representation on user maps. In total, 8 links from this research were commented as either “the perfect hill” or “too steep” of a hill (Figure 17); one additional link was also frequently mentioned – the segment of Fulton Ranch Road which gruelingly ascends over 300 feet from the Blanco River in about half of a mile. Participants comfortability with a link’s hillslope declined significantly around max sustained gradients of 9-10%. Even where gradients peak at 8-12%, if these grades are sustained (evidenced by higher feet per mile), participant’s assessment of the hill rating also decreased. At gradients approaching 20%, even high-mileage bicyclists became uncomfortable. For the moment, the suggestion follows that bicycle route maps should highlight gradients exceeding 8% and mark segments exceeding 100 feet per mile (an average gradient of just 2%).

Table 17: Hillslope derivation results and participant assessments.

				Participants		
Link_ID	Street Name	ft_mile	max.grad	Perfect	Steep	BLOS
8	OLD BASTROP HWY	44	0.096	1	0	3.5
130	OLD BASTROP RD	53	0.060	1	0	2
33	FULTON RANCH RD	79	0.116	1	0	2.5
39	WONDER WORLD / 12	88	0.081	1	0	2.5
19	BURLESON	119	0.087	1	0	2
30	ACADEMY ST	140	0.132	0	1	1.67
35	N LBJ DR	154	0.111	1	1	1
24	SESSOM	236	0.234	0	1	1.2
NA	FULTON RANCH RD	363	0.209	0	1	~1

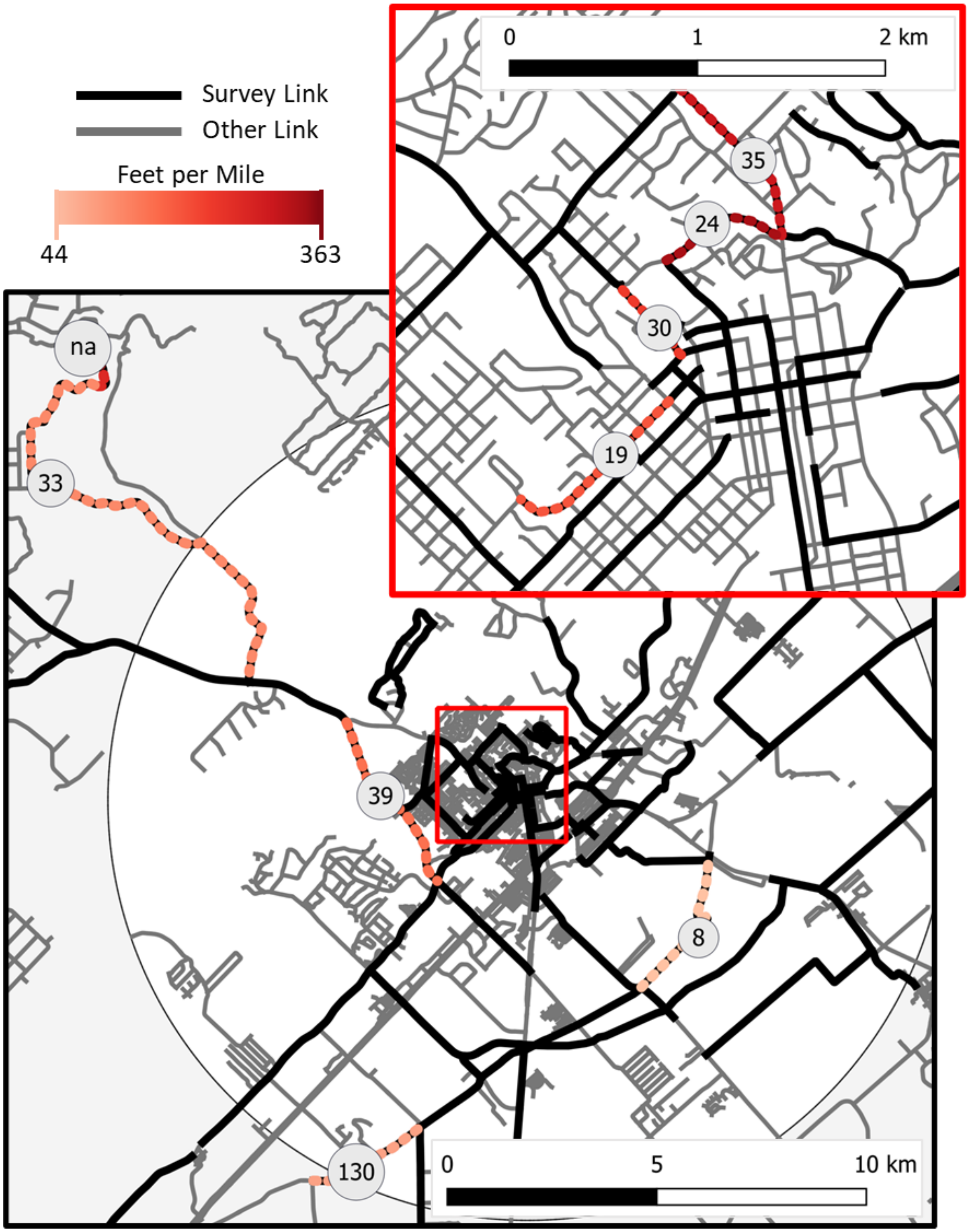

Figure 17: Map of participants' perfect and steep hills. (Author's illustration.)

REFERENCES

Abbas, S. K. S., Adnan, M. A., & Endut, I. R. (2011). Exploration of 85th percentile operating speed model on horizontal curve: A case study for two-lane rural highways. *Procedia - Social and Behavioral Sciences*, *16*, 352–363. https://doi.org/10.1016/j.sbspro.2011.04.456

Asadi-Shekari, Z., Moeinaddini, M., & Zaly Shah, M. (2013). Non-motorised level of service: Addressing challenges in pedestrian and bicycle level of service. *Transport Reviews*, *33*(2), 166–194. https://doi.org/10.1080/01441647.2013.775613

Axhausen, K. W., & Smith Jr, R. L. (1986). Bicyclist link evaluation: A stated-preference approach. *Transportation Research Record*, *1085*(1), 7–15.

Beck, B., Chong, D., Olivier, J., Perkins, M., Tsay, A., Rushford, A., Li, L., Cameron, P., Fry, R., & Johnson, M. (2019). How much space do drivers provide when passing cyclists? Understanding the impact of motor vehicle and infrastructure characteristics on passing distance. *Accident Analysis & Prevention*, *128*, 253–260. https://doi.org/10.1016/j.aap.2019.03.007

Becker, W. E., & Kennedy, P. E. (1992). A graphical exposition of the ordered probit. *Econometric Theory*, *8*(1), 127–131.

Ben-Akiva, M. E. (1973). *Structure of Passenger Travel Demand Models* [Dissertation]. Massachusetts Institute of Technology.

Beura, S. K., & Bhuyan, P. K. (2017). Development of a bicycle level of service model for urban street segments in mid-sized cities carrying heterogeneous traffic: A functional networks approach. *Journal of Traffic and Transportation Engineering (English Edition)*, *4*(6), 503–521. https://doi.org/10.1016/j.jtte.2017.02.003

Beura, S. K., Chellapilla, H., & Bhuyan, P. K. (2017). Urban road segment level of service based on bicycle users' perception under mixed traffic conditions. *Journal of Modern Transportation*, *25*(2), 90–105. https://doi.org/10.1007/s40534-017-0127-9

Bíl, M., Andrášik, R., & Kubeček, J. (2015). How comfortable are your cycling tracks? A new method for objective bicycle vibration measurement. *Transportation Research Part C: Emerging Technologies*, *56*, 415–425. https://doi.org/10.1016/j.trc.2015.05.007

Bogel-Burroughs, N. (2019). Deadliest year for pedestrians and cyclists in U.S. since 1990. *New York Times*. https://www.nytimes.com/2019/10/22/us/pedestrian-cyclist-deaths-traffic.html

Bovy, P. H., & Stern, E. (1990). *Route Choice: Wayfinding in Transport Networks: Wayfinding in Transport Networks*. Springer Netherlands. https://doi.org/10.1007/978-94-009-0633-4

Box, G. E. (1990). George's column: Do interactions matter? *Quality Engineering*, *2*(3), 365–369. https://doi.org/10.1080/08982119008962728

Bradley, P. S., & Fayyad, U. M. (1998). Refining initial points for K-means clustering. *Proceedings of the 15th International Conference on Machine Learning*, 91–99.

Broach, J., Dill, J., & Gliebe, J. (2012). Where do cyclists ride? A route choice model developed with revealed preference GPS data. *Transportation Research Part A: Policy and Practice*, *46*(10), 1730–1740. https://doi.org/10.1016/j.tra.2012.07.005

Buzzelli, M. (2020). Modifiable areal unit problem. *International Encyclopedia of Human Geography*, 169–173. https://doi.org/10.1016/B978-0-08-102295-5.10406-8

Cabral, L. (2019). *Analyzing Network Connectivity by Cyclist Comfort: An Empirical Reappraisal of the Four Types of Cyclists Typology and Level of Traffic Stress Framework* [Dissertation]. University of Alberta.

Cabral, L., & Kim, A. M. (2020). An empirical reappraisal of the four types of cyclists. *Transportation Research Part A: Policy and Practice*, *137*, 206–221. https://doi.org/10.1016/j.tra.2020.05.006

Cavagnaro, H. (2019). *San Marcos restripes Guadalupe Street to include dedicated bike lanes, drawing mixed reviews*. KVUE. https://www.kvue.com/article/news/local/san-marcos-restripes-guadalupe-street-to-include-dedicated-bike-lanes-drawing-mixed-reviews/269-129de0e2-fa5c-4a44-b2ca-7288bd57bf6d

Chaloux, N., & El-Geneidy, A. (2019). Rules of the road: Compliance and defiance among the different types of cyclists. *Transportation Research Record*, *2673*(9), 34–43. https://doi.org/10.1177/0361198119844965

Cox, D. R. (1984). Interaction. *International Statistical Review*, *52*(1), 1–24. https://doi.org/10.2307/1403235

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*, 102368. https://doi.org/10.1016/j.trd.2020.102368

Damant-Sirois, G., Grimsrud, M., & El-Geneidy, A. M. (2014). What's your type: A multidimensional cyclist typology. *Transportation*, *41*(6), 1153–1169. https://doi.org/10.1007/s11116-014-9523-8

Devasurendra, K. W., Wirasinghe, S. C., & Kattan, L. (2020, August). *A Critical Review of Transit Level of Service Measures and an Overview of a Proposed New Approach*. International Conference on Transportation and Development 2020; American Society of Civil Engineers. https://trid.trb.org/view/1735658

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., & Carr, T. (2003). Bicycle commuting and facilities in major U.S. cities: If you build them, commuters will use them. *Transportation Research Record*, *1828*(1), 116–123. https://doi.org/10.3141/1828-14

Dill, J., & McNeil, N. (2013). Four types of cyclists?: Examination of typology for better understanding of bicycling behavior and potential. *Transportation Research Record*, *2387*(1), 129–138. https://doi.org/10.3141/2387-15

Dill, J., & McNeil, N. (2016). Revisiting the four types of cyclists: Findings from a national survey. *Transportation Research Record*, *2587*(1), 90–99. https://doi.org/10.3141/2587-11

Ding, C., & He, X. (2004). *K* -means clustering via principal component analysis. *Twenty-First International Conference on Machine Learning - ICML '04*, 29. https://doi.org/10.1145/1015330.1015408

Ecola, L., Rohr, C., Zmud, J., Kuhnimhof, T., & Phleps, P. (2014). *The future of driving in developing countries*. RAND Corporation.

Emery, J., Crump, C., & Bors, P. (2003). Reliability and validity of two instruments designed to assess the walking and bicycling suitability of sidewalks and roads. *American Journal of Health Promotion*, *18*(1), 38–46. https://doi.org/10.4278/0890-1171-18.1.38

Epperson, B. (1994). Evaluating Suitability of Roadways for Bicycle Use: Toward a Cycling. *Transportation Research Record*, *1438*(1), 9–16.

Fairchild, A. (2021). After uptick in state cyclist deaths, local bicyclists plead for driver awareness. University Star. https://www.universitystar.com/news/after-uptick-in-state-cyclist-deaths-local-bicyclists-plead-for-driver-awareness/article_a576a6f8-2586-11ec-a8b2-bbb2d37f6e0e.html

FHWA. (2017). *2017 National Household Travel Survey (NHTS)*. US Department of Transportation. nhts.ornl.gov

FHWA. (2020). *Federal-Aid Highway Program Funding for Pedestrian and Bicycle Facilities and Programs—Funding—Bicycle and Pedestrian Program—Environment—FHWA*. https://www.fhwa.dot.gov/environment/bicycle_pedestrian/funding/bipedfund.cfm

Forbes, T. W. (1939). The normal automobile driver as a traffic problem. *The Journal of General Psychology*, *20*(2), 471–474. https://doi.org/10.1080/00221309.1939.9710022

Forester, J. (1993). *Effective Cycling*. MIT Press.

Fucoloro, T. (2011, July 22). Adaptive cycling and safe infrastructure can be tools to help overcome disability. *Seattle Bike Blog*. https://www.seattlebikeblog.com/2011/07/22/adaptive-cycling-and-safe-infrastructure-can-be-tools-to-help-overcome-disability/

Galatan, J. (2019, June 3). Op-Ed: Breaking down barriers to disabled cyclists. *Streetsblog New York City*. https://nyc.streetsblog.org/2019/06/03/op-ed-breaking-down-barriers-to-disabled-cyclists/

Gatersleben, B., & Haddad, H. (2010). Who is the typical bicyclist? *Transportation Research Part F: Traffic Psychology and Behaviour*, *13*(1), 41–48. https://doi.org/10.1016/j.trf.2009.10.003

Geels, F. W., Sovacool, B. K., Schwanen, T., & Sorrell, S. (2017). Sociotechnical transitions for deep decarbonization. *Science*, *357*(6357), 1242–1244. https://doi.org/10.1126/science.aao3760

Geller, R. (2009). *Four Types of Cyclists*. https://www.portlandoregon.gov/transportation/article/264746

Germano, A. T., Wright, P. H., Hicks, R. G., & Sanders, P. H. (1973). The emerging needs of bicycle transportation. *Highway Research Record*, 8–18.

Gerrard, A. J. W., & Robinson, D. A. (1971). Variability in slope measurements. A discussion of the effects of different recording intervals and micro-relief in slope Studies. *Transactions of the Institute of British Geographers*, *54*, 45–54. https://doi.org/10.2307/621361

Gillingham, K. (2014). Identifying the elasticity of driving: Evidence from a gasoline price shock in California. *Regional Science and Urban Economics*, *47*, 13–24. https://doi.org/10.1016/j.regsciurbeco.2013.08.004

Golledge. (1999). *Wayfinding Behavior: Cognitive Mapping and Other Spatial Processes*. JHU Press.

Golledge, R., & Garling, T. (2002). Spatial behavior in transportation modeling and planning. In K. Goulias (Ed.), *Transportation Systems Planning* (Vol. 20026459). CRC Press. https://doi.org/10.1201/9781420042283.ch3

Gorobets, A. (2016). Development of bicycle infrastructure for health and sustainability. *The Lancet*, *388*(10051), 1278. https://doi.org/10.1016/S0140-6736(16)31671-3

Griffin, G. P., & Jiao, J. (2015). Where does bicycling for health happen? Analysing volunteered geographic information through place and plexus. *Journal of Transport & Health*, *2*(2), 238–247. https://doi.org/10.1016/j.jth.2014.12.001

Griswold, J. B., Yu, M., Filingeri, V., Grembek, O., & Walker, J. L. (2018). A behavioral modeling approach to bicycle level of service. *Transportation Research Part A: Policy and Practice*, *116*, 166–177. https://doi.org/10.1016/j.tra.2018.06.006

Harkey, D. L., Reinfurt, D. W., & Knuiman, M. (1998). *Development of the bicycle compatibility index: A level of service concept, final report* (FHWA-RD-98-072; p. 94). University of North Carolina.

Heine, J. (2013, December 20). *Who are you calling Fast and Fearless?!* https://www.renehersecycles.com/who-are-you-calling-fast-and-fearless/

Holmes, C. E. (2021). Standing out and blending in: Contact-based research, ethics, and positionality. *PS: Political Science & Politics*, *54*(3), 443–447. https://doi.org/10.1017/S1049096520002024

Huff, H., & Liggett, R. (2014). *The Highway Capacity Manual's Method for Calculating Bicycle and Pedestrian Levels of Service: The Ultimate White Paper*. 62.

Hull, A., & O'Holleran, C. (2014). Bicycle infrastructure: Can good design encourage cycling? *Urban, Planning and Transport Research*, *2*(1), 369–406. https://doi.org/10.1080/21650020.2014.955210

Hunt, J. D., & Abraham, J. E. (2007). Influences on bicycle use. *Transportation*, *34*(4), 453–470. https://doi.org/10.1007/s11116-006-9109-1

Hurst, R. (2009). *Cyclist's Manifesto: The Case for Riding on Two Wheels Instead of Four*. Rowman & Littlefield.

Jensen, M. (1999). Passion and heart in transport—A sociological analysis on transport behaviour. *Transport Policy*, *6*(1), 19–33. https://doi.org/10.1016/S0967-070X(98)00029-8

Jensen, S. U. (2007). Pedestrian and bicyclist level of service on roadway segments. *Transportation Research Record*, *2031*(1), 43–51. https://doi.org/10.3141/2031-06

Jones, E. G., & Carlson, T. D. (2003). Development of bicycle compatibility index for rural roads in Nebraska. *Transportation Research Record*, *1828*(1), 124–132. https://doi.org/10.3141/1828-15

Jones, K. H. (1998). A comparison of algorithms used to compute hill slope as a property of the DEM. *Computers & Geosciences*, *24*(4), 315–323. https://doi.org/10.1016/S0098-3004(98)00032-6

Kaplan, J. A. (1975). *Characteristics of the regular adult bicycle user* (FWHA/R9-76/7 Final Rpt.). Article FWHA/R9-76/7 Final Rpt. https://trid.trb.org/view/63188

Katteler, H., & Roosen, J. (1989). *Substituting other transport modes for cars.* https://trid.trb.org/view/349691

Kay, J. H. (1998). *Asphalt Nation: How the Automobile Took Over America and How We Can Take It Back*. University of California Press.

Kazemzadeh, K., Laureshyn, A., Winslott Hiselius, L., & Ronchi, E. (2020). Expanding the scope of the bicycle level-of-service concept: A review of the literature. *Sustainability*, *12*(7), 2944. https://doi.org/10.3390/su12072944

Klobucar, M. S., & Fricker, J. D. (2007). Network evaluation tool to improve real and perceived bicycle safety. *Transportation Research Record*, *2031*(1), 25–33. https://doi.org/10.3141/2031-04

Ladd, B. (2008). *Autophobia: Love and Hate in the Automotive Age*. University of Chicago Press.

Landis, B. (1994). Bicycle interaction hazard score: A theoretical model. *Transportation Research Record*, *1438*, 3–8.

Landis, B. W., Vattikuti, V. R., & Brannick, M. T. (1997). Real-time human perceptions: Toward a bicycle level of service. *Transportation Research Record*, *1578*(1), 119–126. https://doi.org/10.3141/1578-15

Lawrence, B. M., & Oxley, J. A. (2019). You say one route, we observe four: Using naturalistic observation to understand route-choices in cyclists. *Safety Science*, *119*, 207–213. https://doi.org/10.1016/j.ssci.2019.01.004

Lefeve, B. A. (1954). Speed habits observed on a rural highway. *Highway Research Board Proceedings*, *33*. https://trid.trb.org/view/120665

Li, H., Harvey, J., Wu, R., & Jones, D. (2015). *Bicycle Ride Quality: The Effect of Pavement Treatment Texture*. 318–329. https://doi.org/10.1061/9780784479216.029

Lindsay, J. B., Newman, D. R., & Francioni, A. (2019). Scale-optimized surface roughness for topographic analysis. *Geosciences*, *9*(7), 322. https://doi.org/10.3390/geosciences9070322

Litman, T. (2013). *Evaluating Active Transport Benefits and Costs: Guide to Valuing Walking and Cycling Improvements and Encouragement Programs*. https://trid.trb.org/view/1262285

Lott, D. Y., Tardiff, T. J., & Lott, D. F. (1977). *Bicycle transportation for downtown work trips: A case study in Davis, California*. *629*, 30–37.

Lowe, M. D. (1990). Alternatives to the automobile: Transport for livable cities. *Ekistics*, *57*(344/345), 269–282.

Lowry, M. B., Callister, D., Gresham, M., & Moore, B. (2012). Assessment of communitywide bikeability with bicycle level of service. *Transportation Research Record*, *2314*(1), 41–48. https://doi.org/10.3141/2314-06

Majumdar, B., & Mitra, S. (2018). Development of level of service (LOS) criteria for evaluation of bicycle suitability: A case study of Kharagpur, India. *Journal of Urban Planning and Development*, *144*. https://doi.org/10.1061/(ASCE)UP.1943-5444.0000432

McCullough, R. (2015). *Old Wheelways: Traces of Bicycle History on the Land*. MIT Press.

McKenzie, B. (2014). *Modes Less Traveled—Bicycling and Walking to Work in the United States: 2008–2012* (ACS-25; p. 18). US Department of Commerce.

Moudon, A. V., Lee, C., Cheadle, A. D., Collier, C. W., Johnson, D., Schmid, T. L., & Weather, R. D. (2005). Cycling and the built environment, a US perspective. *Transportation Research Part D: Transport and Environment*, *10*(3), 245–261. https://doi.org/10.1016/j.trd.2005.04.001

Mruck, K., & Breuer, F. (2003). Subjectivity and reflexivity in qualitative research—A new FQS issue. *Historical Social Research / Historische Sozialforschung*, *28*(3 (105)), 189–212.

NHSTA. (2019). *2018 fatal motor vehicle crashes: Overview* (DOT HS 812 826; pp. 1–10). National Highway Traffic Safety Administration; US Department of Transportation.

NHTSA. (2021). *Early estimate of motor vehicle traffic fatalities in 2020.* (DOT HS 813 115). National Highway Traffic Safety Administration; US Department of Transportation. https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/813115

OECD. (2013). *Cycling, Health and Safety*. Organisation for Economic Co-operation and Development. https://www.oecd-ilibrary.org/transport/cycling-health-and-safety_9789282105955-en

Parkin, J. (2012). *Cycling and Sustainability*. Emerald Group Publishing.

Parkin, J., & Meyers, C. (2010). The effect of cycle lanes on the proximity between motor traffic and cycle traffic. *Accident Analysis & Prevention*, *42*(1), 159–165. https://doi.org/10.1016/j.aap.2009.07.018

Petritsch, T. A., Landis, B. W., Huang, H. F., McLeod, P. S., Lamb, D., Farah, W., & Guttenplan, M. (2007). Bicycle level of service for arterials. *Transportation Research Record*, *2031*(1), 34–42. https://doi.org/10.3141/2031-05

Plazier, P. A., Weitkamp, G., & van den Berg, A. E. (2017). "Cycling was never so easy!" An analysis of e-bike commuters' motives, travel behaviour and experiences using GPS-tracking and interviews. *Journal of Transport Geography*, *65*, 25–34. https://doi.org/10.1016/j.jtrangeo.2017.09.017

Potter, I. B. (1898). *Cycle Paths: A Practical Hand-book, Containing the Best Available Information to Guide Members of the League of American Wheelmen and Others in Placing in Substantial Form Their Protest Against Bad Roads by the Construction and Maintenance of Those Temporary Blessings Known as Cycle Paths*. League of American Wheelmen.

Raad, N., & Burke, M. I. (2018). What are the most important factors for pedestrian level-of-service estimation? A systematic review of the literature. *Transportation Research Record*, *2672*(35), 101–117. https://doi.org/10.1177/0361198118790623

Ramsey, K., & Bell, A. (2014). The Smart Location Database: A nationwide data resource characterizing the built environment and destination accessibility at the neighborhood scale. *Cityscape*, *16*(2), 145–162.

Rao, C. R. (1964). The use and interpretation of principal component analysis in applied research. *Sankhyā: The Indian Journal of Statistics, Series A (1961-2002)*, *26*(4), 329–358.

Richter, E. D., Berman, T., Friedman, L., & Ben-David, G. (2006). Speed, road injury, and public health. *Annual Review of Public Health*, *27*(1), 125–152. https://doi.org/10.1146/annurev.publhealth.27.021405.102225

Ridgway, M., Nielson, C., Snyder, D., & Foster, N. (2013). *Calibration of Highway Capacity Manual (2010) Bicycle Level of Service to Urban and Suburban Bikeways* [Transportation Research Board]. https://rns.trb.org/details/dproject.aspx?n=33971

Rietveld, P., & Daniel, V. (2004). Determinants of bicycle use: Do municipal policies matter? *Transportation Research Part A: Policy and Practice*, *38*(7), 531–550. https://doi.org/10.1016/j.tra.2004.05.003

Schipper, L., Saenger, C., & Sudardshan, A. (2011). Transport and carbon emissions in the United States: The long view. *Energies*, *4*(4), 563–581. https://doi.org/10.3390/en4040563

Sener, I. N., Eluru, N., & Bhat, C. R. (2009a). An analysis of bicycle route choice preferences in Texas, US. *Transportation*, *36*(5), 511–539. https://doi.org/10.1007/s11116-009-9201-4

Sener, I. N., Eluru, N., & Bhat, C. R. (2009b). Who are bicyclists? Why and how much are they bicycling? *Transportation Research Record*, *2134*(1), 63–72. https://doi.org/10.3141/2134-08

SFDPH. (2009). *Bicycle Environmental Quality Index (BEQI): Draft Report*. San Francisco Department of Public Health. https://merritt.cdlib.org/d/ark%253A%252F13030%252Fm5qz3r51/1/producer%252F892131162.pdf

SFDPH. (2010). *Bicycle Environmental Quality Index* (Program on Health, Equity, and Sustainability). San Francisco Department of Public Health. https://merritt.cdlib.org/d/ark:%2F13030%2Fm5vq4gtf/1/producer%2F892128603.pdf

Sorton, A., & Walsh, T. (1994). Bicycle stress level as a tool to evaluate urban and suburban bicycle compatibility. *Transportation Research Record*, *1438*(1), 17–24.

Stinson, M. A., & Bhat, C. R. (2004). Frequency of bicycle commuting: Internet-based survey analysis. *Transportation Research Record*, *1878*(1), 122–130. https://doi.org/10.3141/1878-15

Stopher, P. R. (1977). On the application of psychological measurement techniques to travel demand estimation. *Environment and Behavior*, *9*(1), 67–80. https://doi.org/10.1177/001391657791004

Stopher, P., & Stanley, J. (2014). *Introduction to Transport Policy: A Public Policy View*. Edward Elgar Publishing.

Surico, J. (2020, August 13). When street design leaves some people behind. *Bloomberg.com*. https://www.bloomberg.com/news/articles/2020-08-13/do-bike-lanes-have-an-accessibility-problem

Tefft, B. C. (2013). Impact speed and a pedestrian's risk of severe injury or death. *Accident Analysis & Prevention*, *50*, 871–878. https://doi.org/10.1016/j.aap.2012.07.022

Thigpen, C. G., Li, H., Handy, S. L., & Harvey, J. (2015). Modeling the impact of pavement roughness on bicycle ride quality. *Transportation Research Record*, *2520*(1), 67–77. https://doi.org/10.3141/2520-09

Thorin, E. (2017). Life [ageing] is like riding a bicycle. To keep your [coronary and heart] balance you must keep moving. *The Journal of Physiology*, *595*(12), 3701–3702. https://doi.org/10.1113/JP274297

Tomlinson, D. (2003). The Bicycle and Urban Sustainability. *FES Oustanding Graduate Student Paper Series*, *7*(6). https://yorkspace.library.yorku.ca/xmlui/handle/10315/18107

Transportation Research Record. (2010). *Highway Capacity Manual, 2010* (5th ed.).

Turner, S., Hottenstein, A., & Shunk, G. (1997a). *Bicycle and Pedestrian Travel Demand Forecasting: Literature Review* (FHWA/TX-98/1723-1; p. 54). Texas Transportation Institute.

Turner, S., Shafer, C., & Stewart, W. P. (1997b). *Bicycle suitability criteria: Literature review and state-of-the-practice survey* (TX-97/3988-1; p. 56). Texas Transportation Institute.

Veillette, M.-P., Grisé, E., & El-Geneidy, A. (2019). Does one bicycle facility type fit all? Evaluating the stated usage of different types of bicycle facilities among cyclists in Quebec City, Canada. *Transportation Research Record*, *2673*(6), 650–663. https://doi.org/10.1177/0361198119844741

Wang, K., & Akar, G. (2018). The perceptions of bicycling intersection safety by four types of bicyclists. *Transportation Research Part F: Traffic Psychology and Behaviour*, *59*, 67–80. https://doi.org/10.1016/j.trf.2018.08.014

Wang, X., Lindsey, G., Schoner, J. E., & Harrison, A. (2016). Modeling bike share station activity: Effects of nearby businesses and jobs on trips to and from stations. *Journal of Urban Planning and Development*, *142*(1), 04015001. https://doi.org/10.1061/(ASCE)UP.1943-5444.0000273

Warren, S. D., Hohmann, M. G., Auerswald, K., & Mitasova, H. (2004). An evaluation of methods to determine slope using digital elevation data. *CATENA*, *58*(3), 215–233. https://doi.org/10.1016/j.catena.2004.05.001

Wegman, F., Zhang, F., & Dijkstra, A. (2012). How to make more cycling good for road safety? *Accident Analysis & Prevention*, *44*(1), 19–29. https://doi.org/10.1016/j.aap.2010.11.010

White, M. J. (2004). The “Arms Race” on American roads: The effect of sport utility vehicles and pickup trucks on traffic safety. *The Journal of Law and Economics*, *47*(2), 333–355. https://doi.org/10.1086/422979

WHO. (2018). *Global status report on road safety 2018*. World Health Organization.

Wilson, S. S. (1973). Bicycle technology. *Scientific American*, *228*(3), 81–91.

Winters, M., Friesen, M. C., Koehoorn, M., & Teschke, K. (2007). Utilitarian bicycling: A multilevel analysis of climate and personal influences. *American Journal of Preventive Medicine*, *32*(1), 52–58. https://doi.org/10.1016/j.amepre.2006.08.027

Xu, D. (2019). Burn Calories, Not Fuel! The effects of bikeshare programs on obesity rates. *Transportation Research Part D: Transport and Environment*, *67*, 89–108. https://doi.org/10.1016/j.trd.2018.11.002

Zuniga-Garcia, N., Ross, H. W., & Machemehl, R. B. (2018). Multimodal level of service methodologies: Evaluation of the multimodal performance of arterial corridors. *Transportation Research Record*, *2672*(15), 142–154. https://doi.org/10.1177/0361198118776112