Title
Estimating Pedestrian Accident Exposure

Permalink
https://escholarship.org/uc/item/7m97h15t

Authors
Greene-Roesel, Ryan
Diogenes, Mara Chagas
Ragland, David D.

Publication Date
2010-05-01

CALIFORNIA PATH PROGRAM
INSTITUTE OF TRANSPORTATION STUDIES
UNIVERSITY OF CALIFORNIA, BERKELEY

Estimating Pedestrian Accident Exposure

Safe Transportation Education and Research Center (SafeTREC)
(SafeTREC was formerly known as the Traffic Safety Center)

California PATH Research Report
UCB-ITS-PRR-2010-32

This work was performed as part of the California PATH Program of the University of California, in cooperation with the State of California Business, Transportation, and Housing Agency, Department of Transportation, and the United States Department of Transportation, Federal Highway Administration.

Final Report for Task Orders 5211/6211

May 2010

ISSN 1055-1425

ESTIMATING PEDESTRIAN ACCIDENT EXPOSURE

Final Report

TO 5211/6211

SafeTREC Safe Transportation Research & Education Center

(SafeTREC was formerly known as the Traffic Safety Center)

CALIFORNIA PATH

for

California State of California Department of Transportation (Caltrans)
Division of Research & Innovation

Estimating Pedestrian Accident Exposure
TO's 5211 & 6211

Final Report

Safe Transportation Education and Research Center (SafeTREC)

(SafeTREC was formerly known as the Traffic Safety Center)

California Partners for Advanced Transit and Highways

for the
California Department of Transportation

This research was funded by the California Department of Transportation.

Executive Summary

We are pleased to present the final report of Caltrans Task Orders 5211 and 6211, “Estimating Pedestrian Accident Exposure.” The project focused on defining pedestrian exposure and evaluating methods for measuring it within the State of California. The project was funded by the California Department of Transportation as part of the California Partners for Advanced Transit and Highways (PATH) Program of the University of California.

Deliverables associated with the project include (I) a protocol report on assessing pedestrian exposure, which is accompanied by a training curriculum and an evaluation of manual pedestrian counting methods; (II) an evaluation and test of automated pedestrian counting methods; and (III) a report on strategies to create a statewide pedestrian exposure database and (IV) a protocol for Pedestrian Exposure Study in Alameda County. The deliverables are discussed in more detail below.

(I) Protocol report, training curriculum, and test of manual counting methods

The protocol report aims to assist transportation engineers and planners with the task of measuring pedestrian exposure for a variety of purposes and contexts. Purposes may include comparisons of the safety effects of pedestrian infrastructure; comparisons of pedestrian risk among different population groups; or comparisons of risk by mode of travel (e.g. walking versus bicycling). The geographic contexts may range from the entire state of California to a specific pedestrian crossing. Because each possible purpose and context will have a unique set of considerations and constraints, the protocol focuses on matching data collection methods with different study needs.

The protocol report guides the user through the tasks of determining an appropriate definition for pedestrian exposure; choosing the method of measurement that best suits the data collection purpose; devising a sampling strategy; and estimating annual pedestrian exposure from short samples of pedestrian volume. To accompany the report, we created a six-module training curriculum in powerpoint format. The course could be administered by Caltrans staff or local officials to educate engineers and planners about the task of measuring pedestrian exposure.

We also conducted two supporting research efforts to support development of the protocol. The first was a review of state-of-the-art pedestrian volume modeling methods used to estimate pedestrian exposure, including sketch plan, network analysis, and microsimulation models. The review was published in the 2006 Transportation Research Board Meeting CD-Rom as “Pedestrian Volume Modeling for Traffic Safety and Exposure Analysis: The Case of Boston, Massachusetts” and is attached to this report.

The second supporting research effort we conducted was a detailed field test of manual pedestrian counting methods. We compared the accuracy and effectiveness of counts obtained from field observers and from manual review of video recordings. The results of the test are attached as an appendix to the protocol

report, and was also published as “Pedestrian Counting Methods at Intersections: a Comparative Study” (Section 1: Appendix B) in the 2007 edition of the Transportation Research Record (Vol. 2002).

(II) Evaluation and test of automated pedestrian counting methods

Several automated pedestrian detection technologies have emerged in recent years, some of which can also be adapted for the purpose of pedestrian counting. These devices have the potential to allow pedestrian data collection over extended periods, and to reduce the labor costs associated with data collection.

We reviewed existing technologies using information from the literature, and identified five technologies that could be adapted for the purpose of counting pedestrians. We described each of these in our report on automated pedestrian counting methods (II). Based on the results of the review, we selected the passive infrared sensing technology as the most promising candidate for further study, because it is commercially available, not sensitive to lighting conditions, easy to install, and has been used successfully in outdoor environments in the United States. We conducted a test of this technology and included the results as an appendix to the report on automated pedestrian counting methods.

(III) Pedestrian exposure database: Approaches to a Statewide Pedestrian Exposure Database

Volume data is routinely collected for motorized modes but is not for non-motorized modes. Such data is essential for tracking pedestrian exposure and for infrastructure planning purposes. In this deliverable, we explore the possibility of creating a formalized, institutionalized mechanism for pedestrian data collection through a statewide pedestrian volume database. This database would meet a variety of data needs for different stakeholder groups. One of its principal purposes would be to allow safety professionals at the state and local levels to estimate pedestrian exposure to risk at specific sites.

In the report, we discuss the technical and institutional challenges inherent in creation of a pedestrian exposure database; possible sources of a pedestrian network inventory; and possible approaches to data collection. In addition, we recommend further steps for pursuing database development.

(IV) Alameda County Pedestrian and Bicycle Counting Protocol

This document describes the methods that will be used to collect pedestrian and bicycle counts at a sample of roadway intersections in Alameda County. There are two immediate purposes of this counting effort: a) obtain a sample of counts that can be used as a basis for predicting the number of pedestrians and bicyclists at all 531 intersections of Caltrans roadways in the county, and b) demonstrate that the data collection and modeling methods used in this pilot study have the potential to be applied to Caltrans roadways statewide. Ultimately, the predicted pedestrian and bicycle volumes can be used to represent exposure in a crash risk analysis. This will allow Caltrans and Alameda County to evaluate and prioritize pedestrian and bicycle

safety needs more accurately at each intersection. The methods used in this effort can be repeated by the County at regular intervals to track changes in pedestrian and bicycle activity over time.

During the research process, we identified several areas for further research. Two of these in particular stand out. First, we determined that the goal of a statewide pedestrian database could be furthered through research into a pilot database. This could be achieved either by collecting a sample of pedestrian volumes at locations in the state highway network, which could then be entered into the TASAS database, or by developing and sampling a GIS-based inventory of the pedestrian network in one of the Caltrans districts (e.g. District four). Second, we determined that the phenomenon of pedestrian "safety in numbers" has very important implications for the measurement of pedestrian risk and deserves immediate study. This phenomenon potentially undermines the usefulness of pedestrian collision rates as a proxy for pedestrian risk. Further research is needed to determine whether the safety in numbers phenomenon is a result of pedestrian or driver behavior; built environment factors; or other sources.

The "Alameda County Pedestrian and Bicycle Counting Project Summary" is a 5-page document outlining the final effort of this task order. It contains the section "Extrapolating Weekly Pedestrian Intersection Crossing Volumes from 2-Hour Manual Counts" and "A Pilot Model for Estimating Pedestrian Intersection Crossing Volumes," highlighting the research design, findings, and considerations.

Keywords: Pedestrians, Exposure, Intersections, Pedestrian Counts, Pedestrian Traffic, Pedestrian Accidents, Risk Analysis, Pedestrian Volume, Pedestrian Movement

FINAL REPORT TABLE OF CONTENTS

ESTIMATING PEDESTRIAN ACCIDENT EXPOSURE

Protocol Report

MARCH 2007


Photo by: Dan Burden

SafeTREC Safe Transportation Research & Education Center

(SafeTREC was formerly known as the Traffic Safety Center)

(SafeTREC was formerly known as the Traffic Safety Center)

The mission of the UC Berkeley Traffic Safety Center is to reduce traffic fatalities and injuries through multi-disciplinary collaboration in education, research, and outreach. Our aim is to strengthen the capability of state, county, and local governments, academic institutions, and local community organizations to enhance traffic safety through research, curriculum and material development, outreach, and training for professionals and students.

ESTIMATING PEDESTRIAN ACCIDENT EXPOSURE

Protocol Report

Prepared for Caltrans under
Task Order 6211

Prepared by
RYAN GREENE-ROESEL
MARA CHAGAS DIOGENES
DAVID R. RAGLAND

University of California Traffic Safety Center

University of California
Berkeley, CA 94720
Tel: 510/642-0655
Fax: 510/643-9922

ACKNOWLEDGEMENTS

The University of California Traffic Safety Center (TSC) appreciates and acknowledges the contributions of the following participants.

TSC staff:
Noah Radford
Marla Orenstein
Judy Geyer
Kara MacLeod
Tammy Wilder

PATH staff:
Ashkan Sharafsaleh, Research Engineer
Steve Shladover, Research Engineer
Fanping Bu, Research Engineer

Specialized consultants:
Charles Zegeer

Funding for this project was provided by Caltrans.

LIST OF CONTENTS

LIST OF TABLES AND FIGURES

1. PREFACE

1.1. Purpose of the Protocol

Walking is a healthful, environmentally benign form of travel, and is the most basic form of human mobility. Walking trips account for more than 8 percent of all trips taken in California, making walking the second most commonly used mode of travel after the personal automobile (Caltrans, 2002). In addition, many trips made by vehicle or public transit begin and end with walking.

In spite of the importance and benefits of walking, pedestrians suffer a disproportionate share of the harm of traffic incidents in California. As noted above, walking trips make up just 8 percent of all trips in the state, but 17 percent of all traffic fatalities are suffered by pedestrians. In 2004, 694 pedestrians were killed in the state of California and 13,892 were injured (California Highway Patrol, 2004).

To address this problem, significant resources are focused on countermeasures that aim to reduce the risk of pedestrian injury. Because resources are limited, risk analysis is necessary to develop cost-effective countermeasures (Høj and Kröger, 2002).

In the field of pedestrian safety, risk analysis involves assessing factors that contribute to the danger that a pedestrian is struck by a vehicle. These factors may include physical characteristics of the street, such as lack of sidewalks; behavioral issues, such as pedestrian or driver alcohol use; as well as other environmental variables. In order to fully understand how these factors contribute to risk, it is necessary to collect information on pedestrian exposure. Collection of pedestrian exposure information is an essential component of risk analysis.

Pedestrian exposure is a concept that refers to the amount that people are exposed to the risk of being involved in a traffic collision. In principle, pedestrians are exposed to this risk whenever they are walking in the vicinity of automobiles. There are many metrics that can be used to measure pedestrian exposure, but pedestrian volumes are the most frequently used.

Although many state, regional, and local agencies have developed methodologies to collect pedestrian volume data, there is no consensus on which method is best (Schneider et al., 2005; Schweizer, 2005). This is because there is no "one size fits all" method of counting pedestrians. Rather, the choice of strategy depends on a complex range of factors, including the characteristics of the area being studied; the resources available for data collection; and the specific purpose of data collection.

This protocol aims to improve pedestrian data collection in the state of California by providing information and guidance for each decision point in the data collection process. Each chapter represents one of these decision points, and each will guide the user through important considerations relevant to the data collection stage. In addition, each chapter provides a combination of real-world and hypothetical example scenarios to illustrate the issues discussed in the text.

The first chapter, "Pedestrian Exposure," discusses the issue of how to select a definition of pedestrian exposure that is appropriate to the study purposes, resources, and chosen counting method. It also discusses the meaning of pedestrian exposure and its importance in pedestrian risk analysis.

The second chapter, "Area-Wide Methods," describes three general approaches to measuring pedestrian exposure for defined geographic areas, such as cities or counties. This chapter assists users in understanding the strengths and weakness of different methods of measuring pedestrian exposure over wide areas, and introduces users to existing sources of data on pedestrian activity.

The third chapter, "Site-Specific Methods," focuses on commonly used methods for counting pedestrian activity directly at specific sites, such as intersections or crossings. The performance of these methods is evaluated in terms of their relative cost, convenience, accuracy, and ability to collect a range of data points.

The fourth chapter, "Data Collection Planning at Intersections," assists users with the task of planning data collection at specific sites. It describes the statistical issues that must be addressed when designing a pedestrian data collection strategy, such as how to choose which sites to study and how to determine the number of sites to be studied.

The fifth chapter, “Estimating Annual Pedestrian Volumes,” describes a method for converting short pedestrian counts into an annual measure of pedestrian volume using statistical analysis of pedestrian flow patterns. This method can be used to reduce the time and cost associated with developing an annual measure of pedestrian exposure, which is necessary to determine the annual pedestrian risk at a site.

Taken together, these chapters will assist the user in measuring pedestrian exposure for a variety of purposes and contexts. The purposes may include comparisons of the safety effects of pedestrian infrastructure; comparisons of pedestrian risk among different population groups; or comparisons of risk by mode of travel (e.g. walking versus bicycling). The geographic contexts may range from the entire state of California to a specific pedestrian crossing. Because each possible purpose and context will have a unique set of considerations and constraints, this protocol focuses on matching data collection methods with different study needs.

1.2. Who Should Use this Protocol

This protocol is intended to be used by traffic engineers and planners, consultants, and researchers interested in measuring pedestrian exposure. Although unaffiliated users will benefit from reading the protocol, it is most appropriate for those who are associated with an institution that has the resources necessary to mount a data collection program.

1.3. How to Use this Protocol

As discussed above, each chapter is aimed at a particular aspect of the data collection process. Some users may wish to read only the section that is most relevant to their needs. However, because the issues in the chapters are closely inter-related, many users will benefit from reading the entire document.

Users should understand that this protocol is not a “how-to” guide for measuring pedestrian exposure. Although many specific methods and equations are provided, the intention is to educate the user about the data collection process rather than to provide a set of instructions. This is because, as mentioned above, measuring pedestrian exposure is a complex task that is constrained by the study resources, purposes, and context. This protocol aims to inform the user about the data collection strategies available to them, and to assist them in choosing which one best meets their needs.

2. PEDESTRIAN EXPOSURE

Before seeking to measure pedestrian exposure, it is important to have a clear understanding of the concept and its relationship to pedestrian risk. This chapter discusses the meaning of exposure in the context of risk analysis for pedestrian safety, and presents several common measures of pedestrian exposure used in the transportation safety field.

As this guide will demonstrate, there is no single best measure of pedestrian exposure, but some measures are better adapted to specific needs and purposes, such as comparing infrastructure; comparing risk among populations; or evaluating the change in pedestrian risk over time. This chapter will assist users in selecting an appropriate measure of exposure to match their needs.

2.1. Understanding Exposure and Risk

In epidemiology, exposure refers to a person's contact with a potentially hazardous situation or substance. For example, each time you fly in an airplane, you are exposed to ionizing radiation. Each time you cross a street, you are exposed to the possibility of being injured by a vehicle. Exposure can also be understood as a "trial event" in during which a harmful outcome might occur.

Risk is an abstract concept that refers to the probability a harmful event will occur given a certain number of trials. In pedestrian safety, each "trial" is a unit of exposure such as a minute spent walking or a road crossing Table 2.1 describes the relationship between exposure and risk.

Table 2.1: Exposure versus Risk

Exposure	Contact or amount of contact with potentially harmful situation (x)	(x)
Risk	Probability of collision/injury/fatality (c) per unit of exposure.	$P(c \mid x)$

The likelihood that any given trial event will result in a particular outcome is a function of the “chance set up”. In transport safety, the “chance set up” is the transportation system itself, including its physical characteristics, users, and environment. Any one of these characteristics might influence the likelihood that a given trial event – such as a pedestrian crossing – will result in a collision (Hauer, 1982).

Risk and exposure are theoretical concepts that can only be indirectly estimated through the use of proxy measures. In the field of traffic safety, risk is typically represented by a simple ratio between collisions, injuries or fatalities, and exposure for a specific geography and time period (Chu, 2004). This ratio is referred to as the “collision rate” or the “accident rate”. See Section 2.6 for a discussion of the limitations of collision rates as a proxy for risk.

$$\text{Collision rate} = \frac{\text{Number of collisions in a specified time and place}}{\text{Amount of exposure in a specified time and place}} \qquad (1)$$

If one finds that risk is higher at one intersection than another, it suggests that something in the “chance set up” (e.g. higher traffic speeds at one intersection) explains the difference. In this way, risk analysis is used to identify dangerous aspects of the transportation environment.

A short list of some of the factors thought to be associated with pedestrian risk include:

✓ Pedestrian characteristics including age and gender (Evans, 1991; Keall, 1995), and socioeconomic status and ethnicity (Ogden, 1997; Kraus et al., 1996). These characteristics may be related to distance and time traveled; pedestrian behavior; and awareness of the road environment.

✓ Pedestrian behavioral characteristics, such as risk-taking behavior, propensity to jaywalk, etc (Campbell et al., 2004).

✓ Trip characteristics: time of day/year, purpose, time elapsed between drinking alcohol and commencement of trip (Keall, 1995).

- ✓ Area characteristics related to transportation service and land use (Herms, 1970; Ossenbruggen, 1999).
- ✓ Roadway features such crosswalks and alternative crossing treatments, signalization, signing, pedestrian refuge islands, provisions for pedestrians with disabilities, bus stop location, and school crossing measures (Campbell et al., 2004).

2.2. Incorporating Exposure into Risk Measurement

Exposure is a crucial component of risk measurement. If the absolute number of injuries or fatalities is presented without controlling for exposure, it is easy to come to erroneous conclusions about risk.

The following graphs are provided to illustrate the importance of incorporating pedestrian exposure into measurement of risk. Figure 2.1 shows the number of pedestrians killed in New Zealand between 1988-1991, ordered by age and gender. These “raw” counts make it seem that children under twenty are most in danger of being killed.

However, when the raw counts are presented as a function of exposure, measured as the hours spent walking, a very different picture emerges (Figure 2.2). The age categories with the highest risk are those aged 80 and above and those ten and younger. Adolescents aged 15-20 do not have elevated risk levels; rather, the high numbers of fatalities in this category are due the fact that adolescents spend more time walking than other age groups.

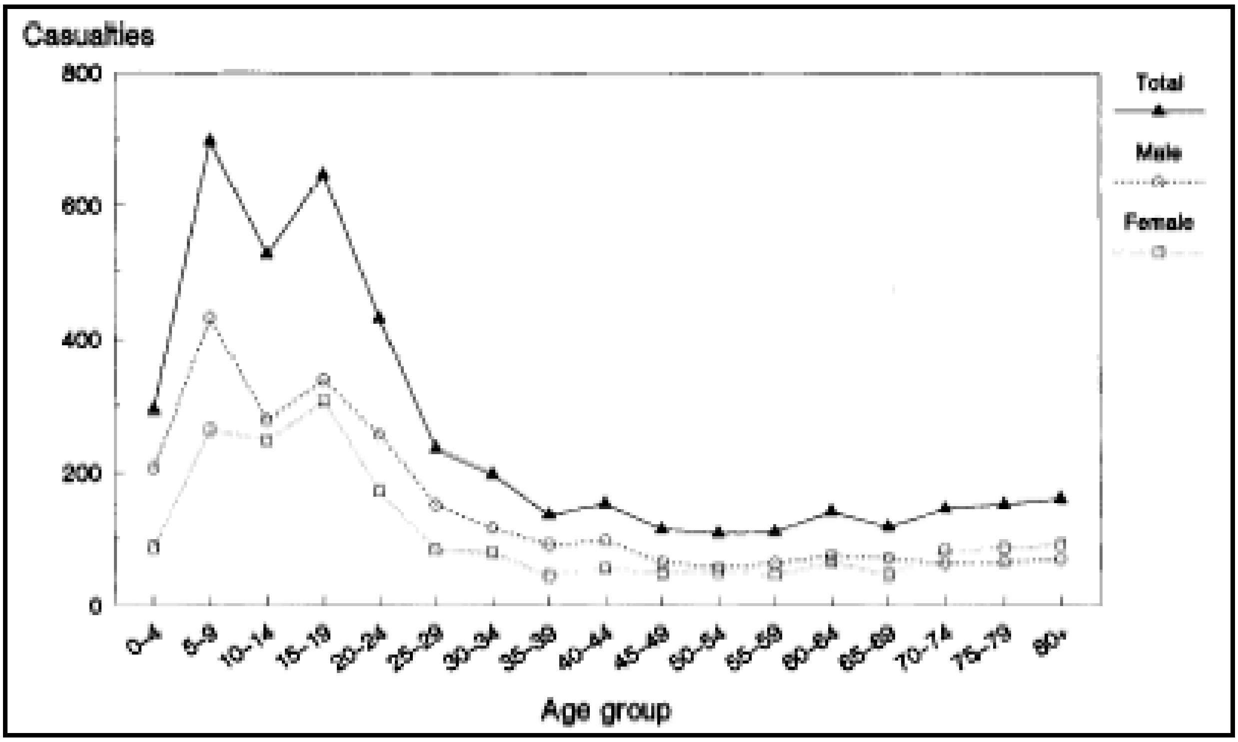


Figure 2.1: Number of pedestrians injured or killed in New Zealand, 1988-91 (Keall, 1995)

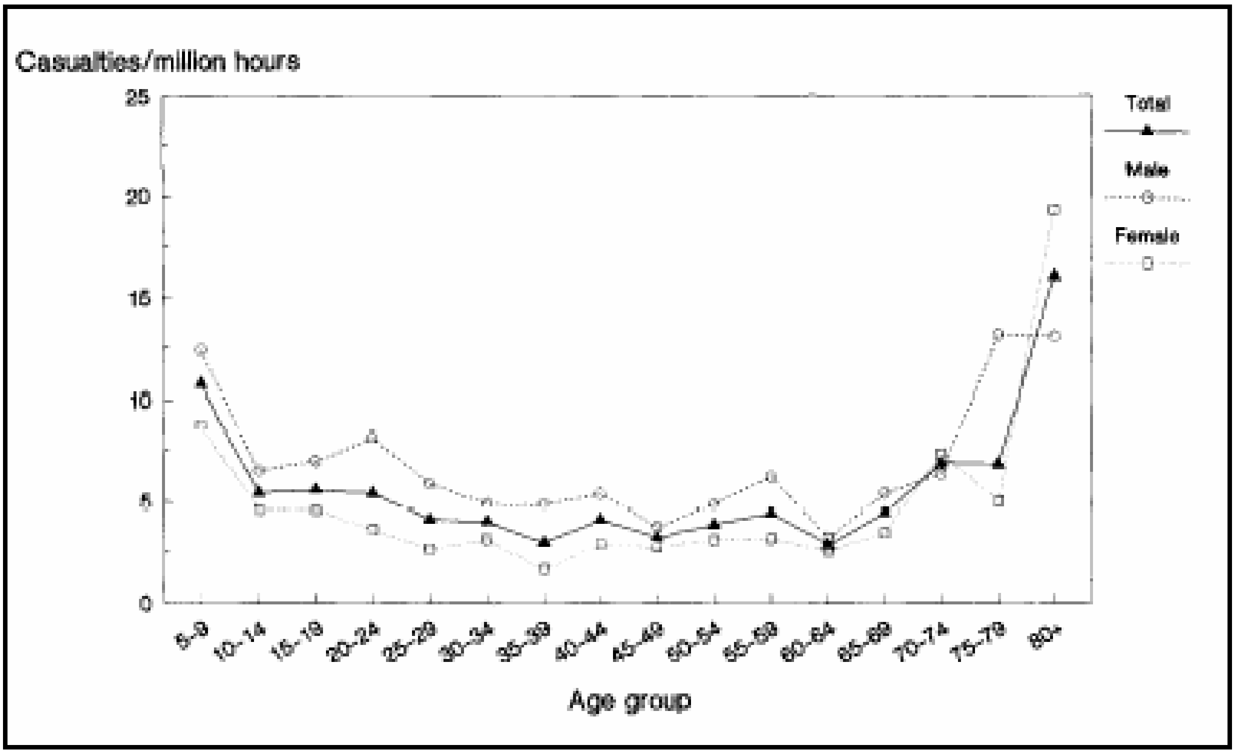


Figure 2.2: Number of pedestrian casualties per million hours walked in New Zealand 1988-91 (Keall, 1995)

When constructing a pedestrian safety risk measure, it is important to keep the following points in mind:

- ✓ The numerator and denominator in a risk measure must be consistent (Hauer, 2001); if exposure is in person-hours of pedestrian travel then the event in the numerator should be the number of pedestrians that experienced a collision or injury.

- ✓ The risk measure should reflect the type of risk being studied (Hakkert and Braimaister, 2002), such as whether the risk being studied is for an individual, or for a defined social group (Jorgensen, 1996).

- ✓ The denominator of the risk measure (pedestrian exposure) must reflect the intended purpose of the risk measure (Hakamies-Blomqvist, 1998). For example, a risk measure used to compare risk between different modes of travel should have a denominator (exposure measure) that is comparable across all modes.

- ✓ The denominator of the risk measure should reflect the target population being studied.

2.3. Defining Pedestrian Exposure

Pedestrian exposure is an abstract concept that reflects the opportunity for a potentially harmful pedestrian-vehicle interaction to occur; in other words, it is the number of trial events that could result in an injury or collision. It is very difficult to measure directly, since this would involve tracking the movements of all people at all times.

Instead, pedestrian exposure must be approximated using an appropriate proxy measure. Examples of measures used to represent pedestrian exposure at the micro level include pedestrian volume (Davis et al., 1988); the product of pedestrian and vehicle volumes at an intersection (Cameron, 1982) or roadway segment (Knoblauch et al., 1984); and the square root of that product (TRL, 2001). Measures used to represent exposure at the macro level in the U.S. include pedestrian distance traveled and pedestrian trips made (Pucher and Dijkstra, 2000, 2003); and the number of streets crossed (Roberts et al., 1996). In Europe, the most common measures include the number of pedestrian trips made; time spent walking; and distance walked (ETSC, 1999).

In situations where travel-based measures of exposure are unavailable, population-based measures are sometimes used to approximate exposure (NHTSA, 2004). These may include population density (Qin and Ivan, 2001), and population divided by the percent of workers who reported that they usually walked to work in the last week (STPP, 2002, 2004).

The choice of exposure measure strongly impacts the resulting calculation of risk. For example, researchers at the Surface Transportation Policy Project used “miles traveled” as the denominator in estimating risk to pedestrians across the nation in the 2004 *Mean Streets* report. They concluded that walking is about twenty times more dangerous than riding in passenger cars, trucks, or on public transit (STPP, 2002, 2004). This conclusion can be distorted by the fact that walking is much slower per mile than other forms of transportation. If the researchers had used as the measure of exposure the amount of time spent traveling, rather than miles traveled, they may have reached different conclusions.

To illustrate further, Table 2.2 presents pedestrian collision rates in the European Union calculated using two different exposure measures: person-kilometers traveled and person-hours of travel. When person-kilometers walked is the measure of exposure, pedestrian travel appears to be many times riskier than travel by car. When person-hours spent walking is the exposure measure, then pedestrian travel appears to have the same risk as vehicle travel.

Table 2.2: Fatality Risks over Distance and Time for Travel Modes in the EU

Travel mode		108 person km	108 person hours
Road	Total	1.1	33
	Bus/Coach	0.08	2
	Car	0.8	30
	Foot	7.5	30
	Cycle	6.3	90
	M/C,MOPED	16.0	500
Trains		0.04	2
Ferries		0.33	10.5
Planes		0.08	36.5

Source: ETSC, 1999

2.4. Measures of Pedestrian Exposure

Presented in Table 2.3 is an exploration of some of the common ways that pedestrian exposure is measured. For each of these exposure measures, an explanation and examples are provided; common and appropriate uses are discussed; and benefits and limitations are explored. Not all possible ways of estimating pedestrian exposure are described.

Table 2.3: Common Metrics Used to Describe Pedestrian Exposure

	Explanation
Population	Number of residents of a given area, or number of people in a demographic group.
Number of pedestrians	Number of pedestrians observed in a given area during a fixed interval.
Trips	Number of distinct trips taken by an individual pedestrian.
Distance traveled	Total distance traveled by an individual pedestrian *or* aggregate distance traveled by all pedestrians in a fixed area.
Time spent traveling	Total time traveled by an individual pedestrian *or* aggregate time traveled by all pedestrians in a fixed area.

These examples will illustrate that there is no single best definition of pedestrian exposure. However, it is important to choose the definition of exposure that best matches the needs and purposes of the study. The chosen exposure measure should be compatible with the measurement devices being used and the target population being studied within a geographic area. The choice of exposure measure will also be determined in part by the amount of available resources, as some measures of exposure are more costly to collect than others.

2.4.1. Exposure based on population data

Population refers to the number of people who live in a given area, or the number of people who make up a particular demographic group. Because it is relatively easy and cheap to estimate, population data is often used as a simple proxy for pedestrian exposure.

There are a large number of issues that make the use of population highly unreliable as an exposure estimate. First of all, actual physical exposure to traffic is unlikely to be evenly distributed throughout the population. Second, time spent as pedestrians, or distance traveled, are not represented or accounted for in any way. Third, population does not necessarily relate directly to the actual number of people walking on the streets.

For example, some tourist sites attract a large number of people who are not accounted for by residential or employment population density, but who may still be involved in traffic collisions (Ivan et al., 2000). Models of pedestrian risk based on population provide only the roughest approximation, and are probably unreliable. Table 2.4 summarizes the issues related to exposure measures based on population.

Table 2.4: Exposure Based on Population Data

APPROPRIATE USES	✓ Used as an alternative to exposure data when cost constraints make collecting exposure data impractical ✓ Used to compare jurisdictions over time because population data is available for many geographies and time periods
HOW DATA IS GATHERED	✓ Population data for most cities is available on an annual basis through the American Community Survey (ACS). The ACS is administered by the U.S. Bureau of the Census and is accessible online (U. S. Census Bureau, 2006)
PROS	✓ Easy and low-cost to obtain; available for most geographies and time periods ✓ Adjusts for differences in the underlying resident population of an area – for example, sparsely populated suburbs versus densely populated inner-city areas ✓ Provides a crude adjustment for amount of vehicle traffic on the streets, since areas where more people live also tend to be areas where more people drive ✓ May be the only way to represent exposure if direct measurements cannot be taken
CONS	✓ Does not accurately represent pedestrian exposure ✓ Does not account for the number of people who travel as pedestrians in the area ✓ Does not provide information about amount of time or distance that members of the population were exposed to traffic
COMMON MEASURES	✓ Number of people in a given area: neighborhood, city, county, state or country ✓ Number of people in a particular demographic group: by age, sex, race, immigrant status or socioeconomic status
EXAMPLES	✓ In 2001, pedestrian collisions killed 20 people per million in California, but only 7 people per million in Nebraska. (FARS and U.S. Census data from 2001). ✓ In 2004, the male pedestrian fatality rate per 100,000 population in United States was 2.22, while the female pedestrian fatality rate was 0.95 per 100,000 population (NHTSA, 2004).

2.4.2. *Exposure based on pedestrian volumes*

Pedestrian exposure can be measured by the number of pedestrians that pass through a fixed point during a specified time interval. This is a common exposure metric, as it is relatively simple to assess through established manual and automated counting methods. This exposure measure is explained in more detail on Table 2.5.

Table 2.5: Exposure Based on Pedestrian Volume

APPROPRIATE USES	✓ Estimating pedestrian volume and risk in a specific location. ✓ Assessing changes in pedestrian volume or characteristics due to countermeasure implementation at that site.
HOW DATA IS GATHERED	✓ Manual or automated counts of pedestrians.
PROS	✓ Counts are simpler to collect than other measures such as time or distance walked. ✓ Automated methods for counting number of pedestrians are improving.
CONS	✓ Does not differentiate pedestrians by walking speed, age, or other factors that may influence individual risk. ✓ Does not account for the amount of time spent walking or the distance walked ✓ Not easily adapted to assess exposure over wide areas (for example, a city).
COMMON MEASURES	✓ Average number of pedestrians per day, sometimes called Average Annual Number of Pedestrians (Zeeger et al., 2005; Cameron , 1976, Hocherman et al., 1988) ✓ Number of pedestrians per time period, e.g., hour (Davis et al., 1988; Cove and Clark, 1993)
EXAMPLES	✓ The average daily pedestrian traffic at marked crossings was 312 pedestrians per site (Zeeger et al., 2005). ✓ Between 7:00 am and 10:00 am, 203 pedestrians crossed Rose Street at the intersection of Shattuck Avenue.

While the “number of pedestrians” is the term most frequently used to refer to this exposure variable, that terminology is not, strictly speaking, accurate. A more precise term is ‘number of pedestrian crossings’, since a single pedestrian can contribute to the count more than once if that person passes through the measurement point more than one time during the observation period (such as during an outbound journey, and then again on the return). In addition, it is important to distinguish whether the crossing is over a roadway or over an arbitrary line on a sidewalk. Statistics suggest that crossing the street might be more dangerous than walking along the road, so that crossing exposure should be distinguished from roadside or sidewalk exposure (Evans, 1991; Ossenbruggen, 1999).

Key to the accurate measurement of the number of pedestrians is a good operative definition of what constitutes an entry into the area, and what constitutes a pedestrian. For example, should a mother pushing an infant in a stroller be counted as one pedestrian, or two?

Any fixed point can be used. However, in practice, intersection crossings are often used as the fixed point. The reason for this is that crossing the street is an activity with a relatively high risk. In a study of pedestrian crash types across several states, Hunter et al. (1996) found that about a third of crashes involving a pedestrian occur at intersections, whereas only about 8 percent of all crashes occurred while the pedestrian was walking along the roadway.

A major assumption made in using an intersection as a fixed point is that each crossing represents a fixed unit of risk, independent of crossing distance or location within the crossing.

2.4.3. Exposure based on trips

Exposure based on number of trips estimates the number of walking trips taken by an individual, regardless of the distance or time the journey takes. Trips may be taken for the purpose of commuting to work or school, for social visiting, for utilitarian purposes such as shopping, for walking a dog, or walking purely for recreation. This information is generally gathered by surveying a representative subset of a population. Because other survey questions are usually asked at the same time, each trip can be linked to information regarding trip purpose, time of day, etc.

Number of trips as assessed by survey is usually difficult to relate to pedestrian collision data on a small-area scale. However, the data is useful to assess exposure over wide areas, especially when combined with other datasets, such as U.S. Census information or land use data, enabling additional analyses of factors affecting walking patterns.

Number of trips may not be the most useful metric for risk analysis purposes, but it is commonly used for assessing pedestrian behavior and activity, for making comparisons between large jurisdictions, and for examining changes over time (Table 2.6).

Table 2.6: Exposure Based on Trips

APPROPRIATE USES	✓ Assessing pedestrian behavior in large areas, such as cities, states, or countries. ✓ Examining changes in pedestrian behavior over time. ✓ Making comparisons between jurisdictions. ✓ Assessing common characteristics of walking trips, such as purpose, route, etc.
HOW DATA IS GATHERED	✓ Data is gathered through use of surveys, such as the National Household Travel Survey (2001)
PROS	✓ Appropriate for use in large areas. ✓ Best metric to assess relationship of walking with trip purpose ✓ Trips can be assessed as a function of person, household and location attributes.
CONS	✓ As with most surveys, a large number of respondents are needed to adequately represent the underlying population. ✓ Unlikely to provide information at the level of detail needed to assess risk at specific locations ✓ Pedestrian trips are often underreported in surveys (Schwartz and Porter, 2000)
COMMON MEASURES	✓ Average number of walking trips made by members of a population per day, week or year. ✓ Proportion of walking trips taken for particular purposes, such as commuting or shopping.
EXAMPLES	✓ In US, the percentage of all work trips made by walking fell from 10.3% in 1960 to only 2.9% in 2000 (Pucher and Dijkstra, 2003). ✓ While in the Mid-Atlantic States 15.8% of all trips are made by the walking mode, in the East South Central and West South Central states this percentage is around 6% (Pucher and Renne, 2003). ✓ In US, 38% of all pedestrian trips are made for social and recreational purposes and 32% for going to school and church, while 10% represent work trips (Pucher and Renne, 2003).

2.4.4. Exposure based on distance

Exposure based on distance, or distance traveled, represents the distance that pedestrians walk while exposed to vehicular traffic. This exposure measurement can be assessed on the level of the individual or on the level of the geographic area. On the individual level, exposure based on distance is expressed as the total or average distance that an individual pedestrian travels in a fixed time period, such as a day, week, or year. Typically the risk is stated in terms of the number of deaths per 100 million person miles traveled (Chu, 2003). As with the measurement of number of trips, assessment of this exposure measure is carried out through surveys of a representative sample of the population. It is also possible to attach walking measurement devices, such as pedometers, to a sample of pedestrians.

On the geographic level, distance traveled is measured directly by aggregating the pedestrian distance traveled within a defined area during a fixed time period. This version of distance traveled is defined as the number of pedestrians counted, multiplied by the distance across the intersection. In this instance, the focus is on the total pedestrian-miles traveled, not the number of unique individuals traveling, and each individual may contribute distance more than once, if they pass through the observation area more than one time.

Using exposure based on distance to estimate risk, through either of the methods presented above, relies on the assumption that risk is a function of distance traveled. That means that other things being equal, crossing a roadway with four lanes carries twice the risk of crossing a roadway with two lanes.

The metric does not differentiate in terms of walking speed or other factors that could moderate the risk associated with distance. This potentially distorts the risk associated with walking when compared to other modes. One person-mile of walking represents far more exposure to vehicle traffic than one person-mile of riding in a passenger vehicle because of the differences in travel speeds between the modes (Chu, 2003). Thus, using a distance-based measure of exposure when comparing risk between modes may distort the results of the comparison. Table 2.7 presents more details about exposure measure based on distance.

2.4.5. Exposure based on time

Time exposure data has long been used for measuring risk (Jonah and Engel 1983; Anderson et al., 1989; ETSC, 1999). It has also been used to compare risk in different social groups or between travel modes. Keall (1995) estimated the risks of traffic collision for different sex and age groups by combining road collision data with survey data using the exposure measures “time spent walking” and “number of roads crossed”. Chu (2003) proposed a time-based comparative approach to examining the fatality risk of walking and vehicle travel because time-based measures take into account the speed differences between walking and riding in a passenger vehicle.

Exposure based on time incorporates not only the distance traveled, but also adjusts for walking speed. Like distance traveled, time traveled can be measured on the individual level through surveys or through direct measurement at specific locations.

Time spent walking at a crossing, for example, might be measured by multiplying the number of pedestrians by the average crossing time. It can also be measured by adding the crossing times of each individual. In comparing two individuals, all other characteristics being equal, the measure will account for different walking speeds.

To better characterize the exposure measure based on time, Table 2.8 presents its appropriate uses and examples.

Table 2.7: Exposure Based on Distance

APPROPRIATE USES	✓ Estimating exposure at the micro or macro level. ✓ Estimating whether risk increases in a linear manner with distance traveled. ✓ Assessing how crossing distance affects risk
HOW DATA IS GATHERED	✓ For individual level exposure, through surveys such as the National Household Travel Survey (2001) ✓ For aggregate level exposure, measurement of the length of the area of interest, combined with a manual or automatic count of the number of pedestrians.
PROS	✓ Can be used to measure exposure at the micro and macro levels ✓ More detailed than pedestrian volumes or population data ✓ Can be used to compare risk between different travel modes ✓ Common measure of vehicle exposure
CONS	✓ Does not take into account the speed of travel and thus cannot be reliably used to compare risk between different modes (e.g. walking and driving) ✓ Assumes risk is equal over the distance walked ✓ Must typically assume that each pedestrian walks the same distance in a crossing or along a sidewalk
COMMON MEASURES	✓ Average miles walked, per person, per day. ✓ Total aggregate distance of pedestrian travel across an intersection.
EXAMPLES	✓ The 2001 fatality rate per 100 million miles traveled in the U.S. was 1.3 for drivers and their passengers and 20.1 for pedestrians (STPP, 2004). ✓ Between 1990 and 2000, the share of Americans walking to work fell from 3.9% to 2.9% (U. S. Census 2000 Summary File 3, Census 1990 Summary Tape File 3.)

Table 2.8: Exposure Based on Time

APPROPRIATE USES	✓ Estimating total pedestrian time exposure for specific locations. ✓ Comparing risks between different modes of travel (e.g. walking vs. riding in a car). ✓ Estimating whether risk increases in a linear manner with walking time. ✓ Comparing risk between intersections with different crossing distances and between individuals with different walking speeds.
HOW DATA IS GATHERED	✓ The number of persons passing through an area multiplied by the time traveled. ✓ Time spent on walking activities reported on surveys.
PROS	✓ Accounts for different walking speeds ✓ Allows for accurate comparison between different modes of travel. ✓ Can be used to measure exposure at the micro and macro levels ✓ More detailed than pedestrian volumes or population data
CONS	✓ Time based measures assume risk is equal over the entire distance of a crossing. Only a small portion of time spent walking on roadways represents real exposure to vehicle traffic. This portion would include time spent crossing roads, walking on the road surface, or possibly walking along the roadside where there are no curved sidewalks (Chu, 2003). ✓ Time spent on walking can be over estimated in surveys, because people perceive that they spend more time walking than they actually do (Chu, 2003). ✓ Walking may also be under-reported in surveys, because people may forget walk trips or may purposely choosing not to report. Both of these reasons are related to the fact that walking trips are relatively short. These very short trips may not register in the memory of respondents or the respondents may think that these short trips are unimportant (Chu, 2003)
COMMON MEASURES	✓ Average time walked, per person, per day or year. ✓ Total aggregate travel time of pedestrian travel across an intersection.
EXAMPLES	✓ In 2001, the U.S. annual per capita minutes traveled was 2,139 minutes (Chu, 2003).

2.5. Choosing an Appropriate Exposure Measure

Exposure can be estimated in a number of different ways for almost any situation, as summarized in Table 2.3. These different ways of assessing exposure lead to different risk estimates, each of which may be correct but each may convey a different meaning. When determining the best exposure measure for a given purpose, key considerations include:

- ✓ **What is the chosen method of measuring exposure? Does it match the study purpose?** Surveys will yield individual-level measures of exposure such as person-trips or person-distance walked, while direct observation will yield geographic-level measures of exposure such as number of crossings or distance walked within a defined area.

- ✓ **Where is the exposure to be measured?** If exposure is measured at a facility such as a pedestrian crossing or along a sidewalk, then the exposure measure should be a micro-level measure, such as number of crossings.
- ✓ **What are the study resources?** Some exposure measures, such as time and distance, more accurately portray pedestrian risk than pedestrian counts alone. However, time or distance spent as a pedestrian will likely be more costly to collect than simpler measures of exposure.

The following section lists examples of study purposes and provides guidance on the choice of exposure measure for each.

2.5.1. Comparing safety infrastructure and countermeasures

When comparing the effects of infrastructure and/or countermeasure on pedestrian risk, the ideal measure of exposure will be collected directly in the area where the infrastructure and/or countermeasure are in place. This will allow an objective connection to be established between the site and pedestrian risk, and will allow a consistent numerator and denominator in the pedestrian risk measure. That is, the numerator will reflect the number of pedestrian-vehicle incidents occurring at the specific site and the denominator will reflect the number of “trials” occurring in the vicinity of the countermeasure. It should be noted however that surveys can in theory be used to track pedestrian use of infrastructure, although they are not well-adapted for this purpose. For example, the New Zealand Travel Survey of 1988-89 asked respondents to keep a diary recording the number of crossings made at ‘zebra-style’ pedestrian crossings (Keall, 1995).

The exposure measure should also be appropriate to the type of infrastructure being studied. If the effect of enhanced crossing devices is being studied, than the pedestrian crossing is an appropriate measure of exposure. Zeeger et al. (2005), for example, used the number of pedestrian crossings as the unit of exposure in a study comparing risk at marked and unmarked crossings. If the effect of new sidewalks along the length of a block are being studied, then pedestrian distance walked along the block would be a better measure of exposure.

2.5.2. Compare risk between groups of pedestrians

If the purpose of the study is to compare risk among different groups of pedestrians, the measure of exposure should be linked to individual-level attributes such as age; racial or ethnic group; income category; and so on. For example, Keall (1995) estimated the risks of collision for different sex and age groups by combining road collision data with survey data using the exposure measures “time spent walking” and “number of roads crossed”. These attributes are most easily collected through surveys, although it is possible to estimate certain pedestrian characteristics such as age and gender through direct observation.

2.5.3. Compare risk among different modes of travel

When comparing risk among different modes of travel, the best exposure measure reflects the different travel speeds of the modes being compared. For that reason, it is best to use time spent traveling to compare risk among different travel modes.

Because different modes use different infrastructure, it may be difficult to record and compare geographic-level measures of time spent traveling by various modes such as automobiles, airplanes, bicycles, and pedestrians. Recording the individual-level use of these modes by survey is more commonly used to compare risk.

2.6. Collision Rates as a Proxy for Risk

Although an in-depth discussion of risk measurement is outside the scope of this paper, it is important to be aware of possible pitfalls associated with using exposure data in simplistic risk analysis.

As noted above, exposure data is commonly used to calculate collision rates, namely the number of collisions in a given time and place divided by an exposure measure. The calculation of collision rates rests on the assumption that the number of collisions is proportional to exposure. In other words, it assumes that, all other things being equal, a place with more pedestrians should have more pedestrian-vehicle collisions, and that the number of collisions should increase at a constant rate as the number of pedestrians increases. Figure 2.3 illustrates this assumption.

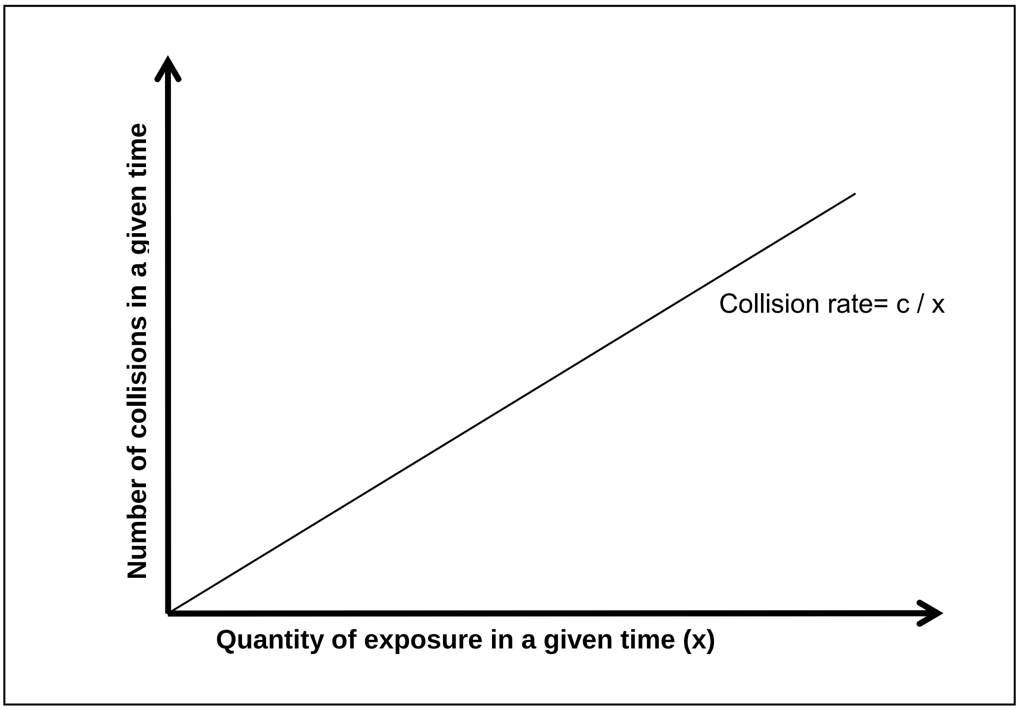


Figure 2.3: Assumed relationship between exposure and number of collisions

Although the assumption that collisions increase as a linear function of exposure is commonly made, there is substantial evidence to suggest that it is erroneous. Jacobsen (2003) has shown that pedestrian-vehicle collisions vary non-linearly with the number of pedestrians. In other words, risk appears to drop off when more pedestrians are present. Similarly, Lee and Abdel-Aty (2005) showed that pedestrian-vehicle collisions vary non-linearly with vehicle volumes. Collisions increase when more vehicles are present, but the rate of increase declines at high traffic volumes. The non-linear relationship may be due to more cautious driver behavior or reduced speed when many road users are present.

The calculation of collision rates without taking into account the non-linear relationship between exposure and collisions can lead to spurious conclusions in safety studies.

Hauer (1995) illustrated the pitfalls of collision rates using the following diagram (Figure 2.4). Accidents increase with exposure, but the rate of increase is not constant. The resulting curve is referred to as the “Safety Performance Function” of the roadway. It may be empirically measured over time with the collection of accident data in periods of differing exposure.

Hauer (1995) shows how the collision rate (the slope of the curve) at point “B” in the diagram is lower than that at point “A” simply by virtue of the fact that the exposure has risen from 3,000 to 4,000 vehicles. If this fact is not taken into account, one could incorrectly conclude that a safety countermeasure was the cause of the decline in accident rates, when a change in exposure was alone responsible.

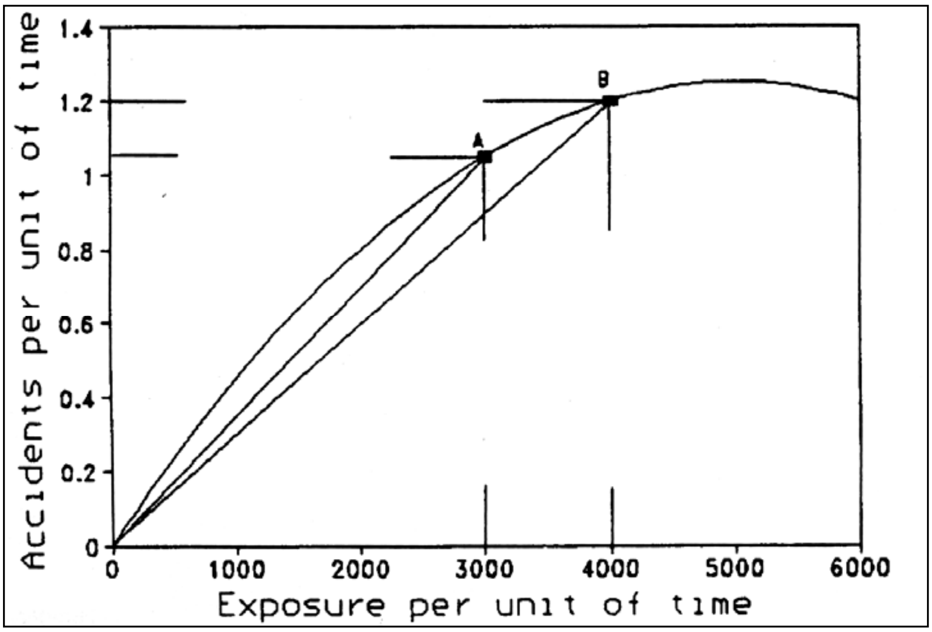


Figure 2.4 Non-Linear Relationship Between Exposure and Accidents (Hauer, 1995)

The best method of coping with the problems of accident rates is to discard them in favor of more complex models of risk. However, since risk modeling is often too costly for practical applications, accident rates are likely to remain common currency. Given that fact, it is sufficient to be aware that the usefulness of accident rates in measuring risk may be undermined in situations where exposure has changed substantially. Future studies of the relationship between pedestrian volumes and collisions are needed to define typical safety performance functions for pedestrian collisions. This will help identify the level of pedestrian exposure associated with a decline in collision rates.

2.7. Converting Between Exposure Measures at Pedestrian Crossings

As noted above, study resources may constrain the choice of exposure measure. For example, in areas with large numbers of pedestrians, recording the actual time each pedestrian spends at a crossing will require multiple observers, whereas recording the pedestrian volume will require fewer observers. In many cases, however, the estimated time a pedestrian spends crossing a street will provide a better indication of exposure than will a simple volume measurement.

In these cases, it is possible to convert the pedestrian crossing volume into an estimate of the aggregate distance crossed or time spent crossing. This can be achieved through the following equations (1) and (2).

Ped distance traveled (feet) = no. of crossings * distance crossed (ft) (2)

Ped time walked (seconds) = Ped distance traveled (ft) / 4 (ft/s)[1] (3)

Transforming pedestrian volume into time spent traveling or distance traveled at a crossing should be conducted for estimation purposes only. It should not be considered the “true” time spent traveling for the following reasons.

- ✓ Pedestrian crossing speed is not static but varies by pedestrian age; gender; pedestrian compliance with intersection controls; weather conditions; and signal cycle length (Knoblauch et al., 1996). One study noted that as many as 19 percent of pedestrians actually run across the intersection (Fitzpatrick et al., 2006).

- ✓ Pedestrians crossing distance is not static because some pedestrians may cross at an angle or walk outside the painted crossing.

- ✓ Pedestrian crossing speed alone does not fully account for crossing time because pedestrians who wait for signals to change require a “startup” time of approximately 3 seconds to begin walking (Knoblauch et al., 1996).

It should also be noted that this conversion should only be attempted for constrained areas where pedestrian distance walked can be estimated with reasonable accuracy.

---

[1] Pedestrian speed as indicated in the Federal Highway Administration 2003 Manual on Uniform Traffic Control Devices with Revision 1 Incorporated, published 2004

Observing pedestrian distance walked along a roadway, for example, is prone to error because individual pedestrians can stop, change directions, or enter and exit buildings, thus changing their distance traveled.

3. AREA-WIDE METHODS

The previous chapter illustrated the fact that there are several possible definitions of pedestrian exposure, and that the definition used in any given study is, to some extent, a function of the measurement instrument and the geographic context. This report identifies two main geographic contexts where measurement of pedestrian exposure takes place: wide areas, such as neighborhoods, cities, or the state, and specific sites, such as intersections or pedestrian crossings. These contexts can overlap when pedestrian exposure at specific sites is sampled in order to estimate exposure over a wide area.

This chapter discusses three general approaches to estimating area-wide pedestrian volumes. The first strategy involves directly sampling pedestrian activity at a representative set of sites throughout an area. The second strategy involves using surveys to gauge how much individuals report having walked in a given area. Surveys of this kind have already been implemented in some metropolitan areas and on the state level in California. The third strategy involves using modeling techniques to estimate pedestrian volumes from a combination of direct counts, surveys, and secondary data. The strengths and weaknesses of each of the methods listed above are discussed, and examples of each are provided.

3.1. Direct Sampling

Direct samples of pedestrian volume can be used to estimate pedestrian activity over a wide area. To achieve this, it is necessary to develop a strategy to sample volumes systematically through time and space. A systematic sampling design could be used to develop an estimate of the average volume at intersections in an area, for example. An in-depth discussion of representative sampling methods may be found in chapter 5, “Data Collection Planning at Intersections.”

The direct sampling approach to measuring area-wide pedestrian volumes has some distinct advantages. Direct measurements of pedestrian activity are based on real observations, rather than reported behaviors, so they avoid the problem of under-reporting of short pedestrian trips common to surveys (Schwartz and Porter, 2000). Direct measurements capture the activity of all pedestrians at the sampled site, regardless of age or economic status, although they do not capture the rich demographic information typically included in surveys. Direct measurements allow the linkage of pedestrian activity to site-specific factors such as intersection design.

Despite these advantages, there are very few examples of direct measurement approaches. This may be because of the lack of good inventories of the pedestrian network, which are necessary to devise a sampling scheme. The Institute of Transportation Engineers Pedestrian and Bicycle Council, with the assistance of Alta Planning and Design, have attempted to implement a program of pedestrian volume sampling over wide areas. This effort, known as the National Pedestrian and Bicycle Documentation Project, aims to establish a nationally consistent methodology for performing pedestrian and bicycle counts; to promote the performance of counts on official counting days during the second week of September; and to input counts into a national database (Alta Planning and Design, 2006). The project has resulted in collection of pedestrian volumes in a few cities throughout the nation. However, since there is no spatial sampling scheme associated with the project, the resulting volumes cannot be used to estimate pedestrian volumes over wide areas. The likelihood that the project will generate systematic, routinely collected pedestrian counts is small given its voluntary nature.

The best example of direct volume sampling comes from outside the pedestrian realm. The Federal Highway Administration has developed a Traffic Monitoring Guide to aid states in the systematic sampling of vehicle volumes. The guide describes a method for sampling every roadway section at least once within a six-year period, and for converting a point-measure of volume (Average Daily Traffic) into a distance-based measure (Vehicle Miles Traveled) based on the length of the roadway segment (FHWA, 2001). Although many states use the methods in the Traffic Monitoring Guide, some states, such as California, use a combination of direct counts and modeling to estimate vehicle volumes (Caltrans, 2005).

3.2. Surveys

Unlike direct sampling methods, surveys conducted at the local, state, and national level are commonly used to quantify pedestrian activity over wide geographic area. Because surveys are able to capture detailed pedestrian characteristics and preferences, they are very useful for studying the pedestrian behavior of specific groups. Surveys are also able to capture detailed trip characteristics such as the number and length of walking trips made by an individual.

In direct sampling, by contrast, it is very difficult to determine the origin and destination of each pedestrian trip, or to determine detailed pedestrian characteristics. However, surveys have certain weaknesses. Surveys do not generally link pedestrian activity to specific infrastructure, such as roadway or sidewalk width, so it is difficult to determine the relationship between infrastructure and pedestrian activity from surveys alone. It is also difficult to determine whether the walking trips reported in surveys were made in areas where the pedestrian was exposed to traffic. Lastly, walking trips are commonly underreported in surveys, because individuals do not always remember short walking trips (Schwartz and Porter, 2000). For example, individuals may not report walking to access transit as a separate trip.

Survey data is available for many different types of geographies and time periods. When seeking information about pedestrian exposure over a wide area, it is important to know whether relevant survey data has already been collected. For that reason, this section focuses on describing existing pedestrian-related surveys and the type of information available from each. Three types of existing surveys are identified and evaluated: (i) health-related surveys; (ii) travel surveys; and (iii) the Journey-to-Work portion of the U.S. Census. These characteristics are also summarized in Table 3.2.

There will be cases where existing surveys will not always meet the data needs of the user. For example, there is no existing data source that provides an estimate of pedestrian exposure for the state of California as a whole on a frequent basis. In these cases, institutional support and resources are needed to implement more frequent or new data collection efforts.

3.2.1. Health-Related Surveys

Health surveys aim to track health conditions and risky behaviors. Since walking is a form of physical activity, some of these surveys include walking-related questions, which tend to be focused on whether the respondent obtained a healthy amount of physical activity. Therefore, these types of surveys may not contain information on they exact amount of walking or whether walking took place in areas where pedestrians were exposed to traffic.

For example, the California Department of Health Services and the California Department of Transportation sponsored the Pedestrian Characteristics in California Survey in 2003 in order to track health trends. The survey included a question on the amount of time spent walking in a typical week (Schneider et al., 2005). Because the survey is not conducted on a regular basis, it is limited in its ability to track pedestrian volume trends over time, and it does not provide information about the total amount of exposure to traffic.

The Behavioral Risk Factor Surveillance System (BRFSS), an annual telephone survey administered by the Centers for Disease Control, is conducted annually. It includes questions on physical activity, but does not distinguish between walking and other forms of physical activity (BRFSS, 2006). The state of California could choose to add additional questions to the BRFSS in order to gain information about the prevalence of walking in the state.

3.2.2. Travel Surveys

Travel surveys are conducted at the metropolitan, state, and national level for transportation planning purposes. Most rely on travel diaries, in which respondents record detailed information about trips taken during a designated travel period. The detail provided by travel diaries is valuable in estimating pedestrian volume, because it allows volume to be expressed in terms of the amount of time walked, the distance walked, or the number of walking trips made.

The largest travel survey conducted nationally is the National Household Travel Survey (NHTS). The survey is conducted about every six years by the Federal Highway Administration, and records the travel patterns of about 20,000 randomly selected U.S. households. The NHTS reports the number of trips by mode that respondents took in the week the survey was administered. It can be used to quantify pedestrian trips as a share of all trips taken nationally or by major Census division (e.g. Mountain; Pacific, West South Central, etc.). The NHTS is not intended for use at the state or sub-state levels, but states or metropolitan areas can purchase add-ons (NHTS, 2006).

Several states and metropolitan areas also conduct travel surveys to serve local needs (TRB, 2006). In the state of California, travel surveys are conducted in several metropolitan areas and on at the state level. The California Statewide Household Travel Survey (CSTS), a travel survey of 17,040 California households, was conducted between 2000-2001 by the California Department of Transportation (Caltrans). The CSTS quantifies the number, duration, and approximate distance of trips taken by survey respondents on an average weekday for each mode of transportation. It also captures household demographic and economic characteristics.

The CSTS provides a robust estimate of the amount of pedestrian activity in the state of California, and for 17 sub-state regions, for the year 2000. The survey must be used cautiously or not at all for small geographic areas such as cities or counties (Caltrans, 2002). In addition, the CSTS cannot be used to track short-term trends in pedestrian activity because it is not conducted on a regular basis.

Several metropolitan areas in California also collect travel surveys similar to the CSTS and the NHTS. For example, the Metropolitan Transportation Commission conducts the Bay Area Travel Survey (BATS) a study of the travel patterns of approximately 15,000 Households in the 9-county Bay Area. The BATS was conducted in 2000, 1996, 1990, 1981, and 1965. The Sacramento Area Council of Governments and the Southern California Association of Governments also conduct travel surveys about once a decade.

3.2.3. U.S. Census Journey-to-Work and the American Community Survey

The Journey-to-Work component of the U.S. Decennial Census long form contains detailed information about the work-trip characteristics of one in six U.S. households. Respondents are asked about the location of their workplace; their usual means of transportation to work; and the amount of time it usually took them to get to work. The data is free to the public, available online, and covers large and small geographies throughout the nation.

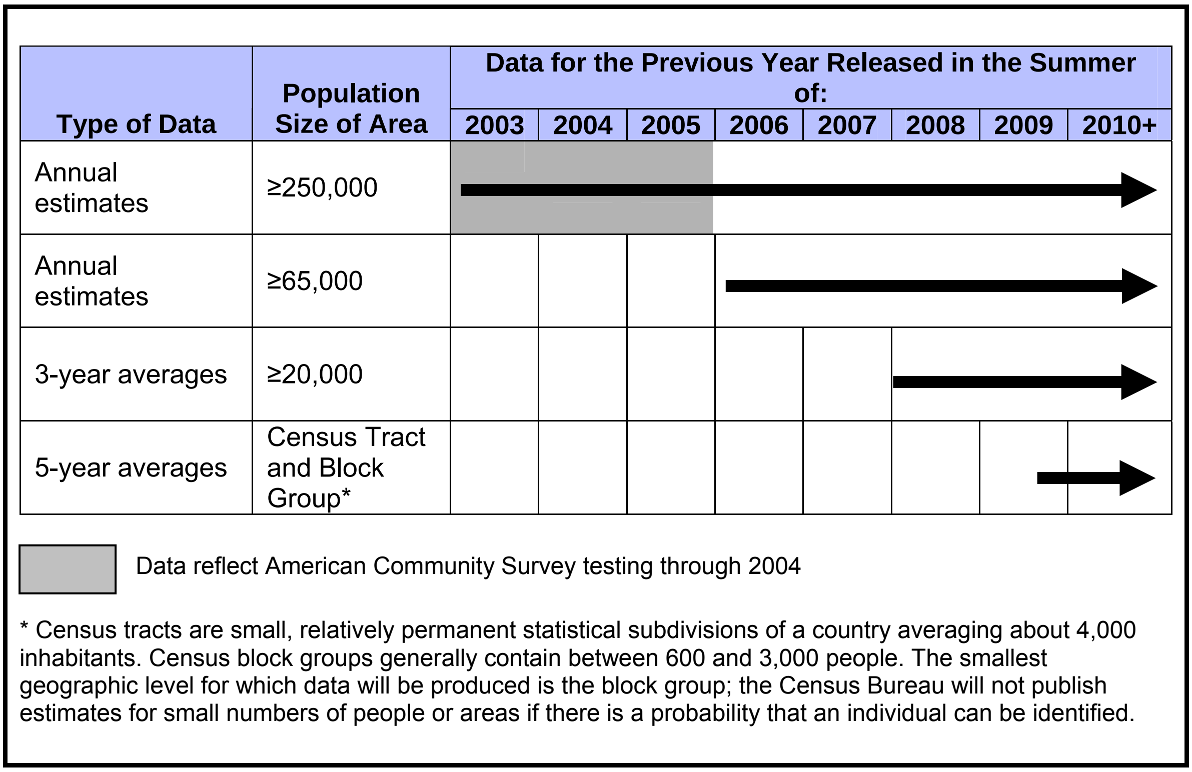
However, Journey-to-Work data has some limitations. The survey questionnaire asks only about which mode of transport the respondent used most frequently to commute to work in the previous week. By doing so, it accounts only for work trips, which make up a minority of all walking trips (Komanoff and Roelofs, 1993), and for employed adults, who make up less than half of the population (U.S. Census Bureau, 2004). Moreover, the form asks how the respondent “usually” got to work, and thus does not capture occasional trips to work made by another mode. Neither does it account for walking trips made as a component of the work trip, such as trips to and from a bus stop. This is because the survey questionnaire asks the respondent to name only the mode they used for the majority of the distance of their trip (U.S. Census Bureau, 2005).

In spite of these weaknesses, Census Journey-to-Work data has been used as proxy for pedestrian exposure because it provides some information about how much people are walking in an area, and is often the only data on walking available at the level of the city. One widely-known report on pedestrian safety, which was published by the Surface Transportation Policy Project, used the percentage of people walking to work and population data from the Census to compare pedestrian risk in metropolitan areas across the nation (STPP, 2002, 2004).

The Census long form that provides Journey-to-Work data is currently being replaced by a new product called the American Community Survey (ACS). Although the information being collected in the ACS is the same as what was collected in the Census long form, the two surveys differ in important ways. The most important difference is that Journey-to-Work data will be available every year through the ACS, rather than once a decade. Another important difference lies in the sample design. Whereas the Census long form data was collected during a specific week in April, the ACS samples households on a rolling basis during each month of the year. This means that ACS data will reflect traveler behavior throughout the year rather than for a specific season. When fully implemented, the ACS will sample about 3 million, or 1 in 10, U.S. households annually.

ACS data are currently available for communities of 65,000 or more on a yearly basis. For smaller communities, it will take between several years to accumulate enough samples to provide data. Beginning in 2008, yearly estimates based on three year averages will be available for communities of 20,000 or more, and beginning in 2010, yearly estimates based on five-year averages will be available at the Census tract and block group level A summary of ACS data availability is displayed in Table 3.1.

Table 3.1: Block Group Level A Summary of ACS Data Availability



Type of Data	Population Size of Area	Data for the Previous Year Released in the Summer of:							
		2003	2004	2005	2006	2007	2008	2009	2010+
Annual estimates	≥250,000	▨ →	▨ →	▨ →	→	→	→	→	→
Annual estimates	≥65,000				→	→	→	→	→
3-year averages	≥20,000						→	→	→
5-year averages	Census Tract and Block Group*							→	→

▨ Data reflect American Community Survey testing through 2004

* Census tracts are small, relatively permanent statistical subdivisions of a country averaging about 4,000 inhabitants. Census block groups generally contain between 600 and 3,000 people. The smallest geographic level for which data will be produced is the block group; the Census Bureau will not publish estimates for small numbers of people or areas if there is a probability that an individual can be identified.

Source: U.S. Census Bureau, 2006

Table 3.2: Characteristics of Existing Pedestrian Related Surveys

Survey	Walking Question	Geographies	Years available
Decennial Census	Usual mode to work	Census tract → nation	1980, 1990, 2000
American Community Survey	Usual mode to work	Census tract → nation	Every year after 2003*
Behavioral Risk Factor Surveillance System	None-possible add on	States, nation	Every year
National Household Travel Survey	Number, length, duration of walk trips	Census divisions, nation	Every 6 years: 1969, 1997, 1983, 1990, 1995, 2001
California State Travel Survey	Number, length, duration of walk trips	Caltrans Districts, state of California	Every 10 years
Metro Area Surveys	Number, length, duration of walk trips	SF, La & Sac metro area	Varies –about every 6-10 years

*ACS release schedule varies by geography; data at the census tract level not available until 2010

3.3. Modeling Methods

Mathematical models can be used to estimate pedestrian volumes by combining key assumptions with existing data. If properly calibrated and tested, models can be powerful tools in estimating pedestrian volumes when direct measurement is not feasible. The advantages and disadvantages of modeling depend to some degree on the model itself, but in general, models have the potential to save time and resources without overly compromising accuracy.

Radford and Ragland (2006) identified three main types of models: sketch plan models, network analysis models, and microsimulation models. The strengths and weakness of each for measuring pedestrian exposure are presented below.

3.3.1. Sketch plan models

Sketch plan models use available data to estimate pedestrian volumes for regional or city-wide planning purposes. These models rely on known or estimated correlations between pedestrian activity and adjacent land uses, such as square feet of office or retail space, and/or indicators of transportation trip generation such as parking capacity, transit volumes, or traffic movements (Schwartz et al., 1999). Some of these models are not capable of producing pedestrian volumes, but rather produce a dimensionless indicator of pedestrian activity.

The city of Sacramento, California, recently used a sketch plan method developed by Fehr and Peers Transportation Consultants (2005) as part of its pedestrian master plan. The method inputs demographic, economic and land use variables associated with walking into Geographic Information Systems software to produce a dimensionless “pedestrian demand index” for each street segment in the city.

3.3.2. Network analysis models

Network analysis models are more complex than sketch plan models because they rely on a map or model of the pedestrian network. As a result, they are capable of estimating volumes for specific street segments and intersections over an entire city or neighborhood. Although the models vary in technique, most use a variation on the four-step modeling approach to generate and distribute trips based upon assumptions about the amount of walking trips in a study area and various route choice algorithms (Senevarante and Morall, 1986; Ben-Akiva and Lerman, 1985; McNally, 2000).

Radford and Ragland (2004) used a network analysis model, Space Syntax, to estimate pedestrian volumes on streets and intersections throughout Oakland, California. The model required input of a pedestrian route map derived from publicly available Census TIGER/line GIS centerline road maps; population and employment data from the U.S. Census and the California Economic Census; and raw pedestrian count data needed to calibrate the model. The model produced reasonable estimates of city-wide pedestrian volume.

The Space Syntax model is also useful for estimating pedestrian flow along corridors. This is very helpful because direct measurement of flow along corridors is difficult. It may be achieved by dividing the road network into small segments, such as a block length, and assuming that flow along the segment is constant. This is not always a fair assumption because of the complexity of pedestrian movement. For example, if a pedestrian is counted at the end of a block, it is uncertain whether she has been traveling for the entire block or if she just exited a building. With vehicle volumes, by contrast, it is often assumed that any vehicle passing through a point has been traveling along the length of the segment (FHWA, 2001). Space Syntax provides an alternative method of estimating flow along many corridors with a small set of samples as input.

3.3.3. Microsimulation models

Microsimulation models use flow principles from physical science to model pedestrian behavior in confined spaces such as the interior of shopping malls or subway stations, on a single or small number of streets, or within building interiors. Microsimulation models provide highly accurate, detailed information about pedestrian movement, but require specialized software, knowledge and extensive data inputs (Radford and Ragland, 2006).

3.3.4. Comparison of modeling techniques

Table 3.3 presents a comparison of these approaches, highlighting their advantages and disadvantages for estimation of wide-area pedestrian volumes. This table was adapted from Radford and Ragland (2006). Each of the modeling approaches discussed in this paper is suited to a different scale of geographic analysis. Sketch plan models are best for broad regional or statewide analysis; network analysis models are appropriate for corridor, neighborhood, or urban area analysis; and microsimulation models are best for a single street or smaller area.

Relevant literature indicates that sketch plans have the most potential to be put into standard use for estimating pedestrian volume throughout the state. While less accurate than other types of models, sketch plans are relatively simple to use and make the most out of existing data sources. A simple, standardized sketch plan method would be an improvement over the current absence of volume estimation methods in many areas.

Microsimulation models are much too complex and costly to be practical beyond the level of the street or intersection. Network analysis models have been successfully used to estimate pedestrian volumes in most large urban areas, but may be impractical in many small cities and rural areas that lack staffing and resources to perform the GIS analysis and calibration necessary to complete the model.

Table 3.3: Comparison of Modeling Methods

	Scale of Application	Advantages	Disadvantages
Sketch Plan	Large scale (city, region, state)	Little data collection required; No specialized expertise needed; Quick estimations.	Aggregate level; Low accuracy.
Network Analysis	Urban and neighborhood level	Good detail; Reasonable accuracy; Limited data requirements; Useful for estimating pedestrian flows along corridors; Appropriate to urban volume analysis.	Model must be calibrated with pedestrian counts; Requires existing GIS data; Must be submitted to sensitivity test.
Microsimulation	Individual Streets or intersections	Highly accurate; Detailed; Allows visualization of pedestrian flow.	Complex; Steep learning curve; Significant initial data requirements.

3.4. Comparison of Methods

This chapter reviewed and evaluated three possible systematic approaches to measurement of pedestrian volumes over wide areas. The choice of area wide counting methods depends on budget constraints and data needs, and the availability of existing data. No single approach is best, but each has strengths and weakness. These are summarized in Table 3.4.

Table 3.4: Comparison of Approaches to Pedestrian Volume Estimation

Approaches	Advantages	Disadvantages
Direct sampling methods	Based on real, not reported pedestrian activity; All pedestrians at each site are sampled; Pedestrian volumes linked to specific sites; If designed appropriately, data could be aggregated from small to large geographies.	Difficult to devise a sampling scheme; Need a good inventory of the pedestrian network; Would require significant manpower; No demographic or attitudinal information captured; No information on distance, length, or time walked.
Survey methods	Can capture demographic and household data; Can capture distance, length, and time walked; Existing surveys could be adapted / expanded.	Walk trips are consistently underreported in surveys; Difficult to link walking to specific infrastructure; Difficult to determine whether walking occurred in areas exposed to vehicle traffic.
Modeling methods	Make the most of available data; Dynamic and flexible; Potential for lowest cost.	Different models may be needed for different geographic areas; Output may be limited to dimensionless measure of pedestrian demand.

4. SITE SPECIFIC METHODS

The previous chapter discussed approaches to measuring pedestrian exposure over wide areas such as cities or states. In many cases it is necessary to collect pedestrian exposure data at specific sites such as intersections, pedestrian crossings, or along a city block. Site-specific measurement of pedestrian exposure is used to identify high collision locations; to evaluate how infrastructure influences pedestrian risk; or to track changes in risk over time at a specific site or sites.

There are three main methods of counting pedestrians at specific sites: (i) field observation (ii) video observation with manual review and (iii) automated methods. This chapter describes these methods and evaluates the strengths and weakness of each.

4.1. Pedestrian Counts at Specific Sites

Pedestrian volumes at specific sites are usually collected directly using either (i) manual counts taken by collectors in the field or through video observation, or (ii) automated counts using specialized equipment. Push button counters are also used to count pedestrians. However, because of their lack of accuracy relative to the other counting methods, push button counters were not reviewed in this protocol. It has been determined that only 35 percent of all pedestrians use push button devices when they are available (Zeeger et al., 1982).

Pedestrian counting methods differ in their cost, convenience, level of data detail, and accuracy. In order to select the most appropriate method for different conditions and study purposes, it is important to understand the strengths and weaknesses of each method.

4.1.1. Manual counting methods

Manual counting methods are frequently used to quantify all types of transportation activity, including vehicle, bicycle, and pedestrian volumes. Manual methods are the most frequently used method of counting pedestrians, particularly for studies that require small samples of data at specific locations, such as pedestrian crossings. The two most common manual counting methods used to measure pedestrian flows at crossings are:

- ✓ *Field observations:* in which pedestrians are observed in the field and counted by hand.

- ✓ *Video-recordings:* in which camera recordings of pedestrian crossings are taken and then processed through playback and manual recording.

Field observations are typically used for periods of less than a day. In this case, the normal intervals for counting are 5, 10, or 15 minutes. The counts are recorded with tally sheets, hand-held computers, or clickers. Tally sheets can include an individual line for each pedestrian and his or her characteristics and/or behavior can be recorded, although not all tally sheets are designed this way. Some include only boxes in which the number of pedestrians crossing within a certain time are recorded. An example of a field sheet used to count pedestrians and make inferences about their characteristics is provided in Appendix A. Hand-held computers (PDAs) are more frequently used to count and classify vehicle movements, but can also be used to collect information about pedestrian flow and movement directions.

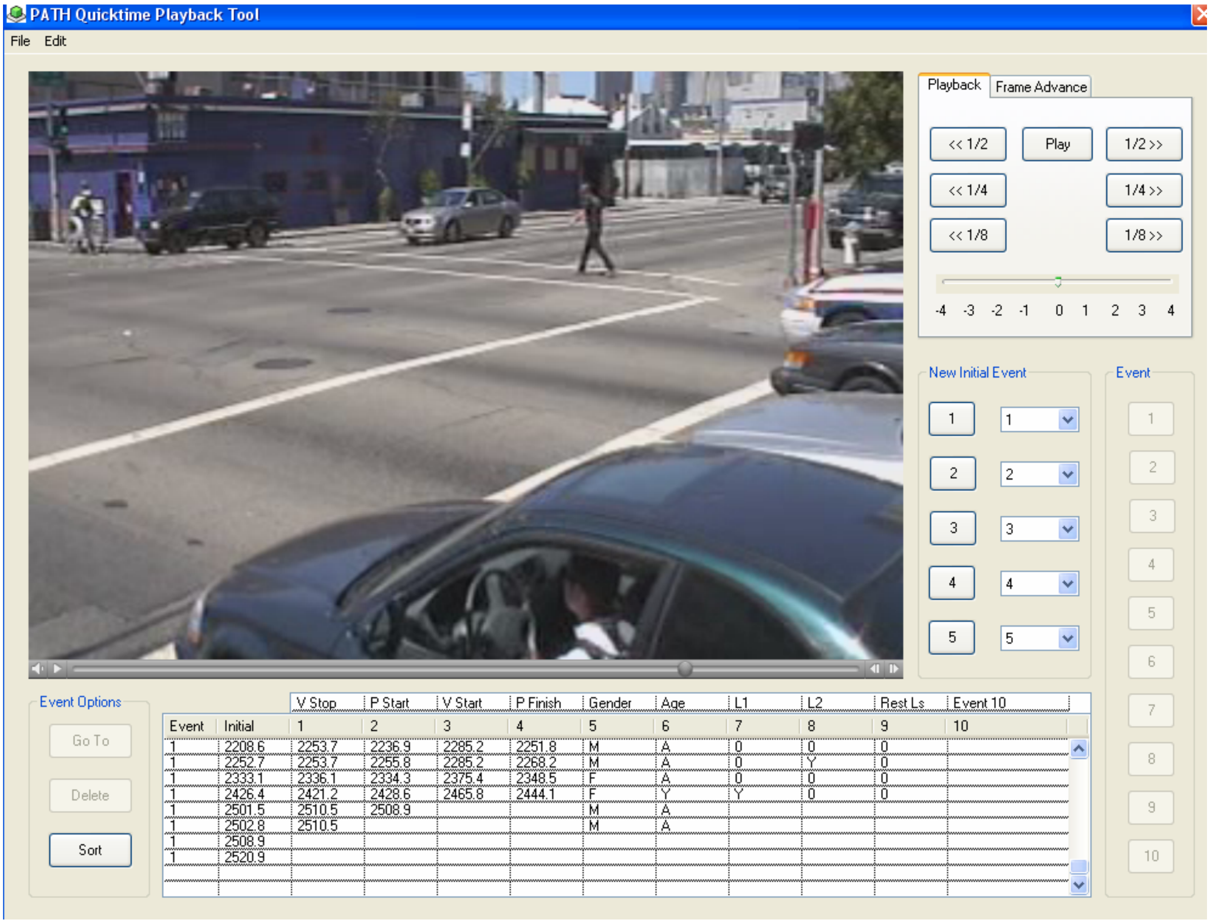
Clickers, Figure 4.1, are appropriate in situations where there is no need to record individual pedestrian characteristics. They are also helpful in areas of high volume, where it is important that the observer have his or her eyes focused on the street. Schweizer (2005) found that a person can count about 2,000 to 4,000 pedestrians in an hour using clickers, and only half that amount without them. Using more than one clicker, the field observer can factor in the difference between males and females or the direction of movement. However, recording these characteristics would decrease the capacity of the field observer.

Manual-video recording uses cameras to record images of pedestrians which are later reviewed by an observer. The observer records the number of pedestrians as well as pedestrian characteristics and behavior, if needed. Detailed review of behaviors, or crowded pedestrian conditions, may require that the observer review the video in variable time (e.g. slowing and speeding the video as needed). Specialized video-playback tools may be used to facilitate review of the videos. One such tool was developed by the Partners for Advanced Transit and Highway Research (PATH), and is depicted in Figure 4.3.

The central issues with the manual-video method of counting pedestrians are the need for a good camera angle and resolution (Figure 4.2) and the long time required to review the video tapes, estimated to be three times the tape length (Diogenes et al., 2007).

Figure 4.1: Field Observation using clickers (Schweizer, 2005)

Figure 4.2: Video-image camera angle and resolution



		V Stop	P Start	V Start	P Finish	Gender	Age	L1	L2	Rest Ls	Event 10
Event	Initial	1	2	3	4	5	6	7	8	9	10
1	2208.6	2253.7	2236.9	2285.2	2251.8	M	A	0	0	0	
1	2252.7	2253.7	2255.8	2285.2	2268.2	M	A	0	Y	0	
1	2333.1	2336.1	2334.3	2375.4	2348.5	F	A	0	0	0	
1	2426.4	2421.2	2428.6	2465.8	2444.1	F	Y	Y	0	0	
1	2501.5	2510.5	2508.9			M	A				
1	2502.8	2510.5				M	A				
1	2508.9										
1	2520.9										

Figure 4.3: Path QuickTime Playback Tool

4.1.1.1. Cost of manual methods

The relative cost of field observations and video-recording counts varies based on the source of labor, the volume of the intersection, and the amount and type of data being collected. Costs can be broken into labor and equipment costs. In general, field observations are labor intensive but have low equipment costs. Video methods have higher equipment costs, and may have equally high if not higher labor costs depending on the amount of staff time taken reviewing video, and on whether the video camera can be left unattended in the field.

Cost of Manual Field Observations:

- ✓ Few equipment costs, though they may be increased if electronic hand-held devices (PDAs) are used to record pedestrian activity. However, use of these devices reduces the labor costs associated with data entry

- ✓ High labor costs. Staff are needed to observe pedestrians in the field and to perform data entry. More staff are needed at high-volume intersections, when several data points are being collected (e.g. pedestrian characteristics), or when detailed pedestrian behaviors are being investigated

- ✓ Training costs vary. The cost of training relates to whether consultant observers are used or whether observers are on staff, and to the need for data quality. Generally, more training can be expected to produce better quality data

- ✓ Costs can be reduced if counts performed by volunteers/students; if counts are integrated in to regularly scheduled vehicle counts (Schneider et al., 2005); and if counts are scheduled efficiently to maximize the use of available labor

- ✓ Example: the District Department of Transportation performs 10-hour counts at intersections across the city. The Department estimated in 2005 that each intersection counted cost between $400 - $500, including the cost of labor for pedestrian and motor vehicle counts and the cost of entering the field data into spreadsheets (Schneider et al., 2005)

Cost of Manual Video Recording:

- ✓ Equipment costs include the price of camcorders, tapes, and recording accessories. Camcorders vary in price depending on the quality required, but range from hundreds to a few thousand dollars. The cost of video tapes varies by number of hours recorded

- ✓ High-resolution or time-lapse video equipment may be required to record detailed pedestrian characteristics, or to monitor more than one crossing at a time. For example, the City of Davis, California, purchased a time lapse video system (including camera, playback system and videotapes) for $7,000 in 1998/99 (Schneider et al., 2005).The cost-burden of video equipment should be assessed over the life of the equipment

- ✓ Costs can be reduced if video counts are combined with other purposes, such as security
- ✓ Staff are needed for initial setup of camera and camera maintenance
- ✓ One staff person per video camera is typically required in the field to prevent vandalism and theft, unless the camera is concealed or made inaccessible
- ✓ Only one staff person is needed to review the video, regardless of the intersection volume, because video can be slowed down and rewound. However, staff may take many hours to review the video if detailed information or a high level of accuracy is required
- ✓ Transportation costs must be paid for staff and video camera. In some cases, a flat-bed truck may be required for set up of the video camera

4.1.1.2. Convenience and data detail of manual methods

Field observations and video-recording differ in their relative convenience and in the data detail that can be collected. Generally, field observers can capture a broad array of pedestrian characteristics and behaviors. Video-recordings are sometimes capable of capturing these details, but not without careful camera positioning and/or high resolution film. Video cameras may be able to record at times inconvenient for field observers, such as night time or weekends; however, this is only possible if the video is positioned or disguised such that it can be left alone without protection from vandalism or theft, and if the video image is unobscured by poor lighting or weather.

Convenience and Data Detail of Manual Field Observations:

- ✓ Staff schedules must be coordinated
- ✓ Inconvenient to collect data during inclement weather or during night/weekend hours
- ✓ Can waste labor time in areas of low volume
- ✓ Possible to capture detailed pedestrian characteristics like age, race, and specific behaviors (Mitman and Ragland, 2007; Diogenes et al., 2007)

- ✓ Difficult to record extra details if pedestrian volumes are high, unless additional staff are used
- ✓ Possible to capture mid-block crossings if observers trained properly
- ✓ Possible for a single staff person to observe multiple crossings if pedestrian volume is low
- ✓ Difficult to record the amount of time it takes pedestrians to cross
- ✓ Possible to record detailed information about the setting or nearby events that are not captured within a camera’s field of view

Convenience and Data Detail of Manual Video Observations:

- ✓ If camera is positioned securely and disguised such that no on-site videographers are required to protect it, data can be collected at inconvenient times such as nights and weekends, assuming there is adequate lighting at the site
- ✓ If camera is rain-proof, it is possible to collect data during inclement weather
- ✓ Difficult to find a suitable place for video camera. Installation and use of cameras requires permits as well as security and safety procedures to protect the camera and those around it. For example, permits are typically needed to park a flat-bed truck near an intersection, and police must be notified so they do not suspect illegal activity.
- ✓ Difficult to capture pedestrian characteristics such as age or behavior without expensive cameras or precise positioning
- ✓ Presence of camera may influence pedestrian behaviors
- ✓ Cannot capture crossings from multiple directions unless multiple cameras are used or camera positioned at a very wide angle, which may compromise the image quality
- ✓ Cannot capture pedestrian behavior outside of the camera’s field of view

- ✓ Possible to capture time and speed
- ✓ Cannot capture detailed information about the setting

4.1.1.3. Accuracy of manual methods

It is important to understand the accuracy of each counting method in order to make adjustments to counts or to choose the method with the desired level of accuracy. Although there are few empirical studies of the error of pedestrian counting methods, it is possible to identify and discuss the sources of error in each. In general, the accuracy of manual counts is affected by the level of observer training and attention, and whether the observer is in the field or reviewing video recordings. Mitman and Ragland (2007) compared the inter-reliability between different field observers and found there is a significant and measurable difference in the data quality produced by observers with different levels of motivation.

In both methods, error can be avoided by choosing observers carefully, conducting adequate training, and matching the collection method with location scenarios (Mitman and Ragland, 2007). However, video-recordings provide additional insurance against lack of observer motivation because they can be reviewed multiple times by different observers to check data quality.

Sources of error in manual field observations:

- ✓ Lack of attention. The motivation and training of field observers may affect their attention in the field.
- ✓ Differences in judgment. The unique personality attributes of field observers may affect their ability to judge pedestrian characteristics and behaviors, such as age and gender.
- ✓ Level of pedestrian activity. The amount of pedestrian activity may impact the accuracy of the count in a variety of ways. Very low or high volumes can impact the observer’s attention and their ability to record all data points. More research is needed to determine the relationship between pedestrian volume and the accuracy of field observations.

- ✓ Amount of data needed. If it is necessary to record several data points for each pedestrian (e.g. gender, direction, age), the quality of the data recorded may decrease if the capacity of the observer is exceeded, or if recording the data requires the observer to take his or her eyes off the street.

- ✓ Length of time collecting data. If the collection period is long, the observer may take unscheduled breaks or get distracted.

Sources of error in manual video recordings:

- ✓ Lack of quality images. The camera angle, positioning, and image resolution affect the quality of the image and therefore the ability of the video observer to discern individual pedestrians and their characteristics.

- ✓ Differences in judgment. As with field observation, the attributes of video observers may affect their judgment of pedestrian characteristics.

- ✓ Lack of attention. As with field observation, the motivation and training of video observers may affect their attention. However, video recordings can be reviewed multiple times to ensure data quality.

- ✓ Traffic composition. Large vehicles may block the view of the crossings and render the video unusable in some instances. In contrast, field observers can adjust their viewing angle in real time to continue the observations and therefore eliminate this issue (Mitman and Ragland, 2007).

- ✓ Level of pedestrian activity. The level of pedestrian activity does not much affect the quality of counts because video can be reviewed in variable time to ensure all pedestrian are counted. However, the level of pedestrian activity may increase the time required to review the video, which may negatively impact the motivation of the video observer.

- ✓ Gaps in data collection. Data may be lost, and accuracy affected, when recording is stalled to change tapes, and if the camera malfunctions or is vandalized during counting.

4.1.2. Automated methods

In general, automated counting of pedestrians is advantageous because it can reduce the labor costs associated with manual methods. It also has the potential to record pedestrian activity for long periods of time that are currently difficult to capture through traditional methods.

Automated methods are commonly used to count motorized vehicles, but are not frequently used to count pedestrians at this time. This is because the automated technologies available to count pedestrians are not very developed, and their effectiveness has not been widely researched. Moreover, most automated methods are used primarily for the purpose of detecting, rather than counting, pedestrians (Dharmaraju et al., 2001; Noyce and Dharmaraju, 2002; Noyce et al., 2006).

A review of pedestrian detection technologies was performed by and Noyce and Dharmaraju (2002) and by Chan et al. (2006). Technologies include piezoelectric sensors, acoustic, active and passive infrared, ultrasonic sensors, microwave radar, laser scanners, video imaging (computer vision). A detailed review of these technologies and their potential for counting, not merely detecting pedestrians is being conducted for this project, and will be presented in the final report.

Of the technologies listed above, those most adaptable to the purpose of pedestrian counting are video imaging (computer vision) and passive infrared devices. Video imaging utilizes intelligent processing of digital images of pedestrians captured with a video camera (Figure 4.4) that is mounted above the area of pedestrian movement. The processor subtracts the static background from the image and then tracks the remaining objects to determine whether they are pedestrians (CLP, 2005).

Passive infra-red devices count pedestrians by tracking the heat emitted by moving objects. The company “Irysis”, based in Great Britain, has developed infrared pedestrian counting devices that can be located either in or outdoors, and are mounted directly above the area of pedestrian activity (Figure 4.5). These sensors have the advantage of being relatively easy to install and configure, and are not affected by lighting conditions since they rely on heat to produce the images (CLP, 2005).

Figure 4.4: Video Imaging for Counting Pedestrian (CLP, 2005)

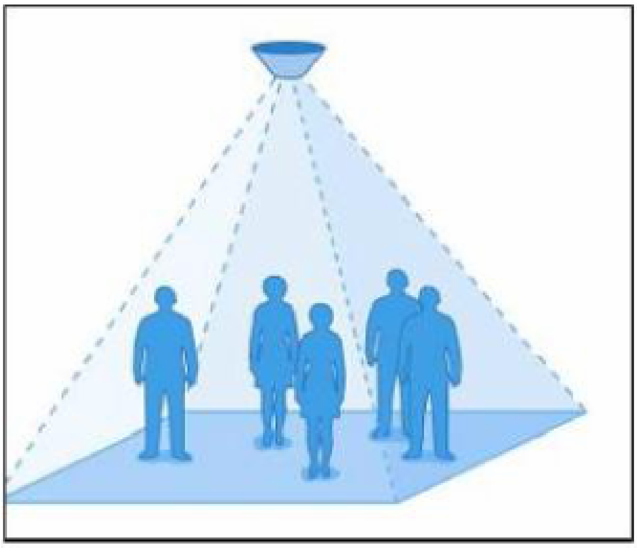


Figure 4.5: Irysis Infrared Pedestrian Counting Device (CLP, 2005)

4.2. Comparison Between Methods

The choice of pedestrian counting method should be based on the accuracy level desired, budget constraints, and the project data needs. For example, manual counts must be used when the effort and expense of automated equipment are not justified or when information about pedestrian characteristics or behavior is required.

To guide the selection of a method, it is important to review the advantages and disadvantages of each in collecting pedestrian exposure data at specific sites (Table 4.1). As specific advantages and disadvantages of the automated methods depend on the particular technology, only general aspects of these methods are highlighted.

It is important to emphasize that little is known about the relative accuracy and reliability of these methods. Field tests were performed within the context of this project to compare the particularities of the manual methods and a summary paper was submitted to the Transportation Research Board Conference (Appendix B). However, further work is needed to draw more specific conclusions about these methodologies.

Table 4.1: Comparison of Methods to Count Pedestrian at Crossings

Method	Advantages	Disadvantages
Field Observations	Relatively low cost; Observer can record detailed pedestrian characteristics and behaviors (Tally sheet)	Labor-intensive; Difficult to control the counting process; Problems at night, in unsafe locations, and during rainy weather; Cannot check accuracy of counts after they occurred;
Video Observations	Small error rate; Can replace several counters; Evaluation can be repeated several times; Possible to observe characteristics of road environment.	Difficult to find suitable place for video camera; May be gaps in the counting process (battery and tape change); Labor intensive (long analysis time) if good data quality is required; Can be hard to identify pedestrian characteristics and behaviors.
Automated Methods	Can collect data for long periods; Data storage is less time consuming.	Capital cost may be high; Specialized training may be required; Can not collect pedestrian characteristics / behavior.

5. DATA COLLECTION PLANNING AT INTERSECTIONS

Another aspect of site-specific measurement of pedestrian volume is the issue of where to collect data. The ideal would be to collect pedestrian volumes at all intersections of a city, but most projects have both budget and time constraints. In this case, a sample of the target population of sites must be selected for study. Nassirpour (2004) points out that there is no uniform standard of quality that must be reached by every sample and that the quality of the sample depends entirely on the stage of the research and how the information will be used. So, the development of a sample design that satisfies the project goals is crucial to obtain the necessary data efficiently.

This chapter describes a simplified set of statistical issues that should be considered when designing a methodology for collecting pedestrian volumes at intersections for different purposes. The proposed methodology is based on the recommendations of the Bureau of Transportation Statistics (BTS, 2003, 2005).

5.1. Sample Design Issues

Sample design is composed of three critical tasks: (i) definition of the target population; (ii) selection of sample technique; and (iii) determination of sample size. All these tasks have as constraints the objectives of the research, the type of the study and the resources available for the study, as shown in the Sampling Strategy Scheme of a Sampling Strategy (Figure 5.1). These constraints will play an important role when selecting the sample technique and determining the sample size.

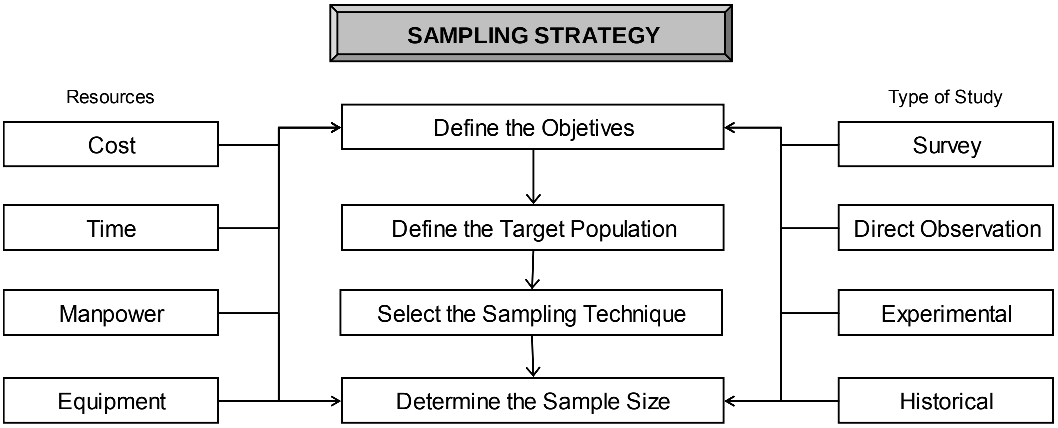


Figure 5.1: Generalized Model of Sampling (Adapted from Aggarwal, 1988 and Nassirpour, 2004)

5.1.1. Definition of target population

The target population can be defined as the complete set of sites from which you need to collect information (Nassirpour, 2004). Determining the population targeted is the first step in the sampling strategy and it is dependent on the study objective. For example, if you want to quantify pedestrian volume in the downtown's intersections, your target population is all the intersections in the downtown area. If you are interested in determining the average pedestrian volume in signalized intersections in California, so all signalized intersections within the state of California is your target population.

When defining the target population you must define the project objectives and specifications clearly to avoid collecting unnecessary data or generating bias. For example, if you want collect pedestrian volumes at marked and unmarked crosswalks you must define how to identify and distinguish between these intersections and define the geography of the study area.

After defining the target population, the operational sampling frame must be constructed. The sampling frame is a list of sampling units from which the sample can be selected at each sampling stage (Aggarwal, 1988). For example, in a study of intersection in the central business district, the sampling frame would be a database of all the intersections within the area. Ideally the target population must be coincident with the available list of sampling units. In situations where a complete database of the sampling units is unavailable, it is necessary to adjust the sample from the frame population to the target population.

In traffic observation studies, the Geographic Information Systems (GIS) and digital road databases are commonly used to develop the sampling frame (Shapiro et al., 2001). GIS can be very useful in defining the sets of intersections that are eligible for sampling, and can also provide additional information about the site, such as the number of pedestrian collisions.

5.1.2. Selection of sampling technique

After selecting the target population it is necessary to choose a sampling technique (Figure 5.2). The first step in selecting this technique is to decide whether to use non-probabilistic or probabilistic sampling.

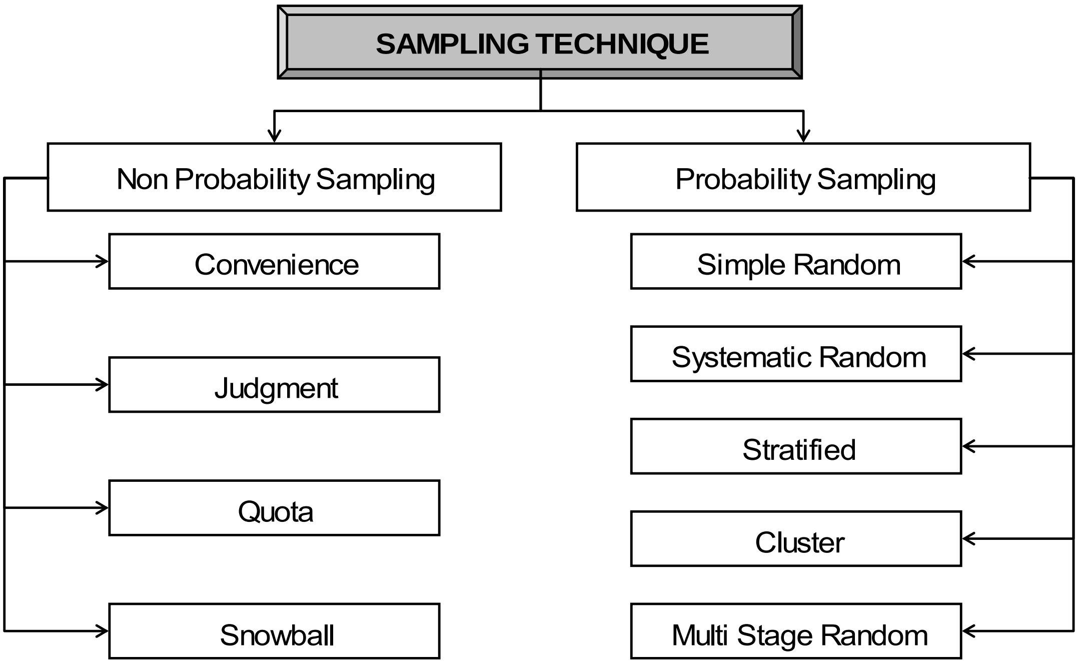


Figure 5.2: Classification of Sampling Techniques (Adapted from Aggarwal, 1988)

The non-probabilistic samples are selected through non-random methods, where the researcher has a lack of control over the sampling error. This type of sampling is most often used in experimental studies or case studies, when the researcher is interested in specific units or individuals and not in making conclusions about an entire population.

Non-probabilistic samples do not require the determination of sample size. Instead, the researcher will typically select a small number of samples based on subjective criteria. Table 5.1 describes in few words some of the existing non-probabilistic sampling techniques, pointing out the advantages and disadvantages of each method.

In contrast to non-probabilistic sampling, probabilistic sampling involves the use of statistical principles to select units or individuals randomly. This allows the researcher to calculate the sampling error and to make inferences about the target population. Probabilistic sampling requires more time and money to design the sample and to calculate the sample size necessary to obtain a representative sample. Table 5.2 describes the most frequently used probabilistic sampling techniques.

It is important to keep in mind that the selection of a sampling technique must be based on the research objectives and on the type of study.

Table 5.1: Non-Probabilistic Sampling Techniques

Non-probabilistic method	***Definition***	***Example***	***Advantage***	***Disadvantage***
Convenience	Obtaining a sample of people or units that are most convenient to study.	Selecting intersections with available collision data	Low Cost; Easy method of sample design.	No representative sample; Not recommended for descriptive or casual studies.
Judgment	Selecting a sample based on individual judgment about the desirable characteristics required of the sampling units.	Selecting signalized intersections because of experience or intuition that they have higher pedestrian flow.	Low cost; Allow to draw some conclusions about the characteristics of the selected sample.	Does not allow drawing general conclusions about the entire population.
Quota	It is similar to the judgment sample, but requires that the various subgroups in a population are represented.	Making sure to select some signalized and some unsignalized intersections in a sample.	Low cost; Allow to draw some conclusions about the characteristics of the selected sample.	Does not allow drawing general conclusions about the entire population, or sample subgroups.
Snowball	Additional survey respondents are obtained from information provided by the initial sample of respondents.	Used when surveying individuals about their behaviors (e.g. how much they walk in specific areas)	Some characteristics about the target population can be known	Requires a lot of time and resources; Used only for surveys.

Table 5.2: Probabilistic Sampling Techniques

Probabilistic method	*Definition*	*Example*	*Advantage*	*Disadvantage*
Simple Random	A sampling procedure that ensures each element in the population will have an equal chance of being included in the sample	When there are enough resources; to inquire about the characteristics of the entire population	Simple; Conclusions about the population can be drawn.	Subgroups within the target population may not be represented in the sample; Larger samples are necessary.
Systematic Random	Samples are randomly selected from a list in order, but not every one has an equal chance of being selected.			The sample may not be representative because of the ordering of the original list.
Stratified	Sub-samples are drawn within different strata. Each stratum is composed of samples with similar characteristics.	When representation of all subgroups within a particular sample is necessary.	More efficient sample (variance differs between the strata); Small sampling error between strata; Smaller samples.	May be difficult to determine characteristics of individuals to appropriate classify them in specific strata.
Cluster	Entire groups, not individuals, are selected to participate in the data collection; Simple random sampling is applied to the representative "clusters" to select the clusters in which all members will participate.	When the population is too big or when there is a lack of information about individual sampling units (e.g. all vehicle occupants in the United States)	Efficient for large numbers. Do not need to identify all units. Smaller samples; Less expensive relative to the population size.	Sample may not be as representative as desired; Error may be greater than with other techniques; Pilot studies may be necessary to identify the clusters.
Multi Stage Random	Stratification techniques within the clusters used to refine and improve the sample. Examples of this kind of sampling: National Safety Belt Survey.			Like cluster sampling but more representative within clusters.

* Based on Nassirpour, 2004 and MRUTC, 2005

5.1.3. Determination of sample size

There are many considerations that come into play when determining the sample size, such the level of precision to be achieved, operational constraints, available resources and the chosen sampling technique. The more accurate the desired results, the greater the sample size required. In order to achieve a certain level of precision, the sample size will depend, among other things, on the following factors (Statistics Canada, 2006):

- ✓ *The variability of the characteristics being observed*: If all intersections have the same pedestrian flow, then a volume count in one would be sufficient to estimate the average pedestrian flow for all the intersections. If intersections have very different flows, then a bigger sample is needed to produce a reliable estimate.
- ✓ *The sampling and estimation methods*: Not all sampling and estimation methods have the same level of efficiency. Operational constraints and the unavailability of an adequate frame sometimes mean that the most efficient technique cannot be used. A larger sample size is needed if the method used is inefficient.

Som (1996) points out other important observations about sample size:

- ✓ Estimates of sample size required to obtain measures with a given precision will often be found to be quite large, when derived on the basis of unrestricted simple random sampling;
- ✓ Small samples have proved useful, not only as pilot studies to full-scale surveys, but also providing interim estimates;
- ✓ An organizations with inadequate resources can start from a small sample and with increasing resources build up a fully adequate sample; the Current Population Survey of the U.S.A., for example, started in 1943 with 68 primary areas which were enlarged to the present 449.
- ✓ It is possible to combine smaller monthly or quarterly estimates into yearly estimates, and the yearly estimates into estimates covering longer periods, to provide estimates with acceptable precision.

- ✓ In the interest of true accuracy, it may sometimes be better to conduct a smaller sample with adequate control than try to canvass a much larger sample but with poor quality data.

In this protocol, examples are given on how to estimate the sample size for collecting pedestrian volumes at intersections for different purposes. However, these examples are based on specific scenarios, and if any variable of the scenario is changed the sample size must be recalculated.

5.2. Sampling Intersections in a City

As presented above, the sample design must be based on the research objective, the type of study and the available resources. Therefore, when planning to collect data about pedestrian exposure at intersections, the data needs and goals must be clearly defined. These considerations include: (i) what data items are needed and how they will be used; (ii) the precision level required for estimates; (iii) the format, level of detail, and types of tabulations and outputs; and (iv) when and how frequently users need the data (BTS, 2005).

Once data needs are defined, the existing data collection systems must be reviewed in order to determine whether all or part of the required data are already available, or could be more easily obtained by adding or modifying other data collection systems (BTS, 2005). Sometimes, manual pedestrian counts can be combined with existing motor vehicle counts at little or no additional cost. This has already been achieved with good results in some U.S. communities such as Albuquerque, NM, Baltimore, MD, and Washington, DC (Schneider et al., 2005). Pedestrian counts can also be combined with other initiatives such as general plans, pedestrian plans, or studies (e.g. the National Seat Belt Survey). When it is not possible to obtain the necessary pedestrian exposure data by adding or modifying the existing data collection system, a sample design is needed.

Data collection and analysis occurs after the data collection methodology has been defined. However, in systematic studies where data collection is performed repeatedly, it is necessary to reevaluate the study objectives and methodology each time data is collected, creating a loop in the data collection planning process. This loop ensures changing conditions are reflected in the study design. Figure 5.1 illustrates this process.

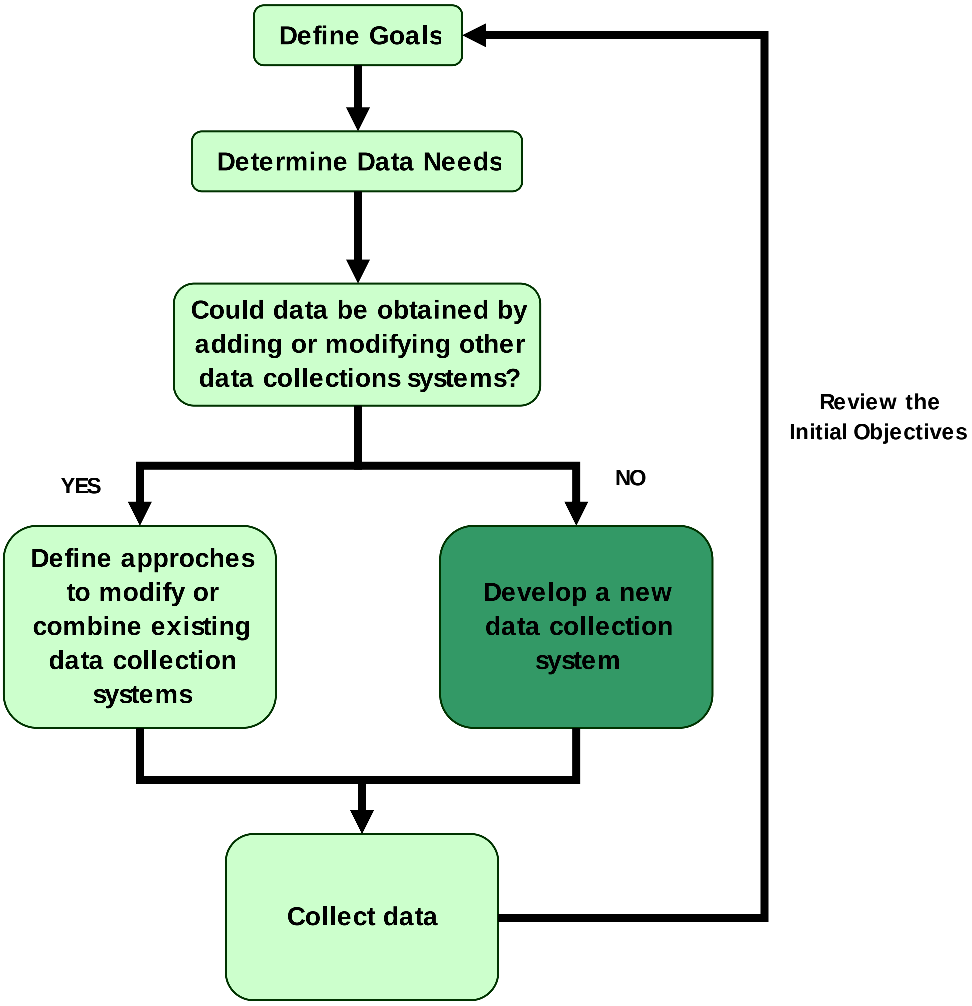


Figure 5.3: Methodology for Planning Pedestrian Exposure Data Collection at Intersections

This chapter focuses on the development of new data collection systems. Three hypothetical scenarios involving the collection of pedestrian exposure data were constructed to illustrate the necessary procedures. These scenarios are intended to be brief sketches of data collection planning. Not all methods and purposes are explored in the scenarios.

To simplify the analysis of the scenarios, we have organized the sampling design in 4 steps, as shown in the Figure 5.4.

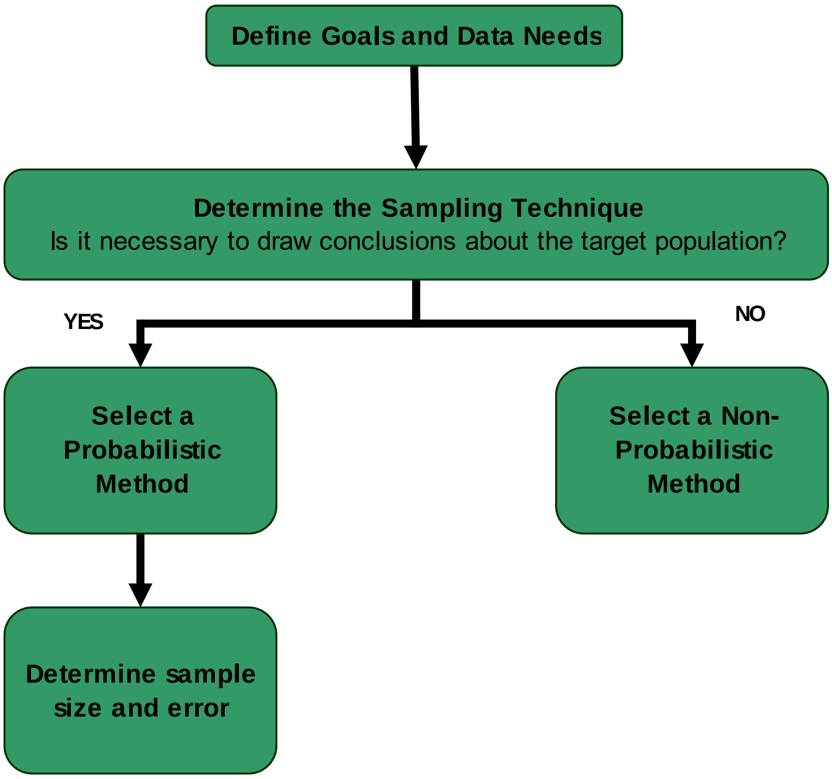


Figure 5.4: Sampling Design Steps for Pedestrian Exposure Data Collection at Intersections

5.2.1. Scenario 1: Evaluate change over time

One of the uses of pedestrian exposure data is to evaluate change over time, such as the change in pedestrian risk in an area or a countermeasure's effectiveness (before-and-after studies, such as Banerjee and Ragland, 2007). In such circumstances, it is common that the researcher is more interested in studying specific sites using non-probabilistic methods to choose where to collect data.

In the first scenario the research goal is to evaluate pedestrian risk among 10 specific intersections before and after signalization. In this case, there is no need to make general inferences about the sample population, and the sites are already chosen using the judgment method (i.e. the intersections that will be signalized). However, the researcher must be aware that when evaluating a temporal series of data it is important to use the same methodologies through time, thus avoiding seasonal influence (Cameron, 1976; Hocherman et al., 1988; Hottenstein et al., 1997).

5.2.2. Scenario 2: Evaluate risk related to infrastructure type

Pedestrian exposure can also be used to compare the safety associated with infrastructure. For example, Zeeger et al. (2005) compared pedestrian risk among marked and unmarked crosswalks. For this purpose, judgment samples or random samples can be used.

The research goal of the second scenario is to determine if pedestrian collision rates at marked mid-block crossings are higher than at unsignalized intersections. The available annual numbers of collisions are aggregated by type of crosswalk in business area of San Francisco. Therefore, the sample frame is marked mid-block crossings and unsignalized intersections in the San Francisco central business district.

To perform the analysis, the annual volume of pedestrians at each type of crossing must be determined. Since the study goal is to understand target population characteristics, a representative sample is needed.

Two random sample sites must be selected: one to determine the annual pedestrian volume at mid-block crossings and one to determine the annual pedestrian volume at unsignalized intersections.

Sites with similar characteristics are expected to have similar pedestrian flows, meaning that the variance in a sample is likely to be relatively low. In this case, a simple random sample technique is appropriate. It is very simple to apply when there is a complete list of all targeted crossings available, and will result in a small sample size when the variance between selected units is low.

Each sample size can be determined by the formula (3).

$$n = \frac{z^2 CV^2}{e^2} \qquad \textbf{(4)}$$

where,
z is the z value, which is derived from the desired confidence level (e.g., 1.645 for 90% confidence level, 1.96 for 95% confidence level, and 2.575 for 99% confidence level);

e is the margin of error (e.g., .07 = + or – 7%, .05 = + or – 5%, and .03 = + or – 3%); and

CV is the coefficient of variance of an attribute in the population (e.g., .10 or 15% for moderate variances).

If a confidence level of 95% (z=1.96) is adopted, with the maximum acceptable error of 5% and a low coefficient of variance (10%) is assumed, the sample size must be 16 crosswalks for each type, totaling 32 intersections. After the first round of data collection the coefficient of variance must be calculated and the sample size must be estimated again, in order to optimize the sample size with a reliable and accurate sample.

The crosswalks must be sampled randomly in each subgroup (mid-block and intersection crossings). It is therefore necessary to have a complete list of all units of the target population classified by subgroup.

5.2.3. Scenario 3: Sampling exposure in a geographic area

Sometimes it is necessary to determine pedestrian exposure in certain area: (i) to compare pedestrian risk between different cities; or (ii) to estimate pedestrian risk for the area. In these cases, a probabilistic approach is necessary to be able to estimate the exposure measure accurately and a stratified sampling technique is most appropriate, since it can provide a sample representative of defined subgroups.

In the third scenario, the main objective is to asses pedestrian risk in the city of Berkeley systematically (the data collection must be repeated every 5 years). The estimate must be representative of the volumes at different types of intersections at different areas. So, a stratified sample must be designed.

Strata must be defined taking into account the similarity of intersection characteristics and geographic sub-areas. One can classify the intersection by type (signalized or non-signalized) or by function (Arterial/Arterial; Arterial/Collector; Arterial/Local; Arterial/Access Ramp; Collector/Collector, Local/Local). There are also many ways to classify geographic areas[2], but in this scenario they are defined in 3 categories: Central Business District; Fringe area; and Suburban and Rural

---

[2] Geographic area classification is explained in greater detail in Chapter 6.

Area. The number of strata will determine the sample size needed, as more strata will require a larger number of samples. For the first year of data collection, it is reasonable to simplify the data collection and use a small number of strata for each stratification variable.

In this scenario, the sample is divided in two stratification variables: (i) intersection type with two classes and geographic area with three classes. Table 5.3 presents these variables, which total six strata (3 x 2). To calculate the number of sites needed within each stratum, the same equation used for scenario 2 can be used (equation 3).

Table 5.3: Stratification Variables

Stratification Variable	*N°. of classes*	*Classes Description*
Intersection Type	2	Signalized Unsignalized
Geographic Area	3	Central Business District Fringe area Suburban and Rural

Adopting a confidence level of 95% (z=1.96), with the maximum acceptable error of 5% and assuming a low coefficient of variance (10%), 15.4 intersections must be selected within each stratum. Therefore, a minimum of 93 intersections (15.4x 6 strata) must be sampled. As in scenario 2, the true coefficient of variance must be calculated and the sample size must be reevaluated after the first round of data collection.

To obtain a more representative sample, we can distribute the total sample size among each stratum proportionally to the target population profile. For example, if in Berkeley 30% of intersections in the central business district are signalized, then 28 intersections with this characteristic must be randomly sampled. However, at least ten units within each stratum should be sampled to maintain statistical reliability.

6. ESTIMATING ANNUAL PEDESTRIAN VOLUMES

In order to determine the annual pedestrian collision risk at a specific site, two pieces of information are needed: the annual number of pedestrian collisions and the annual pedestrian exposure. The numerator of the risk measure, which is the annual number of pedestrian collisions at a site, can be obtained relatively easily from the Statewide Integrated Traffic Records System. The denominator of the risk measure, which is the annual pedestrian exposure at the site, is more difficult to obtain, because it is usually impractical to measure pedestrian volumes continuously for an entire year.

The process of estimating annual pedestrian exposure can be simplified using extrapolation techniques. These techniques allow short samples of pedestrian volume to be converted into a measure of annual pedestrian exposure. The purpose of this chapter is to describe a commonly used method of extrapolating pedestrian volumes and to provide examples of the application of the method.

6.1. Approaches to Estimating Pedestrian Volumes

In theory, the annual pedestrian volume at a site can be obtained by observing and recording pedestrian flow continuously throughout an entire year. In reality, lengthy pedestrian counting periods are impractical because of the time and expense associated with counting (Soot, 1991; Davis et al., 1988; Cove and Clark, 1993; Hocherman et al., 1988).

Various methods of estimating pedestrian volume at a site have been developed in order to reduce the burden of data collection. Some of these strategies do not rely on direct sampling of pedestrian activity, and instead attempt to estimate the activity from land use variables, using similar techniques to the trip-generation methods used to predict vehicle travel (Hottenstein et al., 1997; Otis et al., 1995).

Other strategies rely on extrapolation procedures that convert short pedestrian counts into multi-hourly, daily, or annual estimates of pedestrian flow. There are two main strategies used to achieve the extrapolation of short pedestrian counts. One of these was used by Davis et al. (1988) in Washington, D.C. Pedestrian counts collected at 14 sites over three days were used to develop a set of equations relating short count sample periods of 5, 10, 15 or 30-minutes to expansion periods of 1, 2, 3 and 4 hours. The equations were then validated using data from the remaining sites. It was found that the sample period should be in the middle of the period being sampled, and that the longer the sample period, the more accurate the estimate. The percent error in the estimate ranged from 11.9 percent to 33.6 percent depending on the length of the sampling period.

Although the procedure used by Davis et al. (1988) holds promise, it has some disadvantages. It does not take into account the time of day that the sample was taken, and does not differentiate between different types of sites. It also requires that samples be taken several times during the day in order to obtain a daily estimate.

The second procedure commonly used to extrapolate pedestrian counts involves the development of hourly conversion factors that can be used to expand any hour-long pedestrian count into a daily volume. Because this procedure is relatively simple vis a vis the method used by Davis et al. (1988), and because it takes into account the time of day and the characteristics of the site at which the sample was taken, it has been recommended as a means to extrapolate pedestrian volumes (Soot, 1991).

Moreover, the technique shares some characteristics with the methods of extrapolating short vehicle counts outlined in the Federal Highway Administration’s Traffic Monitoring Guide (FHWA, 2001), which will be discussed below.

The remainder of this section focuses on the second method, which we refer to as the “factoring” method, although it has no specific name in the literature. The factor method involves tracking the temporal and spatial variations in pedestrian volumes in a given area and using them to expand a sample of short pedestrian counts into an annual measure of pedestrian volume

6.2. Temporal and Spatial Variations in Pedestrian Volumes

The factoring method of extrapolating pedestrian counts relies on knowledge or assumptions about how volumes fluctuate at the study site (Soot, 1991). This information is used to create hourly conversion factors that represent each hour’s contribution to the daily flow. For example, if pedestrian flow at a site is perfectly constant, then each hour makes up 1/24, or 4.2 percent, of the day's total.

An hour-long count taken at any site could then be divided by .042 to obtain the daily total. The equation (4) shows the hourly adjustment factor in homogenous pedestrian flow (Zeeger et al., 2005). Similarly, if pedestrian volume were perfectly constant throughout the year, then a day long pedestrian count could be multiplied by 365 to obtain the yearly total.

$$\text{24 Hour Pedestrian Volume} = \text{Hour long count} / 0.042 \qquad (5)$$

The example of homogenous pedestrian flow is useful for illustrative purposes, but does not correspond to reality. Pedestrian volumes are known to fluctuate through time. The pedestrian volume distribution pattern at any given site varies from day to day according to diverse factors such as random variation in weather and day of the week (Hocherman et al., 1988; Hottenstein et al., 1997).

Cameron (1976) found that shopping areas in Seattle, Washington have higher levels of pedestrian activity during the dry summer months, the back-to-school season, and the holiday season, and lower levels during the rainier winter months. On the other hand, areas with little seasonal climate change have little seasonality in pedestrian volume (Hocherman et al., 1988).

In addition to these temporal fluctuations, there are also spatial variations in pedestrian volume. The daily pedestrian volume distribution pattern at one crosswalk may be different from that at a neighboring crosswalk, or in a crosswalk across town. Variations in the volume distribution through space may be produced by land uses surrounding the site (Davis et al., 1988) and the type of pedestrian activity associated with the site (Cameron, 1976).

Although each site is unique, some sites share similar patterns. The unique pattern at a site is sometimes called a “signature” (Soot, 1991). The most comprehensive review of pedestrian volume fluctuation patterns to date was undertaken in 1976 by Cameron. Several hundred days of data were collected, making it possible to track hourly, daily, and seasonal variations in pedestrian volume at each of the sites. It was found that the sites exhibited regular daily and hourly volume fluctuation patterns, and that similar types of sites tended to have similar volume distribution patterns (Cameron, 1976).

Similarities in the pedestrian volume distribution pattern at different sites can be exploited for the purpose of pedestrian volume estimation. Sites which are expected to share a similar pedestrian volume distribution can be treated as a group in order to facilitate the volume estimation process.

If the volume distribution for one site in the group is known, then it can be assumed that all sites in the group share the same distribution pattern. For example, Cameron (1976) classified pedestrian areas by the type of activity at the site: shopper, employee, visitor, mixed, commuter, and special, and identified characteristic pedestrian volume trends for each type of site. Zeeger et al. (2005) grouped sites on the basis of their location in a central business district, residential, or fringe area.

The following section describes how to apply the factoring method using a series of steps. The method involves grouping the sites in an area into strata that share similar pedestrian characteristics, making it similar to the stratified sampling techniques discussed in chapter 5.

6.3. Guide to Estimating Annual Pedestrian Exposure Using the Factoring Method

6.3.1. Select study area

Defining the target area for pedestrian volume monitoring is the first step in performing the factor analysis. Although the analysis can be performed at nearly any geographic scale, it is likely to be most feasible for jurisdictions such as large cities, metropolitan areas, Caltrans Districts, or the state. This is because the procedure requires all-day pedestrian counts, and the time and monetary investments required to collect this data may be harder to justify for small jurisdictions. Larger jurisdictions could achieve a statistical economy of scale by developing adjustment factors applicable to all areas (cities, counties, etc).

However, it is important to be aware of potential tradeoffs between the quality of the results and the size of the study area. One of the sources of error in the calculation of adjustment factors results from differences in the pedestrian volume fluctuation patterns within strata.

Large areas are more likely to contain heterogeneous pedestrian environments that will introduce error into the strata. For example, the city of San Francisco is characterized by mixed land uses, a grid-like street pattern, and high-density development. If one defined three strata within the city (e.g. residential area, employment area, and mixed), one would expect the pedestrian volume fluctuation patterns within these groups to be relatively homogenous, given the consistent character of the urban environment. However, if one defined the same three strata for the entire nine-county San Francisco Bay Area, one would expect a great deal more variation to occur within the strata, and therefore a great deal more error in the resulting volume estimate. Of course, larger jurisdictions may have the resources to account for these variations by selecting and sampling a larger number of strata.

6.3.2. Choose strata (employment center, residential area, mixed/fringe)

As described in the preceding literature review, areas in which the daily pedestrian volume fluctuation pattern is expected or assumed to be homogenous can be grouped into one or more strata. The raw pedestrian volumes at these sites may vary, but similarities in the surrounding land uses, intensity of development, and character of the pedestrian environment create similar temporal variations in pedestrian activity. The strata should be spatially defined, mutually exclusive, and should together equal the study area (Table 6.1). In other words, strata should be defined such that any site in the study area belongs to no more than one strata.

Previous studies have grouped sites by the dominant land use, such as residential, central business district, and fringe area (Zeeger et al., 2005); or by the dominant type of pedestrian at the site, such as shopper, commuter, employee, visitor, and mixed (Cameron, 1976). The ideal selection of strata would account for all the possible sources of variation in activity, and would create a separate stratum for each pedestrian volume fluctuation pattern. The

Table 6.1: Characteristics of Strata

1. Strata are defined by environmental or density variables
2. Sites within each strata are expected to have similar daily pedestrian volume fluctuation patterns
3. Strata are mutually exclusive
4. Sum of strata equal to entire study area

study of Davis et al. (1988), for example, found six unique pedestrian volume variation patterns among fourteen studied sites (Figure 6.1).

It is usually necessary to limit the number of strata groups selected, since each one requires a certain number of samples and is thus associated with a certain cost. This guide proposes three strata, though that number can be increased or decreased depending on the resources available, the desired accuracy of the estimate, and the heterogeneity of the study area.

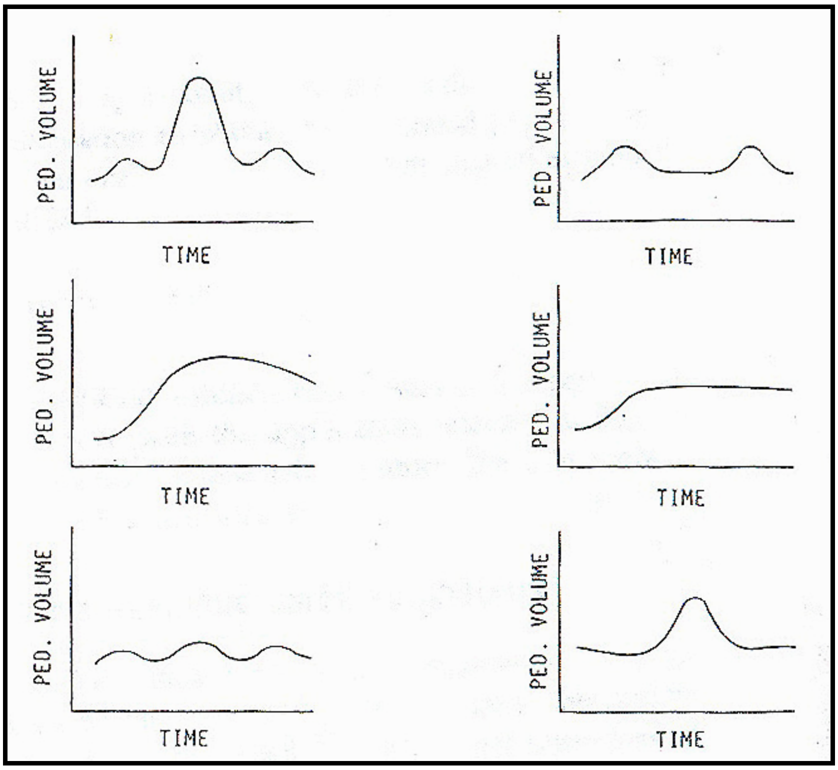


Figure 6.1: 12-hour Pedestrian Volume Distribution Patterns at Sites in Washington, D.C. (Davis et al., 1988)

Although the strata can be defined in a variety of ways, this guide proposes that they be defined in terms of their residential and employment density. The use of residential and employment density has three advantages. First, these data can serve as a simple proxy for more complicated measures of land use mix. Second, these data are readily available for through the U.S. Census Transportation Planning Package. Data may also be drawn from other local or national government sources, such as County Business Patterns data collected by the U.S. Census Bureau. Third, these data can be used to quantitatively define mutually exclusive strata. The definition of strata by land-use type alone (e.g. CBD, residential) is more subjective and is not guaranteed to create groups that are mutually exclusive and sum to the entire study area.

The equation (5) may be used to assign areas to strata on the basis of area density. The formula and list of area types were developed by the Metropolitan Transportation Commission for use in regional transportation demand modeling (MTC, 1997). The original six categories of area type used by MTC are provided, as well as a simplified three-group area type that may be used for this study in Table 6.2.

$$AreaDensity = \frac{P + 2.5E}{AC + AI + AR} \tag{6}$$

where,

P is total resident population within the target area

E is the total employment within the target area

AC is the commercial acreage within the target area

AI is the industrial acreage within the target area

AR is the residential acreage within the target area

Table 6.2: Categories of Area Type

Six-Group MTC Area Type	Simplified Three-Group Area Type
0 Core (Area Density > 300.0) **1 Central Business District (Area Density = 100.0 - 300.0)** **2 Outlying Business District (Area Density = 55.0 - 100.0)** **3 Urban (Area Density = 30.0 - 55.0)** **4 Suburban (Area Density = 6.0 - 30.0)** **5 Rural (Area Density < 6.0)**	1 Central Business District (Area Density > 100.0) 2 Fringe area (Area density = 30.0 – 100.0) 3 Suburban and Rural (Area density = 6 – 30)

6.3.3. *Choose number of factors (hour, day, season, month, year)*

The selection of strata described above reflects the need to account for *spatial* variation in pedestrian volume. This section describes the selection of adjustment factors which account for *temporal* variation in pedestrian activity.

The adjustment factors within a stratum will be used to develop an equation relating a given short count to an estimate of annual pedestrian volume for sites in that stratum. The simplest equation for converting a short count (hourly) into an estimate of annual volume requires a single adjustment factor. This factor must reflect the proportion of the daily volume that the hour makes up in a specific stratum.

The number of adjustment factors required depends on the degree to which the pedestrian volume distribution pattern is expected to change throughout the hour, day, season, month, and year. That is, if the site is located in an area that has significant day-of-week or seasonal variations in pedestrian volume, additional adjustment factors may be necessary to account for those variations. For example, if the short count is taken in a cold month when pedestrian activity is diminished, then simply multiplying the daily estimate by 365 will result in an underestimate of pedestrian activity for the year. A seasonal adjustment factor would help correct for decreases in pedestrian volumes during winter months.

The extent of day-of-week and seasonal variation in pedestrian activity can be estimated by conducting all-day counts of pedestrian activity at a site on several days spread throughout the week and year. The results of such a study could be used to develop adjustment factors that could apply to all the strata, assuming that all strata are similarly affected by day-of-week and seasonal fluctuations in pedestrian volume.

The number of adjustment factors used also depends on resources. Increasing the number of adjustment factors will likely produce a better estimate of annual pedestrian volume, but will require additional sampling to implement. If limited resources make it impossible to develop day-of-week and seasonal adjustment factors, the study can be limited by collecting all counts during a specific time of year (e.g. early fall) and on a specific day (e.g. weekday or weekend day).

6.3.4. Calculate number of day-long counts needed

The number of day-long counts needed within each stratum is a function of the variability of the volume distribution within the stratum. To determine this, a pilot test should be conducted at a sample of sites throughout the study area.

It may occur that there is a great deal of variation in the data collected for each stratum. In this case, the definition of the strata should be examined and possibly readjusted so that each stratum represents, as much as possible, sites within similar pedestrian volume distribution patterns. To facilitate this readjustment, detailed information should be collected on each site sampled during the day-long counts, including the surrounding land uses and type of pedestrian activity.

6.3.5. Collect day-long counts at sites

Day-long counts are collected at sites in the study area in order to determine the daily pedestrian volume fluctuation pattern at each site, which will reflect the daily pattern for all sites in the strata. In theory, it would be ideal to collect day-long counts on every day of the week for the year to determine daily, weekly, and seasonal volume fluctuation patterns for the strata. If this is not possible, then efforts should be made to be consistent in the day chosen for day-long counts. For example, it would be problematic to collect some day-long counts on Friday and others on Tuesday, as the volume distribution pattern will likely differ on each day of the week. Data collection should be avoided on anomalous days of the year, such as holidays, or during times of severe or uncharacteristic weather patterns.

In some cases, lack of automated counting equipment or sufficient resources may make it impractical to collect an entire 24-hour count of pedestrian volume. In these cases, it is advised that 15-hour counts be taken from 7:00 a.m. to 10:00 p.m. Hocherman et al. (1988) found that the period between 10:00 p.m. and 7:00 a.m. represents 3 percent of the daily volume in residential areas and 7 percent of the daily volume in the central business district.

The final result of the data collection should be a table indicating, for each stratum, the mean share of daily volume comprised by each hour in the day, as well as the standard deviation of the sample for each hour.

6.3.6. Develop factor equation

As noted above, the exact form of the factor equation depends on the number of adjustment factors developed during the sampling process. Assuming that only an hourly adjustment factor was developed, the factor equation would yield an average daily volume estimate for a specific day. The factor equation 6 would be used in this case. If a seasonal adjustment factor is developed, then equation 7 can be used (adapted from Hocherman et al., 1988).

$$Aadpv = Cij * K * Di \qquad (7)$$

where

AAdpv = Average daily pedestrian volume for site in strata f

Cij = short-count value in hour i and season j for site in strata f

Di = daily expansion factor for hour I in strata f

K = hourly multiplier: 60/minutes of short count (if less than a one-hour short count is taken)

$$AAdpv = Cij * K * Di * Sj \qquad (8)$$

where

AAdpv = Average daily pedestrian volume in strata f

Cij = short-count value in hour i and season j for site in strata f

K = hourly multiplier: 60/minutes of short count

Di = daily expansion factor for hour I in strata f

Sj = Seasonal correction factor for season j in strata f

6.3.7. Determine optimal length and time period of short count

Although the short count may be taken at any time of day, certain times of day may produce more accurate results. The chosen duration of the short count period will also influence the accuracy of the results and will affect the efficiency of the study.

Length of short count. The optimal length of the short count period is a function of the pedestrian volume at the site and the desired level of accuracy. Haynes (1977) found that the accuracy of a given counting period increases with the volume of pedestrians such that a shorter counting period is required at a high-volume site. A series of curves were developed to aid in the choice of counting period, as shown in Figure 6.2. The curve illustrates that, for example, an hour-long short count does not produce significantly less error than a 40-minute short count in areas with very high pedestrian volume (50 ped / minute).

These curves will be most helpful in urban areas with substantial numbers of pedestrians and will not apply in areas with low numbers of pedestrians. In these areas, it is possible that no pedestrians will be recorded within an hour-long period, resulting in an erroneous average daily pedestrian volume of zero, even if the sample is taken during a peak travel period. To cope with this problem, three possible solutions are proposed:

- ✓ Collect more than an hour of pedestrian volume;
- ✓ Replace the count of zero with a count of .25. This method was used by Zeeger et al. (2005) at sites where an hour-long count produced zero pedestrians. It reflects the fact that pedestrian volume is very low without being zero.
- ✓ Use an alternative method. As noted above, several hours of data may be necessary to develop volume estimates at sites with few pedestrians. When counting pedestrians for several hours is impractical, an alternate method may be required, such as multiple regression techniques (Qin and Ivan, 2001).

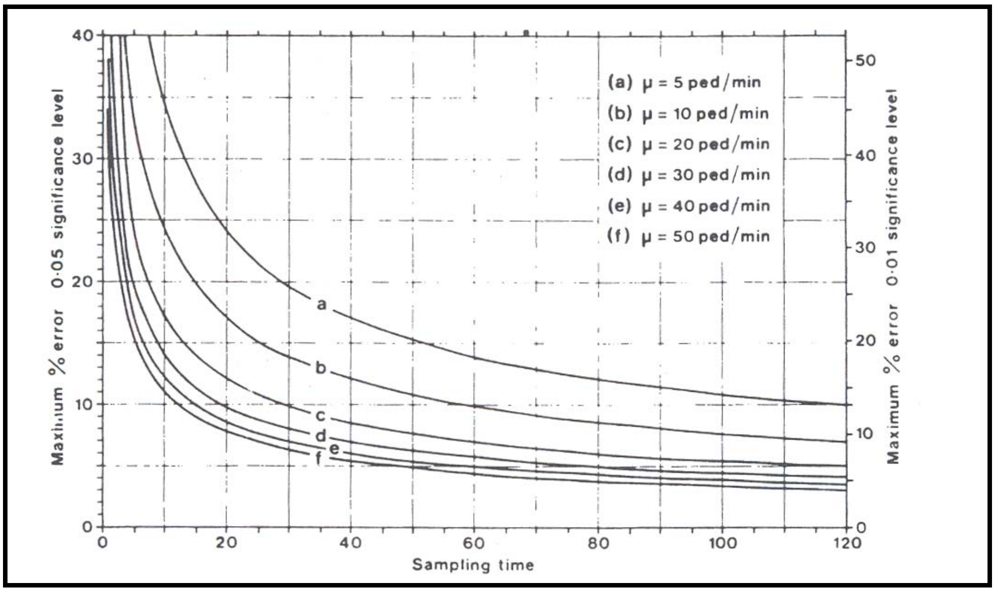


Figure 6.2: Relationship between maximum expected sampling error and sampling time for various levels of pedestrian activity (Haynes, 1977)

Time period of short-count. Three factors should guide the choice of when to sample the short-count at the study site:

- ✓ The expected or known peak hour of pedestrian volume at the site. As noted above, higher pedestrian volumes at a site may reduce the required length of the short count and/or improve the accuracy of the short count. For that reason, there is a benefit to sampling pedestrian volume when volumes are expected to be at their highest.

- ✓ The standard deviation of the hourly adjustment factor. The hourly standard deviations developed for each hourly adjustment factor should be reviewed before sampling short counts. If one or more of the hours was shown to have a high standard deviation, efforts should be made to avoid sampling during that hour, as doing so will produce a less accurate result than sampling during an hour with a lower standard deviation.

- ✓ Sampling schedule. In order to economize resources available for the study, it is important to design a careful sampling schedule. The schedule should minimize the time lost to travel between sample sites. It is also possible to conserve additional time and resources by coordinating the pedestrian volume sampling schedule with vehicle volume sampling schedules (Schneider et al., 2005).

6.3.8. Calculate the error of the estimate

The accuracy of the estimation depends on several factors. Principal among these is the variability of pedestrian volumes at the site. Every real-world site is subject to some random day-to-day variation, but some sites are much more erratic than others. If the flow varies significantly, then a given count is less likely to be representative of the average flow.

Pedestrian volumes in residential areas in Israel were shown to have hourly standard deviation of 2 – 3.5 percent of the daily volume, whereas volumes in central business district were more stable, with a standard deviation of between 1 and 3.5 percent of the daily volume. In addition, pedestrian volumes taken during non-peak periods were shown to be more stable than those taken during peak periods (Hocherman et al., 1988). Thus the problem of random variation in pedestrian volume can be mitigated somewhat by collecting counts during time periods that tend to have less variation in pedestrian volumes, such as non-peak periods.

Error in the factored estimate is also generated by the process of grouping sites on the basis of expected, rather than empirically measured, similarity in the pedestrian volume distribution patterns. Although sites with similar land uses may show similar pedestrian activity, there is likely to be great diversity within the grouping of "central business district", for example. This diversity introduces error into the volume estimate. The amount of error will depend on the extent of diversity within the group. Increasing the number of groups has the potential to decrease the error of volume estimates within each group.

Hocherman et al. (1988) summarized the sources of error in the factoring process with the following equation:

$$Var(aadpv) = K2 \text{ X } f[var(Cij), var(Di), var(Sj) \qquad (9)$$

where:

Var(aadpv) is the variation in the average daily pedestrian volume

K^2 is the square of the hourly adjustment factor

Var(C) is the random day-to-day variation in any given hourly count

Var(D) is the deviation of the daily volume distribution at the location being studied from the volume distribution used to calculate the adjustment factor. It is a function of the homogeneity of sites within the strata

Var(S) is the variation of the seasonality factor used to correct for seasonal variations in pedestrian volume

Another source of error not included this equation is the error that occurs as adjustment factors become outdated. The adjustment factors developed for a group of sites may change from year to year as pedestrian distribution patterns are altered by changing land uses and pedestrian behavior. The extent of this error will depend on the frequency in which adjustment factors are recalculated.

6.3.9. Recalibrate equation

The power of the short-count expansion equation is derived from the assumption that pedestrian activity patterns remain relatively static over time. Over a period of years, however, pedestrian activity patterns will change in response to changing land uses and infrastructure. A site that was once primarily residential may be converted to office uses, for example, resulting in a surge of lunchtime pedestrian activity. Therefore, areas where new or infill development is occurring rapidly should recalibrate more frequently (e.g. every 3 – 5 years) than areas with little development (e.g. 5 –10 years).

6.4. Example Expansion Procedures

This section provides two examples from the literature that used the factoring method described above to estimate pedestrian volumes.

6.4.1. Crosswalk study

The first example comes from a study of 2,000 uncontrolled crossings performed by Zeeger et al. (2005). The crossings were grouped into three types: sites in the central business district (CBD); sites in a fringe area; and sites in a residential area. Sites within each type were assumed to have similar daily pedestrian volume distributions.

Hourly adjustment factors were developed for the three types of sites through the collection of all-day (8- to 12- hour) counts at 22 of the 2,000 sites, as illustrated in Figure 6.3. Counts were not taken during the night time hours (7pm to 7am), but were estimated to represent about 14 percent of the daily total at the site. This estimation was based on the work of Cameron (1976) which found that that the period from 7pm to 7am comprises 14 percent of the 24-hour daily volume at a site. Similarly, Hocherman et al. (1988) found that this period makes up 14.9 percent of the daily volume in residential areas and 18.3 percent of the daily volume in CBD areas.

The pedestrian crossing volume at the remaining 2,000 sites was determined by multiplying a single hour-long count taken at the site by the hourly adjustment factor for that site. Then the daily volume was multiplied by 365 to obtain a yearly volume.

6.4.2. Study of pedestrian volumes in Israel

Hocherman et al. (1988) examined daily pedestrian volume distributions at 72 residential sites and 14 central business district sites in Haifa and Givatayim, Israel, to determine whether the factoring method could be used effectively to extrapolate short pedestrian counts.

It was found that the daily volume distributions at the residential sites were very similar and could be used to calculate an average daily pedestrian distribution at residential sites. The volume distributions at sites in the central business district also showed a clear pattern, with the main differences from residential sites being a smaller morning peak period and a lower hourly variation in pedestrian volume. The authors compared their results with similar distributions in Germany and Australia, and found similarities between the three distributions. Figure 6.4 shows the results of the comparison between the pedestrian volume distributions in these three countries.

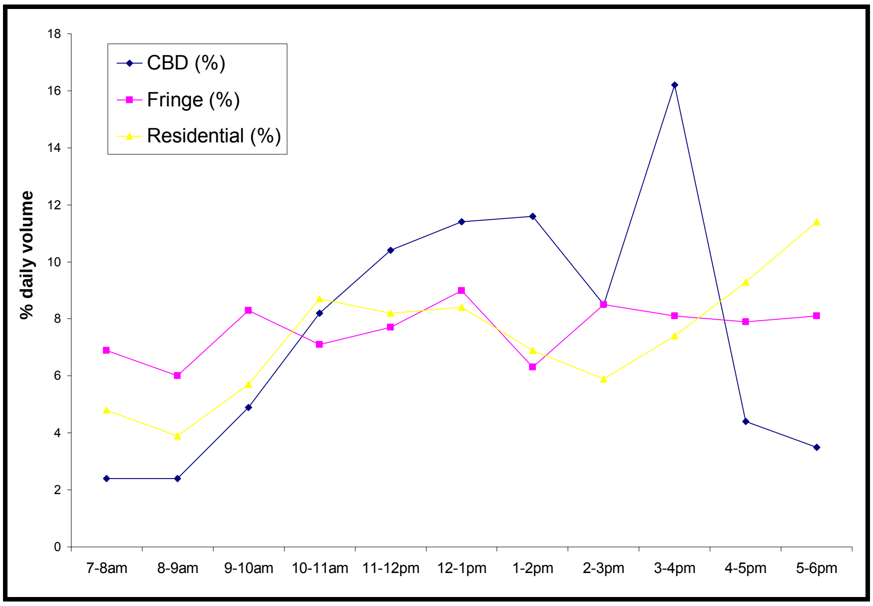


Figure 6.3: Daily volume adjustment factors developed for CBD, Fringe, and Residential Sites (Zeeger et al., 2005)

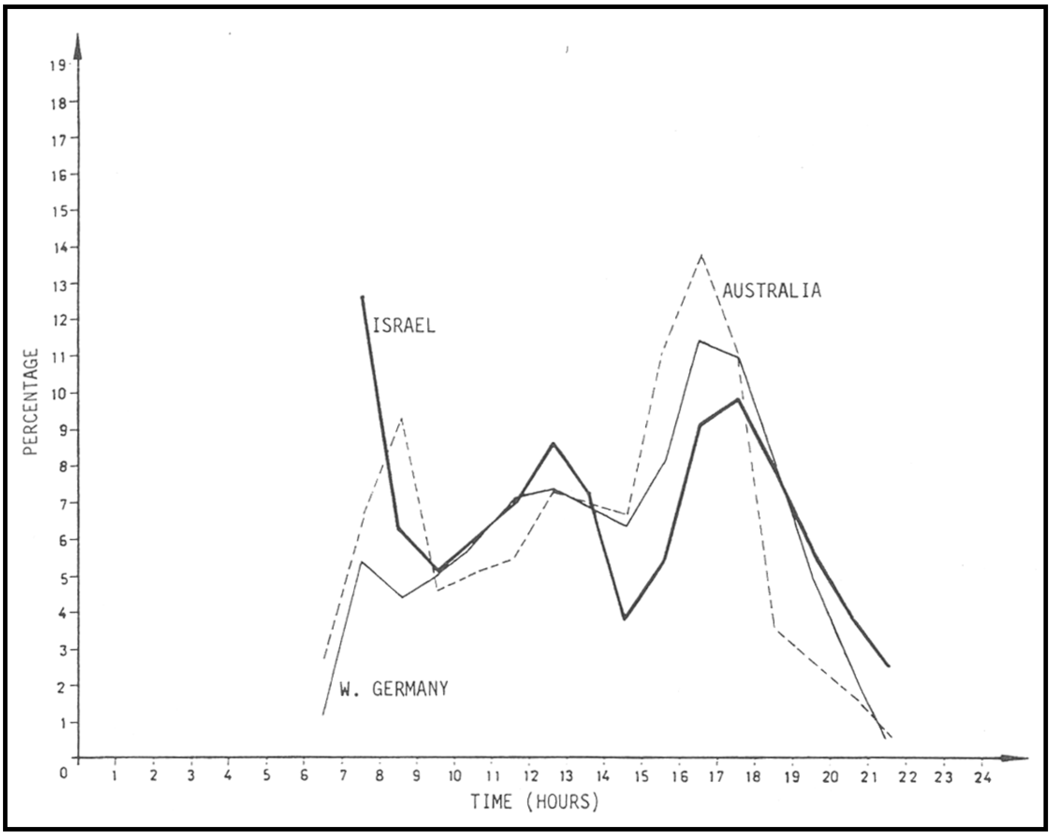


Figure 6.4: Comparison of daily pedestrian crossing volume distributions in Israel, Germany, and Australia (Hocherman et al., 1988)

6.5. FHWA Traffic Monitoring Guide

Although the volume monitoring procedures described in the Traffic Monitoring Guide (FHWA, 2001) involve vehicle volumes only, they employ the factoring method. The methods in the TMG are basically similar to the expansion methods described above, in that they rely on the development of factors to be applied to groups of similar roadways. However, the existence of readily available continuous counting devices makes the vehicle volume estimation process more statistically robust than the pedestrian volume procedures described above.

These devices, also known as Automatic Traffic Recorders (ATRs) are capable of recording volume fluctuation patterns continuously over a period of years. Pedestrian volumes, by contrast, are rarely collected for more than a period of hours or days at a time.

ATRs are typically placed in many locations throughout a state and are used in the development of time-of-day, day-of-week, and seasonal adjustment factors. The ATRs are then matched with groups of roadways on the basis of empirically measured similarities or by expected similarity on the basis of similar functional class or roadway type. The adjustment factors developed for a given group are used to convert short counts, usually of 48 hours or more, into measures of average annual daily traffic.

REFERENCES

Aggarwal, Y. P. **Better Sampling: Concepts, Techniques and Evaluation.** Sterling Publishers Private Limited, 1988.

Alta Planning and Design. **National Bicycle and Pedestrian Document Project**. Alta Planning and Design, Inc. www.altaplanning.com/. Accessed May 21, 2006.

Anderson, P.R., H.J. Montesin, and M.A. Adena. **Road Fatality Rates in Australia 1984-1985**. Federal Office of Road Safety, Canberra, Australia, 1989.

Banerjee, I. and D.R. Ragland. Evaluation of Countermeasures: A Study on the Effect of Impactable Yield Signs Installed at Four Intersections in San Francisco. In **TRB 86th Annual Meeting Compendium of Papers CD-ROM.** Transportation Research Board, 2007.

Behavioral Risk Factor Surveillance System (BRFSS). National Center for Chronic Disease Prevention and Health Promotion. **www.cdc.gov/brfss/about.htm.** Accessed June 30, 2006.

Ben-Akiva, M. and S. Lerman. **Discrete Choice Analysis: Theory and Application to Travel Demand.** The MIT Press, Cambridge, MA, 1985.

Bureau of Transportation Statistics (BTS) **Guide to Good Statistical Practice in the Transportation Field.** U.S. Department of Transportation, Bureau of Transportation Statistics, 2003.

Bureau of Transportation Statistics (BTS) **BTS Statistical Standards Manual.** U.S. Department of Transportation, Bureau of Transportation Statistics, 2005.

California Department of Transportation (Caltrans). **California Statewide Household Travel Survey**, **2000-2001.** Final Report. California Department of Transportation, Division of Transportation System Information 2002.

California Department of Transportation. **California Motor Vehicle Stock, Travel and Fuel Forecast**. California Department of Transportation, Division of Transportation System Information, 2005.

California Highway Patrol. **Annual Report of Fatal and Injury Motor Vehicle Traffic Collisions.** 2004.

Cameron, R.M. Pedestrian Volume Characteristics. Institute of Transportation Engineers **Compendium of Technical Papers**, 1976.

Cameron, M.H. A Method of Measuring Exposure to Pedestrian Accident Risk. **Accident Analysis and Prevention**, Vol. 14, 1982, pp 397-405.

Campbell, B.J., C.V. Zegeer, H.H. Huang and M. J. Cynecki. **A Review of Pedestrian Safety Research in the United States and Abroad.** Publication FHWA-RD-03-042. Office of Safety Research and Development, Federal Highway Administration 2004.

Central London Partnership (CLP) **Automatic Pedestrian Counting Trial**. Stage 3 - Final Report. Central London Partnership, 2005.

Chan, C., F. Bu, and S. Shladover. **Experimental Vehicle Platform for Pedestrian Detection**. California PATH Research Report UCB-ITS-PRR-2006-16, 2006.

Chu, X. The Fatality Risk of Walking in America: A Time-Based Comparative Approach. **Walk21 Conference**: **Health, Equity and the Environment**, Portland, Oregon, 2003.

Chu, X. Measuring the Risk of Walking. Lecture. XVIIth Entretiens Jacques Cartier, EJC Conference, **Urban Road and Public Transit: Who must pay?**, Montreal, 2004.

Cove, L.L. and J.E. Clark. Estimating Pedestrian Volumes on College Campuses. **Transportation Research Record** 1405, Transportation Research Board, 1993, pp. 43-48.

Davis, S.E., L.E. King, and D.H. Robertson. Predicting Pedestrian Crosswalk Volumes. **Transportation Research Record** 1168, Transportation Research Board, 1988, pp 25-30.

Dharmaraju, R., D.A. Noyce, and J.D. Lehman. An Evaluation of Technologies for Automated Detection and Classification of Pedestrians and Bicycles. In **The 71st ITE Annual Meeting. Compendium of Technical Papers**. Institute of Transportation Engineering, Chicago, IL, 2001.

Diogenes, M.C., R.Greene-Roesel, L. Arnold, and D.R. Ragland. Pedestrians Counting Methods at Intersections: A Comparative Study. In **TRB 86th Annual Meeting Compendium of Papers CD-ROM.** Transportation Research Board, 2007.

European Transport Safety Council (ETSC). **Exposure Data for Travel Risk Assessment: Current Practice and Future Needs in the EU**. European Transport Safety Council, Brussels, Belgium, 1999.

Evans, L. **Traffic Safety and the Driver**. Van Nostrand Reinhold, New York, 1991.

Federal Highway Administration (FHWA). **Traffic Monitoring Guide**. Publication FHWA-PL-01-021. U.S. Department of Transportation, Federal Highway Administration, Office of Highway Policy Information, 2001.

Fehr and Peers Transportation Consultants. **City of Sacramento Pedestrian Master Plan: Making Sacramento the Walking Capital** (Public Review Draft) City of Sacramento: Sacramento, 2005.

Fitzpatrick, K., Turner, S. Brewer, M. Carlson, P., Ullman, B. Trout, N. Park, E.S., Whitacre, J., Lalani, N. and Lord, D. **Transit Cooperative Research Program Report 112/ National Cooperative Highway Research Program Report 562**, “Improving Pedestrian Safety at Unsignalized Crossings.” Washington, DC: Transportation Research Board, 2006.

Hakamies-Blomqvist, L. Older Drivers’ Accident Risk: Conceptual and Methodological Issues. **Accident Analysis and Prevention**, Vol. 30, 1998, pp 293-297.

Hakkert, A.S., and L.Braimaister. **The Uses of Exposure and Risk in Road Safety Studies**. Publication R-2002-12. SWOV Institute for Road Safety Research.. Leidschendam, the Netherlands, 2002

Hauer, E. Traffic Conflicts and Exposure. **Accident Analysis and Prevention**, Vol. 14, No 5, 1982, pp. 359-364

Hauer, E. On Exposure and Accident Rate. **Traffic Engineering and Control**, Vol 36, No. 3, 1995, pp. 134-138.

Hauer, E. Computing and Interpreting Accident Rates for Vehicle Types or Driver Groups. **Transportation Research Record**, Vol. 1746, 2001, pp 69-73.

Haynes, R. Sampling Time and Sampling Error in Pedestrian Counts. **Traffic Engineering Control**, Vol. 18, No. 2, 1997, pp. 72-74.

Herms, B.F. **Pedestrian Crosswalk Study: Accidents in Painted and Unpainted Crosswalks**. Police Department, Traffic Bureau, San Diego, 1970.

Hocherman, I., A.S. Hakkert and J. Bar-Ziv. Estimating the Daily Volume of Crossing Pedestrians from Short-Counts. **Transportation Research Record** 1168, Transportation Research Board, 1988, pp. 31-38.

Høj, N.P. and W. Kröger. Risk Analyses of Transportation on Road and Railway from a European Perspective. **Safety Science**, Vol. 40, No. 1-4, 2002, pp. 337-357.

Hottenstein, A., S. Turner and G. Shunk. **Bicycle and Pedestrian Travel Demand Forecasting: Summary of Data Collection Activities.** Publication FHWA/TX-98/1723-1. Texas Transportation Institute, Federal Highway Administration, 1997.

Hunter, W., J. Stutts, W. Pein and C. Cox. **Pedestrian and Bicycle Crash Types of the Early 1990s**. Publication FHWA-RD-95-163, U.S Department of Transportation, Federal Highway Administration Office of Traffic Safety and Operations, 1996.

Ivan, J. N., P. J. Ossenbruggen, X. Qin and Jyothi Pendarkar. **Rural Pedestrian Crash Rate: Alternative Measures of Exposure**. Project UCNR 11-10. United State Department of Transportation. Region I University Transportation Center, 2000.

Jacobsen, P. L. Safety in Numbers: More Walkers and Bicyclists, Safer Walking and Bicycling. **Injury Prevention.** Vol 9, 2003, pp 205-209.

Jonah, B.A. and G.R. Engel. Measuring the Relative Risk of Pedestrian Accidents. **Accident Analysis and Prevention**, Vol. 5, No. 3, 1983, pp.193-206.

Jorgensen, N.O. The Risk of Injury and Accident by Different Travel Modes. **Proceedings of the International Conference on Passenger Safety in European Public Transport**. Transport Safety Council, Brussels, 1996, pp. 17-25.

Keall, M.D. Pedestrian Exposure to Risk of Road Accident in New Zealand. **Accident Analysis and Prevention**, Vol. 27, No. 3, 1995, pp. 729-740

Knoblauch, R.L., H.N. Tobey, and E.M. Shunaman. Pedestrian Characteristics and Exposure Measures. **Transportation Research Record**, Vol. 959, 1984, pp 35-41.

Knoblauch, R. L., M. T. Pietrucha, and M. Nitzburg. Field Studies of Pedestrian Walking Speed and Start-Up Time. **Transportation Research Record,** No. 1538, Pedestrian and Bicycle Research, 1996, pp 27-38

Komanoff, C., and C. Roelofs. **The Environmental Benefits of Bicycling and Walking. National Bicycling and Walking Case Study No. 15.** Publication FHWA-PD-93-015. FHWA, U.S. Department of Transportation, 1993.

Kraus, J.F., E.G. Hooten, K.A. Brown, C. Peek-Asa, C. Heye and D.L. McArthur. Child Pedestrian and Bicyclist Injuries: Results of Community Surveillance and a Case-Control Study. **Injury Prevention**, Vol. 2, 1996, pp. 212-218.

Lee, C. and M. Abdel-Aty. Comprehensive analysis of vehicle-pedestrian crashes at intersections in Florida. **Accident Analysis and Prevention**, Vol 37, 2005, pp. 775-786.

McNally, M.G. The Four-Step Model. In D.A. Hensher and K.J. Button, Ed. **Handbook of Transport Modelling. Pergamon**, Oxford, United Kingdom, 2000, pp. 35-52.

Metropolitan Transportation Commission (MTC) **Travel Demand Models for the San Francisco Bay Area (BAYCAST-90)**. Technical Summary. Metropolitan Transportation Commission, 1997

Midwest Regional University Transportation Center (MRUTC) **Development of a Guide to Statistics for Maintenance Quality Assurance Programs in Transportation**. Research Progress Report for the Quarter Ending: June 30, 2005.

Mitman, M. and D.R. Ragland. Toward Practice-Wide Standards for Pedestrian Data Collection: An Evaluation of Pedestrian Behavioral Data Collection Methods for Multi-Lane Roads with Low Pedestrian Volumes In **TRB 86th Annual Meeting Compendium of Papers CD-ROM.** Transportation Research Board, 2007.

Nassirpour, M. Overview of Sample Surveys (Presentation). Applied Sampling Workshop at the **Annual TRB Conference**, Washington, D.C. , 2004.

National Center for Statistics and Analysis (NHTSA). **Traffic Safety Facts: 2004 Data**. National Center for Statistics and Analysis, Washington, D.C., 2004.

National Household Travel Survey (NHTS). Bureau of Transportation Statistics, U.S. Department of Transportation. **nhts.ornl.gov/2001/index.shtml**. 2001 Accessed May 30, 2006.

National Household Travel Survey (NHTS). Bureau of Transportation Statistics, U.S. Department of Transportation. **[Website] http://nhts.ornl.gov/.** Accessed December 17, 2006.

Noyce, D. and R. Dharmaraju. **An Evaluation of Technologies for Automated Detection and Classification of Pedestrians and Bicyclists.** Massachusetts Highway Department report, 2002.

Noyce, D., A. Gajendranand and R. Dharmaraju. Development of a Bicycle and Pedestrian Detection and Classification Algorithm for Active-Infrared Overhead Vehicle Imaging Sensors. In **Transportation Research Board 85th Annual Meeting**. CD-ROM. Transportation Research Board, Washington , D.C., 2006

Ogden, K.W. The Effects of Paved Shoulders on Accidents on Rural Highways. **Accident Analysis and Prevention,** Vol. 29, 1997, pp. 353-362.

Ossenbruggen, P.J. The Impacts of a Safety Compliance in Highway Design. **Risk: Health, Safety and Environment**, Vol. 10, 1999, pp. 359–379.

Otis, S. C., R. B.Machemehl and H. S. Mahmassani. Predicting Pedestrian Volumes Based on Land Use: a Methodology. **Fifth National Conference on Transportation Planning Methods Applications**. A Compendium of Papers, Vol. 1. Transportation Research Board, 1995.

Pucher, J. and J.L. Renne. Socioeconomics of Urban Travel: Evidence from the 200NHTS. **Transportation Quarterly**, Vol. 7, No. 3, Summer 2003, pp. 49-77.

Pucher, John, and Lewis Dijkstra. Making Walking and Cycling Safer: Lessons from Europe.**Transportation Quarterly**, Vol. 54, 2000, pp 25-50.

Pucher, J. and L. Dijkstra Promoting Safe Walking and Cycling to Improve Public Health: Lessons from the Netherlands and Germany. **American Journal of Public Health**, Vol. 93, 2003, pp 1509-1516.

Qin, X. and J.N. Ivan. Estimating Pedestrian Exposure Prediction Model in Rural Areas. **Transportation Research Record 1773**, Transportation Research Board, 2001, pp 89-96.

Raford, N. and D.R. Ragland. Pedestrian Volume Modeling for Traffic Safety and Exposure Analysis: Case of Boston, Massachusetts. In **Transportation Research Board 85th Annual Meeting**. CD-ROM. Transportation Research Board, Washington , D.C., 2006

Raford, N. and D.R. Ragland. Space Syntax: Innovative Pedestrian Volume Modeling Tool for Pedestrian Safety. In **Transportation Research Record: Journal of the Transportation Research Board**, No. 1878, TRB, National Research Council, Washington, D.C., 2004, pp. 66-74.

Roberts, I, R. Norton and B. Taura. Child Pedestrian Exposure Rates: The Importance of Exposure to Risk Relating to Socioeconomic and Ethnic Differences in Auckland, New Zealand. **Journal of Epidemiology and Community Health**, Vol. 50, 1996, pp 162-165.

Schneider, R., R. Patton, J. Toole and C. Raborn. **Pedestrian and Bicycle Data Collection in United States Communities: Quantifying Use, Surveying Users, and Documenting Facility Extent**. FHWA, Office of Natural and Human Environment, U.S. Department of Transportation, Washington, D.C., 2005.

Schwartz, W. and C. Porter. **Bicycle and Pedestrian Data: Sources, Needs, and Gaps.** Publication BTS00-02. Bureau of Transportation Statistics. U.S. Department of Transportation, Washington, D.C., 2000.

Schwartz, W.L., C. D. Porter, G. C. Payne, J. H. Suhrbier, P. C. Moe and W. L. Wilkinson III. **Guidebook on Methods to Estimate Non-Motorized Travel: Overview of Methods**. Publication FHWA-RD-98-165. FHWA, U.S. Department of Transportation ,1999.

Schweizer, T. Methods for counting pedestrians. In T**he 6th International Conference on Walking in the 21st Century**. Walk21-VI "Everyday Walking Culture". Zurich, Switzerland, 2005.

Senevarante, P. and J. Morall. Analysis of Factors Affecting the Choice of Route of Pedestrians. **Transportation Planning and Technology**, Vol. 10, No. 1, 1986, pp. 147-159.

Shapiro, G., D. Utter, W. Mix, J. Croos, D. Marker, and N. Bondy. Sample Design Issues for Surveys Involving the Observation of Drivers. **Proceedings of the Annual Meeting of the American Statistical Association**, 2001.

Som, R. K. **Pratical Sampling Techniques**. Second Edition, Revised and Expanded. Marcel Dekker, Inc, New York. 635 p. 1996.Som (1996)

Soot, S. Trends in Downtown Pedestrian Traffic and Methods of Estimating Daily Volumes. **Transportation Research Record 1325**, Transportation Research Board, 1991, pp 75-82.

Statistics Canada. **http://www.statcan.ca/start.html** Accessed November 10, 2006.

STPP. **Mean Streets 2004: How Far Have We Come?** Surface Transportation Policy Project, Washington, D.C., 2004.

STPP. **Mean Streets**. Surface Transportation Policy Project, Washington, D.C. 2002.

Transport Research Laboratory (TRL). **Cost-Benefit Analysis of Measures for Vulnerable Road Users**. Deliverable 5 for European Research Project PROMISING. SWOV Institute for Road Safety Research, The Netherlands, 2001.

Transportation Research Board's (TRB) Committee on Urban Transportation Data and Information Systems (ABJ30). **Household Travel Surveys on the WWW** trb.mtc.ca.gov/urban/hhsurvey.html. Accessed July 17, 2006.

United States Census Bureau, **American Community Survey.** 2004. Tables generated using American FactFinder. factfinder.census.gov. Accessed July 17,2006.

United States Census Bureau. **American Community Survey 2005 Questionnaire.** 2005. http://www.census.gov/acs/www/Downloads/SQuest05.pdf. Accessed May 30, 2006.

United States Census Bureau. **American Community Survey Basics.** 2006 www.census.gov/acs/www/SBasics/index.htm. Accessed May 30, 2006.

Zeeger, C.V., K.S. Opiela, and J.J. Cynecki. Effect of Pedestrian Signals and Signal Timing on Pedestrian Accidents. **Transportation Research Record** 847, TRB National Research Council, Washington, DC, 1982, pp. 62-72

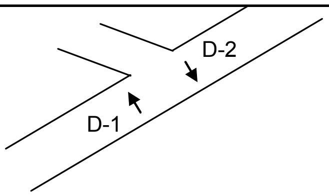
Zeeger, C.V., R. Stewart, H. Huang, P.A. Lagerwey, J. Feaganes and B.J. Campbell. **Safety Effects of Marked versus Unmarked Crosswalks at Uncontrolled Locations: Final Report and Recommended Guidelines**. Publication FHWA–HRT–04–100. Office of Safety Research and Development, Federal Highway Administration, 2005.

APPENDIX A: Example of a Tally Sheet Used to Count Pedestrian

Intersection: ____________________

Data Collected by: ____________________

Data Collected on: ____________________



Period:
( ) 1:00 to 1:30 pm ( ) 1:31 to 2:00 pm ( ) 2:01 to 2:30 pm ( ) 2:31 to 3:00 pm
( ) 3:01 to 3:30 pm ( ) 4:00 to 4:30 pm ( ) 4:31 to 5:00 pm

LEGEND:

Age options:
1 - =<12
2 - 13-18
3 - 19-25
4 - 26-35
5 - 36-50
6 - 51-64
7 - 65+

Direction options:
1
2

Gender options:
M - Male
F - Female

PED #	DIRECTION	AGE	GENDER	PED #	DIRECTION	AGE	GENDER	PED #	DIRECTION	AGE	GENDER
1		( ) =<12		6		( ) =<12		1			
	( ) 1	( ) 13-18	( ) Male		( ) 1	( ) 13-18	( ) Male	2			
		( ) 19-25				( ) 19-25		3			
		( ) 26-35				( ) 26-35		4			
		( ) 36-50				( ) 36-50		5			
	( ) 2	( ) 51-64	( ) Female		( ) 2	( ) 51-64	( ) Female	6			
		( ) 65+				( ) 65+		7			
2		( ) =<12		7		( ) =<12		8			
	( ) 1	( ) 13-18	( ) Male		( ) 1	( ) 13-18	( ) Male	9			
		( ) 19-25				( ) 19-25		10			
		( ) 26-35				( ) 26-35		11			
		( ) 36-50				( ) 36-50		12			
	( ) 2	( ) 51-64	( ) Female		( ) 2	( ) 51-64	( ) Female	13			
		( ) 65+				( ) 65+		14			
3		( ) =<12		8		( ) =<12		15			
	( ) 1	( ) 13-18	( ) Male		( ) 1	( ) 13-18	( ) Male	16			
		( ) 19-25				( ) 19-25		17			
		( ) 26-35				( ) 26-35		18			
		( ) 36-50				( ) 36-50		19			
	( ) 2	( ) 51-64	( ) Female		( ) 2	( ) 51-64	( ) Female	20			
		( ) 65+				( ) 65+		21			
4		( ) =<12		9		( ) =<12		22			
	( ) 1	( ) 13-18	( ) Male		( ) 1	( ) 13-18	( ) Male	23			
		( ) 19-25				( ) 19-25		24			
		( ) 26-35				( ) 26-35		25			
		( ) 36-50				( ) 36-50		26			
	( ) 2	( ) 51-64	( ) Female		( ) 2	( ) 51-64	( ) Female	27			
		( ) 65+				( ) 65+		28			
5		( ) =<12		10		( ) =<12		29			
	( ) 1	( ) 13-18	( ) Male		( ) 1	( ) 13-18	( ) Male	30			
		( ) 19-25				( ) 19-25		31			
		( ) 26-35				( ) 26-35		32			
		( ) 36-50				( ) 36-50		33			
	( ) 2	( ) 51-64	( ) Female		( ) 2	( ) 51-64	( ) Female	34			
		( ) 65+				( ) 65+		35			

APPENDIX B: Comparative Study between Manual Count Methods

Pedestrian Counting Methods at Intersections: a Comparative Study

Submission Date: August 1, 2006

Word Count: 2976

Number of Figures and Tables: 6

Authors:

Mara Chagas Diogenes
Traffic Safety Center
University of California, Berkeley
140 Warren Hall #7360
Berkeley, CA 94709
510-643-7625
maracd@berkeley.edu

Ryan Greene-Roesel
Traffic Safety Center
University of California, Berkeley
140 Warren Hall #7360
Berkeley, CA 94709
510-643-7625
ryangr@berkeley.edu

Lindsay S. Arnold
Traffic Safety Center
School of Public Health
University of California, Berkeley
140 Warren Hall
Berkeley, CA 94720
510-282-5896
larnold@berkeley.edu

David R. Ragland (corresponding author)
Traffic Safety Center
University of California, Berkeley
140 Warren Hall #7360
Berkeley, CA 94709
510-642-0655
510-643-9922 (Fax)
davidr@berkeley.edu

ABSTRACT

Resources for implementing countermeasures to reduce pedestrian collisions in urban centers are usually allocated on the basis of need, which is determined by risk studies. They commonly rely on pedestrian volumes at intersections. The methods used to estimate pedestrian volumes include direct counts and surveys, but few studies have addressed the accuracy of these methods. This paper investigates the accuracy of three common counting methods: manual counts using sheets, manual counts using clickers, and manual counts using video cameras. The counts took place in San Francisco. For the analysis, the video image counts, with recordings made at the same time as the clicker and sheet counts, were assumed to represent actual pedestrian volume. The results indicate that manual counts with either sheets or clickers systematically underestimated pedestrian volumes. The error rates range from 8-25%. Additionally, the error rate was greater at the beginning and end of the observation period, possibly resulting from the observer's lack of familiarity with the tasks or fatigue.

INTRODUCTION

Road collisions are a major public health concern throughout the world. It is estimated that 1.2 million traffic fatalities occur each year worldwide. The problem is especially acute for pedestrians, who face a significantly greater risk of death when involved in traffic collisions than do vehicle occupants *(1)*. Significant resources are focused on countermeasures that aim to reduce the risk of pedestrian injury. Because resources are limited, risk analysis is necessary to develop cost-effective countermeasures *(2)*.

Risk is defined as the frequency of an undesired event or collision per unit of exposure. Pedestrian volume is the exposure measure most frequently used in risk analysis. According to Gårder *(3)*, pedestrian risk should be calculated as a function of pedestrian volume, not just vehicle volume. Although many state, regional, and local agencies have developed methodologies to collect pedestrian volume data, there is no consensus on which method is best *(4, 5)*. To improve the risk monitoring process, it is necessary to define a systematic pedestrian counting method.

The two most frequent types of pedestrian counting methods are direct counts and surveys. Direct counts involve direct observation of pedestrian activity at fixed locations, such as crosswalks or intersections. Surveys indirectly capture pedestrian activity in a geographic area by gathering travel data from a sample *(6)*.

Pedestrian volumes at intersections are usually collected directly using either (i) manual counts, taken by collectors in the field, or (ii) automated counts using specialized equipment. Although motorized vehicles are commonly counted with automated devices, the technology for counting non-motorized modes of transportation, especially pedestrians, is not very developed *(7)*.

The accuracy of these counting methods directly affects the accuracy of the exposure estimate and thus the value of the risk analysis at an intersection. However, few studies have attempted to compare the accuracy of different counting methods. This paper aims to compare the accuracy of three common pedestrian counting methods: (i) manual counts using sheets; (ii) manual counts using clickers; and (iii) manual counts using video cameras.

METHODS

The research was conducted at 10 different intersections in the city of San Francisco, California, during the last two weeks of April and the first week of May, 2006. Field observers collected pedestrian counts with either sheets or manual clickers. Counts were taken for four hours between 1:00 pm and 6:00 pm, with a break of one hour. Video footage of the intersection was recorded simultaneously with the field counts.

Two persons were contracted from a private consulting firm specializing in data collection. One individual made the field observations, and the other operated the video recorder. The contracted staff was the same for all data collection. Sheets were used at eight intersections and clickers at two intersections. The selected intersections had different pedestrian flows, with values varying between 12 and 262 pedestrian crossings per hour based on the video analyses, as shown in Table 1. Figures 1 and 2 present the camera angles used at two of the study intersections.

TABLE 1 Data Collection Schedule and Pedestrian Flow

Intersection	Date	Method	Volume (ped)	Period (hours)	Flow (ped/hour)
France and Mission St.	04/17/2006	Manual with sheets	128	4	32
Admiral Ave. and Mission St.	04/18/2006	Manual with sheets	49	4	12
16th St. and Capp	04/19/2006	Manual with sheets	412	4	103
Geneva and Mission St.	04/20/2006	Manual with sheets	1046	4	262
Folson and 7th St.	04/21/2006	Manual with sheets	334	4	84
Harrison and 7th St.	04/24/2006	Manual with sheets	651	4	163
Market and Castro	04/25/2006	Manual with sheets	579	4	145
Market and Noe	04/26/2006	Manual with sheets	994	4	249
Harrison and 10th St.	05/03/2006	Manual with clickers	161	4	40
Santa Rosa and Mission St.	05/05/2006	Manual with clickers	338	4	85

Before the start of data collection, the researchers supplied the field staff with the following directions:

1. The data collection must be synchronized with the video. The person collecting the data should begin to count the pedestrians when the video begins to run. During the period that the tape is being changed, the observer should stop counting.
2. The field observer must note any problem or interruption in the data collection, such as a break or lack of attention for any reason. These interruptions are important since the main objective was to compare the accuracy of the methods.

3. The field observer must count only pedestrians who cross the street centerline (e.g. the middle of the crossing). He or she should not count bicyclists unless they are walking their bicycle across the intersection.
4. The field observer must stand close to the crosswalk.

Field data were entered into a Microsoft Access 2000 database. For quality control, all database tables were compared with the original field data sheets.

FIGURE 1 Camera angle used at Admiral Ave. and Mission St.

FIGURE 2 Camera angle used at Market and Castro (still from video tape)

Manual with sheets

The field observer received a sheet with three fields: (i) direction of travel; (ii) pedestrian gender; and (iii) age. The observer was instructed to use his best judgment to assign the pedestrian to one of seven age categories.

At the top of the sheet, the observer was instructed to write the following information: (i) name of the intersection; (ii) his/her name; (iii) date of the data collection; and (iv) period of the data collection (check box) – divided in periods of 30 minutes. The field observer was told to concentrate on accurately counting the number of pedestrians, even if it meant leaving gender and age fields blank in crowded intersections.

To improve the analysis, after the fourth day (April 20), the field observer was asked, when possible, to take note of any distinguishing characteristics that would allow an individual to be identified in the video, i.e., clothing color, hair color, parcels or suitcases, exact time, and so on. This information made it possible to determine when the field observer missed or over-counted pedestrians, and to determine whether the manual data collection was properly synchronized with the video.

Manual with clicker

On May 3 and May 5, the field staff collected pedestrian counts using a manual clicker. The observer clicked once for every pedestrian crossing the intersection, regardless of direction. At the end of every 10-minute period, the observer noted the count on the clicker on the data sheet provided.

Manual with Video

The intersections were videotaped using a camera set up on a flatbed truck parked opposite the crosswalk being studied. The camera recorded an image of the crosswalk at an angle that allowed both directions of pedestrian travel to be captured. Video tapes were replaced after each hour.

Researchers involved in the study carefully analyzed the video tapes in order to obtain the most reliable results possible. The researchers tried to identify each pedestrian counted by the field observer. This task was only possible for the days that the field observer noted individual pedestrian characteristics.

The tapes were viewed in variable time, and sometimes viewed more than once if the results were in doubt. On average, one hour of video tape required three hours of video analysis. During the analysis, the researchers paid attention to whether the field counts were synchronized with the videotape and looked for any discrepancies between the field observations and the video images.

DATA ANALYSIS

The purpose of the data analysis was to compare the accuracy of the methods. Because it was not possible to know the exact number of pedestrians on the roadway at any given time, inter-reliability between the methods was used as a proxy for accuracy. The counts derived from the video tapes were assumed to be closest to the actual pedestrian volume.

The comparison used the relative difference between the counts taken through each method to calculate the error:

$$Error = \frac{NPi - NPv}{NPv} \quad (1)$$

where NP_i is the number of pedestrians counted in the field and NP_v is the number of pedestrians counted using the video images. The error was calculated for each interval of data collection (30 minutes for the sheets and 10 minutes for the clickers), as well as for the total number of pedestrians counted at each intersection.

Synchronization of the field counts and video taping was a major issue identified during the video analysis, despite the fact that field staff were directed to synchronize the counting methods. Sometimes the field observer began counting slightly before or after the video camera began recording. When this occurred, it was difficult to compare the counts obtained through each method. To improve the results of the comparison study, counts taken in periods when the field observer was not synchronized with the video were not included in the calculation of the intersection error.

Comparisons of the accuracy of pedestrian gender and age identification were also made, but not included in this paper. The researchers concluded that it was not possible to precisely identify the gender or age of the pedestrians from the video images because of low image resolution.

RESULTS

In the first week of data collection, the field observer did not follow all of the instructions he was given and did not consistently collect data for four-hour periods. For example, he sometimes started counting late; failed to take note of his breaks; and counted bicycles as pedestrians. Despite this, the video tapes were analyzed for the entire counting period (four hours) in order to determine the average hourly pedestrian volume (Table 1).

The results of the comparison reveal that the field observer systematically counted fewer pedestrians than were observed on the video recordings. The average error calculated for the manual counting using sheets was 15%, varying from 9% to 25%, as shown in Tables 2. For the manual counting with clickers, the average error was 11%, varying from 8% to 15% (Table 3). Given the variation in the results, it is not possible to determine which method, with sheets or clickers, is the most accurate.

TABLE 2 Comparison of Counting Methods (Video vs. Sheets)

Period	**Date**						
	4/17/2006	**4/18/2006**	**4/19/2006**	**4/20/2006**	**4/21/2006**	**4/24/2006**	**4/25/2006**
	Error	**Error**	**Error**	**Error**	**Error**	**Error**	**Error**
1:00 to 1:30	Not Counted	Not Counted	-27%	-28%	-16%	-7%	-22%
1:30 to 2:00	150%*	Not Counted	-18%	-6%	0%	-2%	-17%
2:00 to 2:30	-13%	0%	3%	-23%	-17%	-16%**	-29%
2:30 to 3:00	-14%	0%	-28%	-2%	-12%		-26%
4:00 to 4:30	-13%	-22%	-42%	-14%	-8%	-8%	-27%
4:30 to 5:00	-21%	86%*	-67%	-15%	-10%	-11%	-17%
5:00 to 5:30	Not Counted	Not Counted	-25%	-16%	-5%	-3%	-25%
5:30 to 6:00	Not Counted	Not Counted	-49%	3%	-8%	-10%	-31%
Error (Total)	**-15%**	**-11%**	**-21%**	**-12%**	**-10%**	**-9%**	**-25%**

* Not included in the total, because it was not synchronized with the video

**In this period, the field observer failed to record the counts in half hour periods

TABLE 3 Comparison of Counting Methods (Video vs. Clickers)

	5/3/2006						5/5/2006					
	1:00 to 2:00pm						1:00 to 2:00pm					
Error (10 min)	-11%	-43%	-13%	0%	0%	0%	0%	0%	-19%	17%	-8%	100%
Error (hour)	-11%						2%					
	2:00 to 3:00pm						2:00 to 3:00pm					
Error (10 min)	-25%	-67%	0%	100%	-50%	0%	0%	-14%	25%	-31%	-8%	9%
Error (hour)	-23%						-5%					
	4:00 to 5:00pm						4:00 to 5:00pm					
Error (10 min)	0%	17%	33%	-25%	-11%	0%	50%	-25%	-41%	-33%	-40%	-88%
Error (hour)	0%						-32%					
	5:00 to 6:00pm						5:00 to 6:00pm					
Error (10 min)	-20%	0%	38%	-33%	0%	20%	-30%	6%	-64%	-15%	-8%	-88%
Error (hour)	0%						-21%					
Error (4 hours)	**-8%**						**-15%**					

An in-depth analysis of the data revealed that error was often greater at the beginning and end of the data collection period. Possible explanations for this finding include: (i) the observer’s lack of familiarity with the intersection and the counting method at the beginning of the data collection; (ii) the long counting periods, which may have caused the observer to become fatigued and lose attention; and (iii) lack of synchronization with the video that was not possible to identify.

It was assumed that the observer would have more difficulty counting at intersections with high volumes of pedestrians, increasing the error value. However the results revealed that pedestrian flow did not influence the error, since the correlation ($R^2 = 0.1$) between them was weak. Figure 3 presents a graph with the relationship between the error and the pedestrian flow.

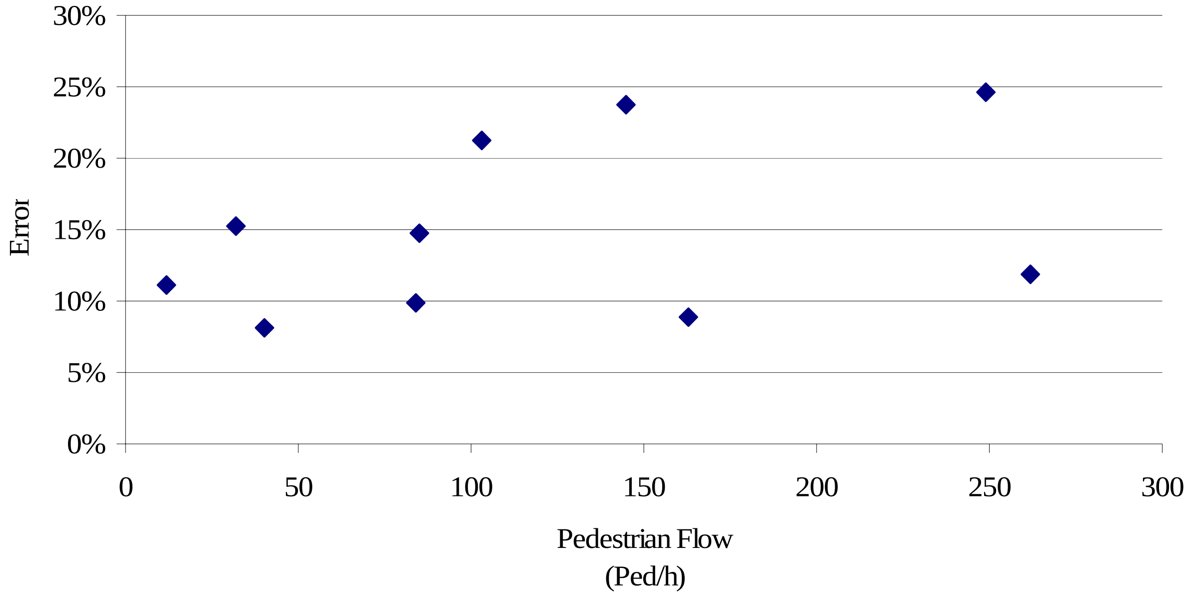


FIGURE 3 Relationship between the error and the pedestrian flow

DISCUSSION

The most significant results of this study were that pedestrian counts taken in the field were systematically lower than counts taken by observing video recordings, and that the accuracy of field counts did not seem to be strongly related to pedestrian flow. These results stem from the fact that the collection of field counts using either sheets or clickers is very difficult to control, and requires planning and organization during the counting day *(5)*.

The level of observer attention is one aspect of field data collection that is difficult to control. In this study, the observer may have become distracted at intersections with little pedestrian activity, but may have been more focused in areas with high activity that demanded his attention. It is also possible that the error was related to the observer’s unique characteristics and motivation. Future studies should use multiple field observers to determine how the characteristics of the observers, such as their experience and background, affect the quality of the pedestrian counts. However, given the budgetary constraints of most transportation agencies, it may be difficult to ensure that field observers have high-level training and experience.

It was expected that manual counts taken with clickers would have very low error because this method allows the observer to keep his attention on the intersection and does not demand that he identify and record pedestrian characteristics. No significant difference was found in the relative accuracy of manual counts using clickers and manual counts using sheets; however, more research is needed to compare the methods.

Although this study suggests that field counts may be less accurate than counts taken with video images, it is often necessary to use field observers to record detailed pedestrian characteristics and behaviors. It is difficult to identify these characteristics on video recordings without adequate image resolution and a well-selected camera angle.

This study suggests that video recordings should be used in situations where the accuracy of the count is of primary importance. However, users of this method should be aware that obtaining an accurate count from video can be very time consuming and requires meticulous attention to the video analysis. Overall, the choice of pedestrian counting method depends on the data collection needs and available resources.

ACKNOWLEDGEMENTS

The authors would like to thank Phyllis Orrick, Jill Cooper, Joseph Zheng, and others at the UC Berkeley Traffic Safety Center for their ongoing support and contributions. Their comments and suggestions helped make this paper possible. This work was partially funded through fellowship support from the Brazilian Foundation for the Coordination of Higher Education and Graduate Training (CAPES) to Mara Diogenes and from the Eisenhower Transportation Fellowship Program to Ryan Greene-Roesel.

REFERENCES

1. Peden, M., R. Scurfield, D. Sleet, D. Mohan, A.A. Hyder, E. Jarawan and C. Mathers. World Report on Road Traffic Injury Prevention. World Health Organization, Geneva, 2004.
2. Høj, N.P. and W. Kröger. Risk Analyses of Transportation on Road and Railway from a European Perspective. *Safety Science*, Vol. 40, No. 1-4, 2002, pp. 337-357.
3. Gårder, P.E. The impact of speed and other variables on pedestrian safety in Maine. *Accident Analysis & Prevention*, Vol. 36, No. 4, 2004, pp. 533-542.
4. Schneider, R., R. Patton, J. Toole and C. Raborn. P*edestrian and Bicycle Data Collection in United States Communities: Quantifying Use, Surveying Users, and Documenting Facility Extent.* FHWA, Office of Natural and Human Environment, U.S. Department of Transportation, Washington, D.C., 2005.
5. Schweizer, T. Methods for counting pedestrians. In *The 6th International Conference on Walking in the 21st Century*. Walk21-VI "Everyday Walking Culture". Zurich, Switzerland, 2005.
6. Schwartz, W. and C. Porter. *Bicycle and Pedestrian Data: Sources, Needs, and Gaps*. Publication BTS00-02. Bureau of Transportation Statistics. U.S. Department of Transportation, Washington, D.C., 2000.
7. Dharmaraju, R., D.A. Noyce, and J.D. Lehman. An Evaluation of Technologies for Automated Detection and Classification of Pedestrians and Bicycles. In *The 71st ITE Annual Meeting*. Compendium of Technical Papers. Institute of Transportation Engineering, Chicago, IL, 2001.

ESTIMATING PEDESTRIAN ACCIDENT EXPOSURE

Automated Pedestrian Counting Devices Report

March 2007


Dan Burden

Traffic Safety Center
Setting New Directions in Traffic Safety

The mission of the UC Berkeley Traffic Safety Center is to reduce traffic fatalities and injuries through multi-disciplinary collaboration in education, research, and outreach. Our aim is to strengthen the capability of state, county, and local governments, academic institutions, and local community organizations to enhance traffic safety through research, curriculum and material development, outreach, and training for professionals and students.

ESTIMATING PEDESTRIAN ACCIDENT EXPOSURE

Automated Pedestrian Counting Devices Report

Prepared for CalTrans under
Task Order 6211

Prepared by
FANPING BU
RYAN GREENE-ROESEL
MARA CHAGAS DIOGENES
DAVID R. RAGLAND

California PATH

1357 S.46th Street, Bldg. 452
Richmond, CA 94804-4648
Tel: 510-665-3552
Fax: 510-665-3537

and

University of California Traffic Safety Center

University of California
Berkeley, CA 94720
Tel: 510/642-0655
Fax: 510/643-9922

ACKNOWLEDGEMENTS

The University of California Traffic Safety Center (TSC) appreciates and acknowledges the contributions of the following participants.

TSC staff:
Noah Radford
Marla Orenstein
Judy Geyer
Kara MacLeod
Tammy Wilder

PATH staff:
Ashkan Sharafsaleh, Research Engineer
Steve Shladover, Research Engineer
Fanping Bu, Research Engineer

Specialized consultants:
Charles Zegeer

Funding for this project was provided by Caltrans.

LIST OF CONTENTS

LIST OF TABLES AND FIGURES

1. PREFACE

1.1. Purpose of the Review

Automated methods are commonly used to count motorized vehicles, but are not frequently used to count pedestrians. This is because the automated technologies available to count pedestrians are not very developed, and their effectiveness has not been widely researched. Moreover, most automated methods are used primarily for the purpose of detecting, rather than counting, pedestrians (Dharmaraju et al., 2001; Noyce and Dharmaraju, 2002; Noyce et al., 2006).

Automated pedestrian counting technologies are attractive because they have the potential to reduce the labor costs associated with manual methods, and to record pedestrian activity for long periods of time that are currently difficult to capture through traditional methods. Data input and storage may also be less time consuming than with manual methods.

On the other hand, the capital costs of automated equipment may be high; specialized training may be required to operate it; and automated devices are generally not capable of collecting information on pedestrian characteristics and behavior. For these reasons, automated devices are not appropriate for all pedestrian data collection efforts.

The choice between which method is more appropriate to collect pedestrian data must be based on the accuracy level desired, budget constraints, and data needs specifications.

1.2. Automated Counting Technologies

Much of the research on automated pedestrian tracking devices has focused on pedestrian detection, not pedestrian counting. Extensive reviews of pedestrian detection technologies were conducted by Noyce and Dharmaraju (2002) and by Chan et al. (2006). Technologies include piezoelectric sensors, acoustic, active and passive infrared, ultrasonic sensors, microwave radar, laser scanners, video imaging (computer vision).

Of the technologies listed above, those most adaptable to the purpose of pedestrian counting are: infra-red beam counters; passive infrared counters; piezoelectric pads; laser scanners; and computer vision technology. None of these devices are widely used for the purpose of counting pedestrians outdoors, but all have some potential to be adapted for that purpose.

This report describes each of these technologies in detail, and discusses some of the technical strengths and weaknesses of each method. It is important to be aware that technical limitations are only one consideration among many when choosing an appropriate counting device. The device “packaging,” such as the method and location of installation may be equally important. For example, the location and accessibility of the device may create liability issues or promote vandalism.

2. BACKGROUND

Automated pedestrian counting capabilities have been developed for a variety of purposes such as traffic planning, retail customer volume statistics and security monitoring. Most of the existing products are developed for indoor environments (e.g. shopping mall, casino, subway station and building entrance etc) or outdoor environments with low density pedestrians (e.g. trails and parks). A few projects have attempted to compare alternative technologies for use in the same environment, but most of published references have focused on individual technologies.

Central London Partnership (CLP) has conducted a project on automatic pedestrian counting technologies to better understand and potentially demonstrate existing products in the outdoor London environment (CLP, 2005). They identified and are testing three commercially available pedestrian counting technologies: computer vision, passive infra-red and vertical laser scanners. They are also conducting manual pedestrian counts to verify the results obtained from automated technologies.

Schneider et al. (2005) presented case studies of pedestrian and bicycle data collection efforts in local communities. Although the purpose of the study was not to evaluate different automatic pedestrian counting technologies, it includes case studies involving different automated counting methods, such as passive infra-red, vertical laser scanner and piezoelectric pad. These devices were used largely in low-density pedestrian environments such as bicycle and pedestrian paths.

The Minnesota DOT sponsored the evaluation of a variety of commercial off-the-shelf (COTS) bicycle and pedestrian detectors as part of a broader project to evaluate traffic detection systems (SRF Consulting Group, 2003). Their report included an extensive literature review on bicycle and pedestrian detection technologies, and they tested four such systems under low volume conditions on a bicycle and pedestrian pathway. Their tests showed that three systems were 100% accurate (one video, one passive IR and one combined ultrasound and passive IR), while one system was 93% accurate (active IR), however they were only presented with one target bicycle or pedestrian at a time, in a simple environment without any significant disturbances.

The University of North Carolina also tested a variety of COTS pedestrian detectors for their ability to automatically trigger walk signals (Hughes et al., 1999). They discussed the issues involved in automatically triggering walk signals, but did not emphasize the strengths and weaknesses of the different detection technologies (Hughes et al., 2000). An analogous study in Israel was reported by Hakkert et al. (2001), again focusing on pedestrian detection but with only passing references to technical performance limitations of the two detection systems that were tested. Note that all of these evaluations have addressed the need for simple detection of pedestrian presence (Is there a pedestrian here?), but not counting how many pedestrians are present or crossing.

The following sections describe technologies that have potential to count pedestrians in an outdoor environment:

- ✓ Infra-red beam counters
- ✓ Passive infrared counters
- ✓ Piezoelectric pad
- ✓ Laser scanner
- ✓ Computer vision

3. INFRA-RED BEAM COUNTERS

Infrared beam counters are one of the most popular types of commercially available counters. The counter is a simple device with low power consumption that can be powered by batteries. It is a popular pedestrian counter for indoor settings.

An infra-red light beam counter is composed of following components: an infra-red beam transmitter, an infra-red beam receiver and a data logger. The transmitter emits a constant infrared beam that is intercepted by the receiver at an appropriate position. When the beam is interrupted by a solid object passing through, a count is registered by the data logger. Infrared beam counters typically operate at a range of around 30 meters.

There are three types of infra-red beam counters. Figure 3.1 shows the infra-red beam counter with separated infra-red beam transmitter and receiver. Figure 3.2 shows the infra-red beam counter with transmitter and receiver in the same housing. A separate reflector is used to bounce back the infra-red beam. Figure 3.3 shows the infra-red beam counter with a two beam setup that can provide the pedestrian traveling direction.

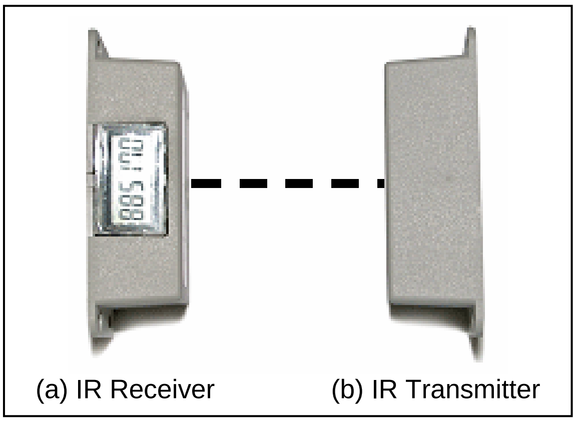

(a) IR Receiver (b) IR Transmitter

Figure 3.1: Infra-red beam transmitter and receiver

(a) IR Receiver and Transmitter (b) IR Reflector

Figure 3.2: Infra-red beam transmitter/receiver and reflector

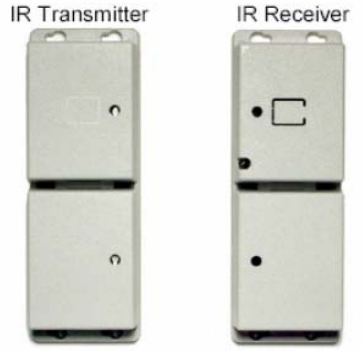


Figure 3.3: Infra-red beam counter with directional counting

The following are some of the major drawbacks of infrared beam counters:

1. Infrared beam counters cannot differentiate between pedestrians and other objects. Vehicles, insects flying close to the transmitter, or even rain drops could block the counting beam and trigger the counter;

2. The transmitter and receiver need to be aligned carefully to ensure the reception of the beam at the receiver end. If either the transmitter or receiver are installed on a flexible structure, strong winds or other disturbances could cause the beam to miss the receiver;

3. When several pedestrians cross the counting beam simultaneously, they are only registered as one count.

4. PASSIVE INFRARED COUNTERS

Passive infrared devices count pedestrians by tracking the heat emitted by moving objects. The earliest infrared counters were based on CCD (charge coupled device) and CMOS (complementary metal oxide semiconductor) technologies. These are very expensive and usually targeted for military use. More recently, pyroelectric sensing technology has been developed as cheap alternative that does not require expensive cooling methods.

Figure 4.1 shows an example of a single sensor pyroelectric people counting device manufactured by Eco-Counter, a company based in France. It operates by detecting the body heat of pedestrians in close proximity (usually within 4 meters).

Figure 4.1: People counter with single pyroelectric sensor (Eco-Counter)

Figure 4.2 shows an example of a double sensor pyroelectric counter that is capable of providing directional counts. The device will register a count when it detects an object with a temperature that exceeds a certain threshold. However, neither the single or double sensor device can distinguish whether the heat source is generated by a pedestrians or a vehicle. It also has difficulty distinguishing individual pedestrians walking closely within a group, so may underestimate pedestrian volumes. The Vermont Agency of Transportation is currently testing the pyroelectric sensor developed by Eco-Counter.

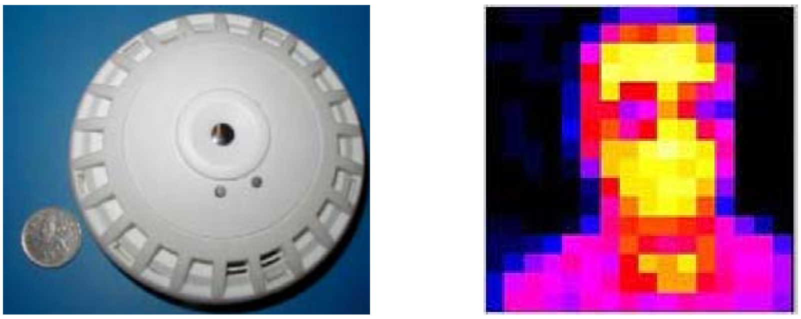
Figure 4.2: People counter with double pyroelectric sensors (Eco-Counter)

The drawbacks of single (or double) pyroelectric sensor based people counters may be addressed by using a pyroelectric sensor array to generate infrared images. Figure 4.3a shows a top mounted people counter with 16 by 16 pyroelectric sensor array manufactured by IRISYS.

The low resolution (16 by 16) thermal images produced by pyroelectric sensors (Figure 4.3b) can be improved to 128 by 128 by using interpolation algorithms. Afterwards, standard video imaging processing techniques can be applied to extract pedestrian counts.

Although the coverage area for single IRISYS passive infrared people counter is small (3.5m-by-3.5m), IRISYS provides an option to connect up to 30 counters using a CAN (Controller Area Network) bus device.

Kerridge et al. (2004) conducted experiments with the IRISYS people counter in an indoor corridor. They monitored the counter's ability to track pedestrian movements. The results showed some loss of tracking ability at higher pedestrian densities, when it became more difficult for the detector to distinguish adjacent pedestrians.


(a)Passive infrared count

(b)Original Image generated by pyroelectric sensor array

Figure 4.3: Passive infrared counter with pyroelectric sensor array (IRISYS people counter)

5. PIEZOELECTRIC PAD

Piezoelectricity, or “pressure” electricity, is the property of certain materials that produce a change in electrical properties with mechanical pressure. For application to pedestrian detection, piezo-cables with piezoelectric material are usually fabricated into a “mat” (Figure 5.1). When a person steps onto the mat, an electrical signal is generated and triggers a count.

Figure 5.1: Piezoelectric pad counter (Eco-Counter)

A piezoelectric pad counter is a simple reliable sensor for pedestrian counting. Several piezoelectric pads can be buried together for large coverage area. Timer systems have also been developed to ensure that only one person is counted even if they make two steps on the pad.

The piezoelectric pad counter does not require complex signal processing. However, it does require physical contact between a pedestrian and the sensor mat. Therefore, the piezoelectric pad counter is ideal when direct physical contact between pedestrian and sensor is assured, such as at a location where pedestrians are channeled into a crossing.

6. LASER SCANNER

A laser scanner is a high-resolution laser range finder (Fuerstenberg and Lages, 2003; Fuerstenberg and Scholz, 2005; Streller and Dietmayer, 2004; Zhao and Shibasaki, 2005). The laser scanner emits infrared laser pulses and detects the reflected pulses. The measurement principle is based on the time-of-flight method, where the distance to the target is directly proportional to the time interval between transmission and the reception of a pulse.

Scanning of the measurement beam is achieved by a rotating prism and covers a viewing angle of up to 360 degrees. The original data from a laser scanner is much like vision image data in the horizontal scanning plane with accurate distance (centimeter level) and azimuth angle information (from 0.25 degree to 1 degree depending on scanning frequency).

A procedure similar to image processing is applied to analyze the laser “image.” First, clustered data points are grouped into different objects by segmentation. Then the objects are classified into different categories according to their characteristics. For example, pedestrians can be classified by the characteristics of their moving legs (Fuerstenberg and Dietmayer, 2004).

There are two classes of laser scanner: horizontal scanning and vertical scanning. Figure 6.1 shows a multifunctional traffic sensor with a horizontal scanning Sick laser scanner (Lotraffic from LogObject AG). It is capable of detecting and counting pedestrian within a 15m radius.

Figure 6.1: Horizontal scanning configuration (Lotraffic with Sick laser scanner)

Figure 6.2 shows a vertical laser scanner with two vertical scanning laser beams (PeCo people counter from LASE GmbH). It is capable of covering a passage width of up to 26m and providing directional counts and classification of pedestrians according to their height.

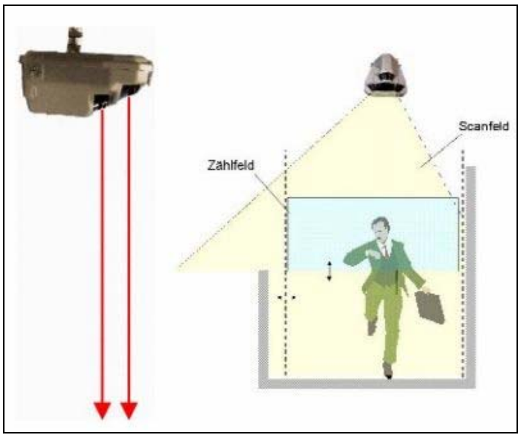


Figure 6.2: Vertical scanning configuration (PeCo from LASE GmbH)

Excellent range accuracy and fine angular resolution make laser scanners suitable for applications in which a high-resolution image of the surroundings is required. However, since they are optical sensors, weather conditions like fog or snow will limit their detection range. The signal processing is a little more complex for laser scanner compared with ultrasonic or microwave radar, therefore a dedicated CPU (central processing unit) may be needed.

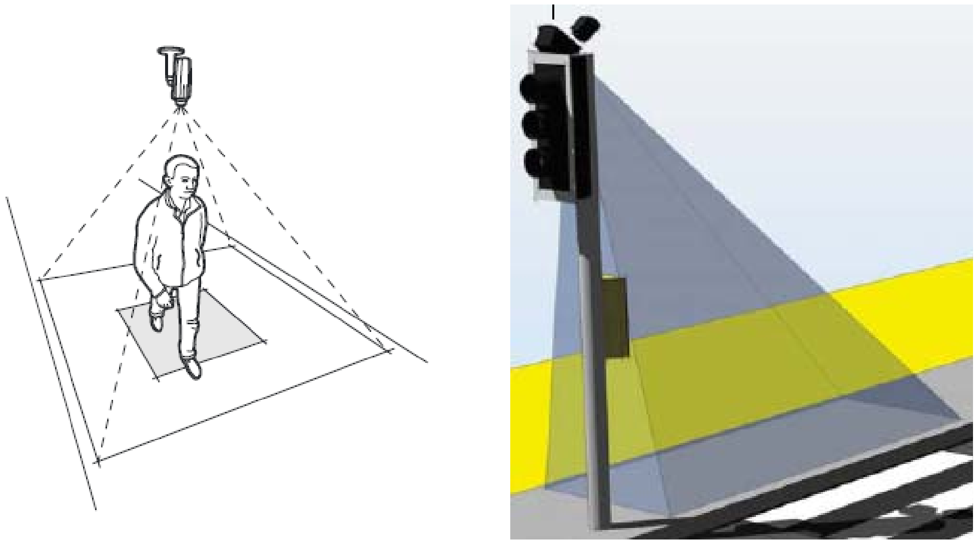
7. COMPUTER VISION

Computer vision utilizes intelligent processing of digital images of pedestrians captured with a video camera to count pedestrians. The processor subtracts the static background from the image and then tracks the remaining objects to determine whether they are pedestrians (CLP, 2005).

Although a video camera can obtain much richer information about the surrounding environment compared with other types of sensors, the image sequences can not be used for anything directly without further interpretation. Extracting useful information from available image sequences is not a trivial task for several reasons (Zhao, 2001; Rabaud and Belongie, 2006):

- ✓ Pedestrian detection and counting involves a complex uncontrolled outdoor environment. Lighting conditions may change due to weather, sunrise or sunset. The cluttered urban environment also makes it difficult to distinguish pedestrian from nearby objects such as buildings, vehicles, poles and trees.
- ✓ A wide range of variations exist in pedestrian appearance because of clothing, pose, occlusion, shadow, motion, size and skin color.
- ✓ Accurate pedestrian counting of high-density pedestrian crowds is extremely difficult due to the potential that pedestrians may occlude one another's images, and because of the need to distinguish among many independently moving bodies.

To count pedestrians using video imaging, the video camera is usually installed at a fixed location. To minimize occlusions among passing pedestrians and simplify the corresponding video processing algorithm, the camera can be mounted directly above the interested area as shown in Figure 7.1a. However, it is not always possible to find a suitable overhead installation position. In this case, video cameras can also be installed on nearby building or signal pole (Figure 7.1b), but there is a risk of occlusion due to the angle of view. To monitor a wide open area, multiple cameras may be needed for overhead installation.



(a) Overhead camera installation (b) Camera installed on signal pole or nearby building

Figure 7.1: Camera Installations

To cope with inherently dynamic phenomena (people enter the scene, move across the field of view of the camera, and finally cross the counting line), the people tracking and counting problem has been decomposed into the following three steps (Rossi and Bozzoli, 1994; Kim et al., 2003):

1. Determine whether any potentially interesting objects have entered into the scene (alerting phase);
2. Track their motion until the counting line is reached (tracking phase);
3. Establish how many people correspond to tracked objects (interpretation phase).

7.1. Alerting Phase

The alerting phase includes foreground objects extraction and pedestrian detection. Since the camera is usually installed on a fixed position, a background image is generated for calibration purposes at the beginning of installation. These images are updated using a very slow recursive function to accommodate background changes such as lighting conditions (Masound and Papanikolopoulos, 2001). Foreground objects including pedestrians and vehicles can be extracted by subtraction from the background. Terada et al. (1999) used distance information acquired from stereo cameras to further improve the performance of foreground object extraction.

Various algorithms have been proposed to detect pedestrians in image sequences acquired from video cameras based on their unique features. Recent research shows two main trends. Motion-based approaches take into account temporal information and try to detect the periodic features of human gait in the movement of candidate patterns. Shape-based approaches rely on shape features to recognize pedestrians.

7.1.1. Motion-based detection approaches

Motion-based approaches use rhythmic features or motion patterns unique to human beings. Yasutomi and Mori (1994) used the Maximum Entropy Method to observe the periodic changes of image intensity caused by walking. Cutler and Davis (2000) used Fourier Transformation with Hanning window to find periodicity in the acquired image sequences.

The United Kingdom’s Defense Evaluation and Research Agency counted pedestrian motions in order to estimate their exposure to risk in traffic (Allsopp and Smith, 1997). The report on this work dates from 1997, when it was necessary to use custom computer hardware. The results showed video imaging was capable of 85% accuracy when sampling 35 pedestrians in 30 minutes. There were concerns about occlusion problems at high pedestrian densities.

There are several limitations to motion-based approaches. First, motion-based schemes cannot detect stationary pedestrians or pedestrians engaged in unusual movements like jumping. Second, the pedestrian’s feet or legs must be visible in order to extract rhythmic features or motion patterns. Third, the recognition procedure requires a sequence of images, which delays the identification until several frames later and increases processing time.

7.1.2. Shape-based detection approaches

Shape-based methods allow recognition of both moving and stationary pedestrians. Papageorgiou and Poggio (1999) represented human shape characteristics using

Harr wavelets. A support vector machine trained with characteristics extracted from example human images is used as a classifier.

In order to detect a partially occluded pedestrian, the same system is modified to first detect components of the human body (e.g. head, torso or limbs) and then the detected body parts are assembled together (Mohan et al., 2001). The proposed system has to search the whole image at multi-scale for pedestrian characteristics. This search process increases the computation cost substantially.

Although the shape-based method is more general, the major drawbacks associated with the shape method are:

1. High false positive rate due to variation of human shape and changing lighting conditions.
2. Heavy computation burden when performing feature matching.

Different approaches can be tried in order to resolve these drawbacks. The single-frame shape match can be combined with motion analysis to reduce false positive rates (Shashua et al., 2004). A specialized system-on-a-chip hardware solution can be used to increase processing speed (Elouardi et al., 2004; Mobileye, 2007). Knowledge about certain sites and situations (e.g. traffic signal, pedestrian crossing, etc.) can be used as *a priori* information to optimize the vision-processing algorithm (Lombardi and Zavidovique, 2004).

The main difficulty associated with shape based approaches is the problem of accommodating the wide range of variations in pedestrian appearances due to pose, various articulations of body parts, lighting, clothing, occlusion, etc. The key issues are to: (i) find a concise yet sufficient human shape feature representation that could achieve high inter-class variability with low intra-class variability; and (ii) maintain a balance between accuracy of detection and processing time.

7.2. Tracking Phase

The purpose of tracking is to establish connections of objects among frames and determine if the pedestrian has reached the counting line. Most tracking algorithms employ different variations of the Kalman Filtering approach (Heikkila and Silven, 2004). In the study of (Masound and Papanikolopoulos, 2001), the ground-plane constraint was imposed on the pedestrian motion model. This constraint assumes that all pedestrian motion is constrained to the ground plane.

7.3. Interpretation phase

The interpretation phase of vision based pedestrian counting involves determining the actual pedestrian count based on the pedestrian objects detected and tracked in the previous phases. For the motion based pedestrian detection technique, different pedestrian objects generated from the detection step may belong to the same person (Antonini and Thiran, 2006). Therefore an overestimation may occur. Clustering techniques can be applied to the detected target trajectories in order to reduce the bias between the number of tracks and real number of pedestrians (Antonini and Thiran, 2006).

Some of the tracked objects may consist of several pedestrians (Heikkila and Silven, 2004). A computer-generated shape of a pedestrian group with two or three persons can be introduced so that such groups can be counted properly. If high-density crowds are present (e.g. during a political demonstration), the occlusion of pedestrians makes it very difficult to distinguish among each individual pedestrian. Face recognition and head detection can be used to improve counting accuracy in these situations (Casas and Folch, 2005; Liu et al., 2005).

One study developed a specialized tracker (a highly parallel version of Kanade-Lucas-Tomasi Feature Tracker) to process the video into a set of feature trajectories. The identified trajectories were then clustered into pedestrian objects (Rabaud and Belongie, 2006). In Kong et al. (2005) the number of pedestrians is first assumed to be proportional to the pixels of pedestrian objects. To further address the occlusion problem, a nonlinear relationship between pedestrian count and the pixels of pedestrian object is established and a neural network is trained to represent such nonlinear relations.

Researchers at the National Institute of Transportation Safety Research in France (INRETS) were able to accurately count the number of passengers passing through specific locations in transit stations using a “linear camera” optical method (Khoudour et al., 1998). They used an infrared camera and an active illumination source on the ceiling, looking down at reflective lines on the ground, and then counted the number of pedestrians passing by, including estimates of their speed. They study resulted in an accuracy of 99% in counting pedestrians passing through a 3 m wide passageway, but noted some loss of accuracy at higher densities when pedestrians were so close that their image “blobs” merged together.

8. COMPARISON BETWEEN AUTOMATED METHODS

There are a variety of technologies capable of counting pedestrians. Each technology has strengths and weaknesses that make it particularly suited to different purposes, budgets, and counting environments. Table 8.1 provides a summary and comparison of the devices discussed in this document.

Table 8.1: Summary of automated pedestrian counting devices

Counter	***Pros***	***Cons***	***Manufacturer and Cost***
Infra-red beam counter	Cheap and widely available commercially; Low power consumption; Easy installation; Highly portable.	Infrared beam counter cannot differentiate pedestrian and other objects; Transmitter and receiver need to be aligned carefully to ensure the reception of beam at the receiver end; Both transmitter and receiver should not be installed on a flexible structure; When several pedestrians cross the counting beam simultaneously, they are only registered as one count.	Jamar Technologies Inc $790
Passive infra-red counter	Cheap and widely available commercially; Low power consumption; Not affected by wet or foggy weather; Counter with multiple sensor arrays could achieve performance comparable with computer vision.	Single or double sensor counter cannot distinguish between individuals and groups; Temperature can affect counter performance; Limited coverage area.	Irisys $1400 for counter with multiple sensor array EcoCounter $2000 for counter, $600 for software
Piezo-electric pad	Low maintenance cost; Low power consumption;	Need physical contact between pedestrian and pad;	Eco-Counter Cost estimate not available

Counter	***Pros***	***Cons***	***Manufacturer and Cost***
	Capable of counting pedestrians on sidewalks.	Sub-surface installation is expensive; Limited coverage area; Some of products cannot differentiate between single pedestrian and group of pedestrians.	
Laser scanner	Accurate range measurement; Can differentiate pedestrian according to their height; Easy setup; Large coverage area.	Expensive; Performance could be affected by different weather conditions.	LASE GmbH Around $9000 for counter only
Computer vision	Large coverage area; Has the potential to count accurately in various conditions such as crowded pedestrians, different lighting conditions; Can be manually reviewed to collect pedestrian characteristics; Easy installation and setup; The video can be recorded for manual review.	Most commercially available products are intended for indoor setting; The difficulty of counting pedestrians in crowded settings has not yet been resolved; The performance can be affected by different environmental conditions if not designed properly.	Video Turnstile Start from $1230

REFERENCES

Allsopp, R.A., and S.M. Smith. Image Processing for the Analysis of Pedestrian Behaviour. In **Fourth World Congress on Intelligent Transport Systems Proceedings**. Berlin, Nov. 1997, Paper No. 2120.

Antonini, G. and J. Thiran. Counting Pedestrians in Video Sequences Using Trajectory Clustering**. IEEE Transactions on Circuits and Systems for Video Technology**, 2006. Vol. 16, No. 8, 2006, pp 1008-1020.

Casas, J. , A. Sitjes and P. Folch. Mutual Feedback Scheme for Face Detection and Tracking Aimed at Density Estimation in Demonstrations. **IEE Proceedings Vision, Image & Signal Processing**, Vol. 152, No. 3, 2005, pp 334-346.

Central London Partnership (CLP) **Automatic Pedestrian Counting Trial**. Stage 3 - Final Report. Central London Partnership, 2005.

Chan, C., F. Bu, and S. Shladover. **Experimental Vehicle Platform for Pedestrian Detection**. California PATH Research Report UCB-ITS-PRR-2006-16, 2006.

Cutler, R. and L. Davis. Robust Real-Time Periodic Motion Detection, Analysis, and Applications. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, Vol. 22, No. 8, 2000, pp. 781-796.

Dharmaraju, R., D.A. Noyce, and J.D. Lehman. An Evaluation of Technologies for Automated Detection and Classification of Pedestrians and Bicycles. In **The 71st ITE Annual Meeting. Compendium of Technical Papers**. Institute of Transportation Engineering, Chicago, IL, 2001.

Elouardi, A. S., A. Bouaziz, A. Dupret, J. Klein and R. Reynaud. On Chip Vision System Architecture Using a CMOS Retina. In **Proceedings of IEEE Intelligent Vehicle Symposium**, Parma, Italy, 2004, pp. 206-211.

Fuerstenberg, K. C. and J. Scholz. Reliable Pedestrian Protection Using Laser-Scanners. In **Proceedings of IEEE Intelligent Vehicle Symposium**, Las Vegas, 2005.

Fuerstenberg, K. C. and K. Dietmayer. Object Tracking and Classification for Multiple Active Safety and Comfort Applications Using a Multilayer Laserscanner. In **Proceedings of IEEE Intelligent Vehicle Symposium**, Parma, Italy, 2004, pp. 802-807.

Fuerstenberg, K. C. and U. Lages. Pedestrian Detection and Classification by Laserscanners. In **Procedures of 9th EAEC (European Automobile Engineers' Cooperation) International Congress**, Paris, June 2003.

Hakkert, A. S., V. Gitelman and E. Ben-Shabat. An Evaluation of Crosswalk Warning Systems. In **80th Annual Meeting of the Transportation Research Board**, Washington DC, January 2001, Paper No. 01-0135.

Heikkila, J. and O. Silven. A Real-Time System for Monitoring of Cyclists and Pedestrians. **Image and Vision Computing**, Vol. 22, 2004, pp. 563-570.

Hughes, R., H. Huang and C.V. Zegeer and M. Cynecki. Automated Pedestrian Detection Used in Conjunction with Standard Pedestrian Push Buttons at Signalized Intersections. In **79th Transportation Research Board Annual Meeting**, Washington DC, January 2000.

Hughes, R., H. Huang and C.V. Zegeer. ITS and Pedestrian Safety at Signalized Intersections. **ITS Quarterly**, Vol. 7, No. 2, 1999, pp. 37-45.

Kerridge, J., A. Armitage, D. Binnie and L. Lei, “Monitoring the Movement of Pedestrians Using Low-Cost Infrared Detectors: Initial Findings”, **Transportation Research Record** 1878, Transportation Research Board, 2004, pp. 11-18.

Khoudour, L., J.P. Deparis and B. Meunier, Pedestrians Counting in Public Transport as an Exploitation Help for Operators. In **ITS America Eighth Annual Meeting**, Detroit, 1998.

Kim, J., K. Choi, B. Choi, J. Lee and S. Ko. Real-Time System for Counting the Number of Passing People Using a Single Camera. In **25th Annual meeting of the German Association for Pattern Recognition, DAGM Symposium,** Magdeburg , Germany, 2003, pp. 466-473.

Kong, D., D. Gray and H. Tao. Counting Pedestrians in Crowd Using Viewpoint Invariant Training. In **Procedings of British Machine Vision Conference**, 2005.

Liu, X., P. Tu, J. Rittscher, A. Perera and N. Krahnstoever. Detecting and Counting People in Surveillance Applications. In **Proceedings of the IEEE Conference on advanced Video and Signal Based Surveillance**, 2005, pp. 306-311.

Lombardi, P. and B. Zavidovique. A Context-Dependent Vision System for Pedestrian Detection. In **IEEE Intelligent Vehicle Symposium**, 2004, pp. 578-583.

Masound, O. and N. Papanikolopoulos. A Novel Method for Tracking and Counting Pedestrians in Real-Time Using a Single Camera. **IEEE Transactions on Vehicular Technology**, Vol. 50, No. 5, Sept. 2001, pp1267-1279.

Mobileye. **http://www.mobileye.com**. Accessed February 10, 2007.

Mohan, A., C. Papageorgiou and T. Poggio. Example-Based Object Detection in Images by Components. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, Vol. 23, No. 4, 200, pp. 349-361.

Noyce, D. and R. Dharmaraju. **An Evaluation of Technologies for Automated Detection and Classification of Pedestrians and Bicyclists.** Massachusetts Highway Department report, 2002.

Noyce, D., A. Gajendranand and R. Dharmaraju. Development of a Bicycle and Pedestrian Detection and Classification Algorithm for Active-Infrared Overhead Vehicle Imaging Sensors. In **Transportation Research Board 85th Annual Meeting**. CD-ROM. Transportation Research Board, Washington , D.C., 2006

Papageorgiou, C. and T. Poggio, Trainable Pedestrian Detection. **In Proceedings of International Conference on Image Processing**, 1999, pp. 35-39.

Rabaud, V. and S. Belongie. Counting Crowded Moving Objects. In **Proceedings of IEEE Computer Society on Computer Vision and Pattern Recognition**, 2006, pp. 705-711.

Rossi, M. and A. Bozzoli, “Tracking and Counting Moving People”, In **Proceedings of IEEE International Conference on Image Processing**, 1994, pp. 212-216.

Schneider, R., R. Patton, J. Toole and C. Raborn. **Pedestrian and Bicycle Data Collection in United States Communities: Quantifying Use, Surveying Users, and Documenting Facility Extent**. FHWA, Office of Natural and Human Environment, U.S. Department of Transportation, Washington, D.C., 2005.

Shashua, A., Y. Gdalyahu and G. Hayun. Pedestrian Detection for Driving Assistance Systems: Single-frame Classification and System Level Performance. In **Proceedings of IEEE Intelligent Vehicle Symposium**, 2004.

SRF Consulting Group, Inc. **Bicycle and Pedestrian Detection**. Minnesota DOT Research Report, Feb. 2003

Streller, D. and K. Dietmayer. Multiple Hypothesis Classification with Laser Range Finders. In **Proceedings of IEEE Intelligent Transportation Systems Conference**, 2004, pp. 195-200.

Terada, K., D. Yoshida, S. Oe, and J. Yamaguchi. A Method of Counting the Passing People by Using Stereo Image. In **Proceedings of IEEE International Conference on Imaging Processing,** 1999, pp. 338-342, 1999.

Yasutomi, S. and H. Mori. A Method for Discriminating of Pedestrian Based on Rhythm. In **Proceedings of IEEE International Conference on Intelligent Robots and Systems**, 1994, pp. 988-995.

Zhao, H. and R. Shibasaki. A Novel System for Tracking Pedestrians Using Multiple Single-row Laser Range Scanners. **IEEE Trans. On Systems, Man and Cybernetics**, Vol. 35, No. 2, 2005, pp. 283-291.

Zhao, L. **Dressed Human Modeling, Detection and Parts Localization**. Ph.D. Thesis, The Robotics Institute, Carnegie Mellon University, July 2001.

ESTIMATING PEDESTRIAN ACCIDENT EXPOSURE

Approaches to a Statewide Pedestrian Exposure Database

March 2007

Traffic Safety Center
Setting New Directions in Traffic Safety

ESTIMATING PEDESTRIAN ACCIDENT EXPOSURE

Approaches to a Statewide Pedestrian Exposure Database

Prepared for CalTrans under
Task Order 6211

Prepared by
RYAN GREENE-ROESEL
MARA CHAGAS DIOGENES
DAVID R. RAGLAND

University of California Traffic Safety Center

University of California
Berkeley, CA 94720
Tel: 510/642-0655
Fax: 510/643-9922

ACKNOWLEDGEMENTS

The University of California Traffic Safety Center (TSC) appreciates and acknowledges the contributions of the following participants.

TSC staff:
Noah Radford
Marla Orenstein
Judy Geyer
Kara MacLeod
Tammy Wilder

PATH staff:
Ashkan Sharafsaleh, Research Engineer
Steve Shladover, Research Engineer

Specialized consultants:
Ron Van Houten
Charles Zegeer

Funding for this project was provided by Caltrans.

LIST OF CONTENTS

LIST OF TABLES AND FIGURES

1. INTRODUCTION

1.1. Transport System Usage Data

One of the activities of the transportation sector is to inform policy through data collection, especially data on usage of the transport system. These data are essential to government agencies, academics, and non-profit organizations that monitor the performance of the transportation system. The data also provide important information for allocating transportation funds effectively.

The collection of vehicle, transit, and aviation system usage data is mandated by the federal government. National and state-wide databases exits to store the data, which is used by numerous individuals and organizations for the purposes outlined above. For example, states are required by law to collect data on motor vehicle traffic volumes on state highways and to submit them to the Federal Highway Administration's Highway Performance Monitoring System (HPMS) database. The volumes are collected according to a standardized procedure outlined in the FHWA Traffic Monitoring Guide. The volumes are then used in the allocation of billions of dollars of formula-grant federal surface transportation funds (FHWA, 1999).

Transit system usage data are also collected systematically. Transit agencies around the country are required by law to submit ridership data to the Federal Transit Administration's National Transit Database (FTA's NTD) through an internet-based reporting system. Database statistics are used to distribute over $4 billion of FTA funds to transit agencies in urbanized areas (UZAs), primarily through FTA's Urbanized Area Formula Program and Nonurbanized Area Formula Program. Data are submitted via an online reporting system (FTA, 2007).

1.2. Lack of Pedestrian System Usage Data

Although walking is the second most frequently used form of travel on a per-trip basis after the private automobile (Hu and Reuscher, 2004), no such federal or state laws mandate the collection of pedestrian volume data. A few cities and states routinely collect pedestrian count data, but most collect them only sporadically or not at all (Schwartz and Porter, 2000).

The only widely-available measures of pedestrian activity exist in the form of travel surveys conducted at the state, metropolitan, and national level. While these provide information about pedestrian trips made by individuals and households, they do not provide information about the usage of specific pedestrian facilities. In other words, data are available on who is walking but not where they are walking. This makes it difficult for governments or organizations to justify pedestrian facility investments, to monitor pedestrian safety, or to allocate transportation funds on the basis of pedestrian activity.

The need for a consistent, widely-available source of pedestrian system usage data has long been recognized. The Institute of Transportation Engineers Pedestrian and Bicycle Council and Alta Planning and Design have called the lack of usage data one of the greatest challenges facing the bicycle and pedestrian field (Alta Planning and Design, 2006). A report by the Bureau of Transportation Statistics listed the need for better pedestrian system usage data as the most pressing of national pedestrian and bicycle data priorities (Schwartz and Porter, 2000).

1.3. A Pedestrian Volume Database

This report discusses approaches to addressing the need for better and more widely available pedestrian volume data in the state of California. While a variety of approaches could be used, this report focuses on the strategy of a statewide pedestrian volume database.

This database would meet a variety of data needs for different stakeholder groups. One of its principal purposes would be to allow safety professionals at the state and local levels to estimate pedestrian exposure to risk at specific sites.

Since exposure data is essentially equivalent to facility usage data, a pedestrian exposure data would be used for many purposes beyond risk analysis. Facility usage data might be used by municipalities to pinpoint new infrastructure needs, or to determine whether new infrastructure encourages more pedestrian activity. Facility usage data might also be used by advocacy groups as a means to promote new facility investments.

If the database includes information beyond pedestrian volumes, such as facility characteristics (e.g. the availability of sidewalks and intersection crossings) or planning variables (e.g. land uses and population densities), it may be used as a means to improve pedestrian demand modeling techniques or to investigate the relationship between pedestrian environmental quality and pedestrian demand. Furthermore, if facility funding data are included, the database may also be used as a means to track spending on pedestrian projects.

In short, there is a wide range of usage for a pedestrian volume database. In designing the database, it is important to maximize its utility to pedestrian stakeholder groups while recognizing the costs associated with increased complexity.

1.4. Decision Points

Creation of a pedestrian volume database for the state of California involves several major decision points. This report examines these decision points and provides a range of database approaches given different funding and institutional constraints, and describes the challenges that will need to be addressed in the database development process.

Chapter 2 discusses the technical and institutional challenges inherent in creation of a pedestrian exposure database. Chapter 3 discusses the need for an inventory of the pedestrian network as a starting point for the database, and present two existing sources for the network. Chapter 4 presents a range of approaches to data collection process, and suggests data points that might be appropriate for inclusion in the data collection process. Chapter 5 discusses how pedestrian demand modeling might be used to estimate pedestrian volumes with limited data inputs. Chapter 6 summarizes the report and provides recommendations for future development of the database.

2. CHALLENGES

Before discussing possible strategies for developing a statewide database of pedestrian volumes, it is important to consider the challenges that have prevented the creation of such a database up to this point. These challenges fall into two main categories: technical and institutional. Technical challenges are those arising from the characteristics of the pedestrian network and pedestrian travel. Institutional challenges are those arising from the need to coordinate and fund the collection of pedestrian volume data. The following section discusses these challenges and possible means to overcome them.

2.1. Technical Challenges

There are a number of technical challenges related to the design and implementation of a database of pedestrian volumes. Most are related primarily to the complexity and size of the pedestrian network relative to the vehicle network.

2.1.1. Lack of statewide inventory of the pedestrian network

Any database capturing usage of the pedestrian network must build upon an inventory of the network. But at this time, no complete statewide network exists. One way to overcome this challenge is to use portions of the vehicle road network, such as the state highway system, as a proxy for the pedestrian network. Another approach would be to use Geographic Information Systems (GIS) based roadway centerline files, which are publicly available from the U.S. Census, as a proxy for the pedestrian network. These alternatives are discussed in more detail in the next chapter.

2.1.2. Pedestrian network is distinct from vehicle network

Both of the approaches listed above depend to some extent on using the vehicle network as a proxy for the pedestrian network. Although a great deal of the pedestrian network (e.g. sidewalks) overlaps with the vehicle network, the two are distinct. Unlike vehicles, pedestrians are not constrained by the boundaries of the roadway, and can move off-road through parks, trails, driveways, and buildings with relative ease (Radford and Ragland, 2006). For these reasons, it is not entirely correct to use the vehicle network as a substitute for the pedestrian network. It would be preferable to use the vehicle network as a starting point for the pedestrian inventory and to gradually modify it to better represent travel pathways available to pedestrians.

2.1.3. Pedestrian network is very large

Conducting a sampling program over a wide area, such as a state, presents major a technical challenge, in addition to some institutional challenges (discussed below). One method of addressing this challenge is to sample only certain parts of the network, such as a limited set of cities or intersections, and to estimate volumes in the remaining areas using modeling techniques. This technique is already applied in vehicle volume estimation.

For example, Caltrans samples vehicle volumes on the state highway system, which is a subset of the entire road network. Volumes on the remaining local roads are estimated using Caltrans’ Motor Vehicle Stock Travel and Fuel Forecast model. The model estimates current and future vehicle miles traveled using inputs such as income and fuel consumption data (Caltrans, 2005).

2.1.4. Pedestrian movement is complex

Relative to vehicle movement, pedestrian movement is very complex. Whereas vehicles move along a small number of restricted pathways and can only execute a limited number of turning movements, pedestrians move freely through their environment. They are able to turn abruptly, reverse directions, and pause at will. It is difficult to identify when a pedestrian trip begins and ends, as the pedestrian is able to combine multiple sub-journeys that involve pauses of indeterminate length (Kerridge et al., 2001).

The complexity of pedestrian movement makes measuring pedestrian travel along corridors difficult, because it is difficult to know the path taken by each pedestrian. For example, if a pedestrian is counted at the end of a block, it is uncertain whether she has been traveling for the entire block or if she just exited a building. With vehicle volumes, by contrast, it is often assumed that any vehicle passing through a point has been traveling along the length of the segment (FHWA, 2001).

2.1.5. Difficult to link accident data with pedestrian network.

For a database of pedestrian safety to be useful, it must be easy to link to other data sources, particularly accident records data. For example, the pedestrian volume at a site should be linked to the accident records associated with the site.

Unless the pedestrian database is confined to the state highway system, it may be difficult to link it automatically with accident data from the current California Statewide Integrated Traffic Records System (SWITRS). SWITRS data is automatically linked by postmile to the state highway inventory, but is not automatically linked to local roads (Boehm, 2007).

However, this problem may be overcome in the future as efforts are made to pinpoint all SWITRS accidents using Geographic Information Systems. The University of California at Berkeley Traffic Safety Center is currently working on a project to geocode all SWITRS accidents (State-level Geocoding of SWITRS Data).

2.1.6. Pedestrians are found mostly in urban areas.

Unlike vehicles, pedestrian are found mostly in urban areas in the United States. Techniques for measuring and estimating pedestrian volumes in urban areas do not necessarily function well in rural areas. Thus a variety of data collection strategies may be needed to obtain pedestrian volume estimates for the entire state.

2.2. Institutional Challenges

For any database to function successfully, roles and responsibilities for data collection, maintenance and storage must be clearly defined. The following section describes some of the questions that must be addressed.

2.2.1. Need for hosting institution

Hosting a database involves data input, maintenance and cleaning as well as provision of data to interested parties. A recent paper describing the feasibility of a new federal database of airline passenger surveys estimated that 0.3 to 0.6 of one full-time staff person would be needed to conduct basic hosting responsibilities, which include data input and cleaning; outreach to data collection organizations; and marketing to data users (Gosling and Hansen, 2006).

If the proposed database is to cover the entire state, a state agency will need to play a role in data hosting. It is also possible that the database could be made up of component parts hosted by sub-state agencies such as county governments or Caltrans districts.

2.2.2. Need for institutionalized data collection

The responsibility for data collection is a major issue, as data collection is a costly and time consuming activity. Data collection alternatives include volunteer data collection/submission; state-mandated local agency data collection; or state-level data collection. These possibilities are described in more detail in Chapter 4.

2.2.3. Need for data collection resources

The cost of collecting and maintaining data depends on the quality of data desired; the data collection approach; and the sampling scheme. A state-sponsored, yearly census of pedestrian volumes on all roadways will cost far more than sporadically collected counts submitted by volunteers. In general, it is more expensive to mount a systematic data collection program than an unsystematic one, but the former will yield more meaningful results.

2.2.4. Need for institutional commitments to automated counting

Caltrans systematic vehicle volume data collection program relies on automated vehicle counting devices (loop detectors) installed at a representative set of locations around the state. These devices collect continuous counts, and also provide essential information on daily, weekly, and seasonal variations in vehicle volumes. This information makes it possible to convert short counts of vehicle volumes taken at other sites into yearly estimates of vehicle volumes.

At this time, there are a limited number of devices capable of automatically measuring pedestrian volumes in an outdoor urban setting, and they have not yet been well-tested. This presents a challenge to the routine collection of automated pedestrian count data. However, this research has identified technologies capable of collecting pedestrian counts over a long period. Institutional commitment is needed to ensure that these devices are used to continuously monitor pedestrian volumes in certain locations.

2.3. Balancing Constraints

This chapter illustrated the fact that there are a number of possible responses to the technical and institutional challenges arising in the design of a statewide pedestrian database. These challenges are summarized in Table 2.1. The following sections will describe approaches to a pedestrian volume database that seek to balance the constraints of limited resources, uncertain institutional support, and technical complexity.

Table 2.1: Summary of technical and institutional challenges

Technical	Lack of statewide inventory of the pedestrian network; Pedestrian network is distinct from vehicle network; Pedestrian network is very large; Pedestrian movement is more complex than vehicle movement; Difficult to link accident data to specific locations; Pedestrians are found mostly in urban areas.
Institutional	Need institutional commitments to host database; fund data collection; collect data; install automated devices.

3. BUILDING AN INVENTORY

The first step in developing a pedestrian exposure database is to build an inventory of the pedestrian network. The inventory provides the framework within which pedestrian volumes and other data can be stored.

3.1. Defining the Inventory

Crucial in the development of an inventory of the pedestrian network is a clear definition of its scope:

- ✓ Does it include off-road pathways? City parks? Underground subway connections? Indoor pedestrian malls? Overpasses?
- ✓ Should sidewalks on each side of a roadway be distinguished from one another?

The answer to these questions is to some extent a function of the size of the inventory. Given a fixed amount of resources, large-scale inventories (e.g state level) will have less detail than small-scale inventories (block-level).

Consideration of the purpose of the inventory will also help define its scope. If the sole database purpose is to provide a basis for measuring pedestrian exposure to traffic accidents, then the inventory should not include off-road paths or parks where vehicles are not present. On the other hand, if the inventory serves multiple purposes, it may need to include trails and paths.

Once the inventory is defined, it will need to be broken into discrete elements. Elements could be line segments representing a length of roadway or pedestrian path; points, representing an intersection or pedestrian crossing; or a combination of both. Each element, whether it represents a point or a segment, would be assigned a unique identifier and would appear as a record within the database. Information such as segment attributes (e.g. length) and pedestrian volume would then be assigned to the element.

3.2. Inventory Source

As described in the previous chapter, the task of constructing a detailed inventory of the pedestrian network represents a major challenge. The challenge can be lessened somewhat by using existing inventories of the roadway network as a basis upon which to build an inventory of the pedestrian network. In the state of California, there are two existing inventories of the roadway network that could serve as a proxy for the pedestrian network:

- ✓ Topologically Integrated Geographic Encoding and Referencing (TIGER) Geographic Information Systems (GIS) files;
- ✓ The Caltrans Transportation System Network-Traffic Accident Analysis and Surveillance System (TSN-TASAS) database.

TigerLine roadway centerline files are produced by the U.S. Census in GIS format, and are freely available for download by county (U.S. Census Bureau, 2007). The files include all roadways, but do not include much detail on roadway geometry.

The TSN-TASAS database is the product of a recent conversion of the previous TASAS database into an Oracle (relational database) framework. The database is owned and maintained by Caltrans, and only includes state owned and maintained roads. The attributes of the two inventories are described in detail in Table 3.1.

Table 3.1: Attributes of existing California road inventories

	TSN-TASAS	TigerLine GIS
Source	Caltrans	United States Census Bureau
Availability	Data available by request from Caltrans; GIS shapefiles available internally by County, Caltrans District, and State.	Free download; Files stored by county.
Format	Recently transitioned to Oracle database; Linked to GIS shapefiles.	Transitioning to Oracle database in 2007; Linked to GIS shapefiles.
Scope	Includes only state owned and maintained roadways; Most urban streets not included; Includes roughly 20,000 intersections, 13,000 ramps, and 24,000 km of highway segments.	Includes all roadways.
Data fields (not all are listed)	Location information; Highway group (divided/undivided); Average Daily Traffic; Federal-aid system designation; Access control type.	Road name; Address range; Segment length.
Special features	Linked to SWITRS (accidents coded by postmile).	Points can be automatically geocoded using a address-coding service.

Sources: Caltrans, 2004, 2007; Bohem, 2007; Prevost, 2007; U.S. Census Bureau, 2007.

3.3. Improvements to the Inventory

Over time, the geometry of the original road-based inventory can be modified and improved to better reflect the pedestrian network. For example, in one study of pedestrian volumes, a GIS road network was modified by adding cut-throughs and pedestrian malls (Radford and Ragland, 2004).

Along with physical modifications to the geometry of the base road network, the inventory can be improved through the addition of other environmental attributes that are useful in predicting pedestrian travel.

Possible data points include pedestrian facility factors, such as short block lengths, and pedestrian accessibility factors, such as the local land use mix and development intensity. Research has shown that these factors are associated with higher rates of pedestrian travel in some neighborhoods (Cervero and Radisch, 1996).

One example of this type of detailed data collection is the Washington State Bicycle and Facility Inventory, which was conducted in 2002-2003. The state department of transportation collected extensive information on pedestrian facility characteristics and amenities, such as the presence of sidewalks and crosswalk markings, along over 7,000 miles of state-owned roadways, and input the results into a GIS-based pedestrian facility inventory (Schneider et al., 2005).

It may also be useful to pedestrian stakeholder groups to have funding information associated with some inventory elements. New pedestrian trails added to the inventory could contain a description or code indicating the major sources of funding for the facility and the funding amounts. Funding databases of this type already exist in some states and at the federal level. For example, the New Jersey Department of Transportation Bicycle and Pedestrian Project Database contains data on the location, funding source, and funding amount of bicycle and pedestrian projects in the state in a searchable web-based format (NJ DOT, 2005). Rails-to-Trails Conservancy, a national pedestrian / bicycle advocacy group, maintains a web-based searchable database of all bicycle and pedestrian projects funded through the federal Transportation Enhancements program. The database includes information on the facility location and amount of funding (RTC, 2007).

3.4. Summary of Base Inventory Advantages and Disadvantages

The following table summarizes the advantages and disadvantages of using the TIGER and TSN-TASAS inventories as a base for the statewide pedestrian network. The major advantage of TIGER files is that they include all roads (highways and local roads). The major advantage of TSN-TASAS files is that they are already state-maintained and are linked to accident records.

Table 3.2 : Summary of inventory advantages and disadvantages

	TIGER files	**TSN-TASAS**
Advantages	Includes all roadways; Allows automatic geocoding of points using address-based referencing system; Freely available and widely accessible.	Linked to accident database. Includes data on roadway geometric features; Linked to SWITRS accident records; Linked to roadway AADT.
Disadvantages	Not linked to SWITRS accident records; Not linked to roadway AADT; Does not include road geometry.	Does not include local roads where most pedestrians are present; Database not easily accessed.

4. DATA COLLECTION STRATEGIES

Once an inventory has been chosen, inventory elements, such as road segments or intersections, can be sampled for pedestrian volume. The amount and quality of data collected depends on available resources and the institution responsible for data collection. The following section describes several possible data collection approaches, each of which involves a tradeoff between expense and data quality.

4.1. Volunteer Data Collection Approach

The lowest cost mechanism for obtaining pedestrian volume data would be to provide an online repository to which local institutions (e.g. cities) or organized volunteer groups could submit previously gathered pedestrian volume data in a manner that could be viewed by others. A recent report funded by the Federal Highway Administration urged the creation of such a repository at the national level (Schneider et al., 2005).

While this would certainly provide a useful resource to researchers, and would prevent the duplication of pedestrian counts by local municipalities, it would be limited in its usefulness for systematic estimation of pedestrian exposure to risk or tracking of pedestrian facility usage over large areas. In the absence of mandates or incentives, it would be difficult to ensure that local institutions would submit data, or to ensure that data would be collected in a consistent manner.

Improvements in data quality could be obtained by requesting that agencies or volunteers follow a consistent format when collecting data. The Institute of Transportation Engineers and Alta Planning and Design have developed a standardized pedestrian data collection approach, known as the “Pedestrian and Bicycle Documentation Project,” and have tested it at a limited number of sites around the country using volunteer labor (Alta Planning and Design, 2006). Data are collected according to a consistent protocol, and data is collected at specific times of the year. However, no spatial sampling scheme is used.

4.2. Small Sample Data Collection Approach

As compared to volunteer data collection, institutionalized data collection would likely yield better data coverage, consistency, and quality, but would come at a greater expense. A state agency could organize the data collection process, or could institute a mandate requiring local jurisdictions to submit data. A recent example of such an arrangement occurred in 2003 in the New York Metropolitan Region. The New York Metropolitan Transportation Council worked with ten metropolitan area counties to develop and implement a coordinated bicycle and pedestrian count program at 100 locations throughout the metro area. Program costs, including data collection and analysis, amounted to approximately $300,000 (Schneider et al., 2005).

The cost of institutionalized data collection could be reduced somewhat by limiting the amount of data collected. Instead of attempting a wide coverage of the pedestrian network, a small number of samples of pedestrian volume could be collected and input into a model that would estimate volumes as the remaining sites. Future data collection could be used to calibrate and refine the model. This alternative is discussed in detail in Chapter 5: Pedestrian Demand Estimation.

4.3. Large Sample Data Collection Approach

The most costly data collection approach would be for a state or sub-state agency to collect a large, statistically representative sample of volumes in the pedestrian network. Caltrans' Vehicle and Data Collection Systems Unit conducts this type of systematic, frequent sampling of traffic volumes on every road segment in the state highway network. Representative sampling techniques are discussed in greater detail in the accompanying pedestrian exposure protocol report.

For reasons outlined in the “challenges” section of this report, such frequent, routine sampling, across the state would be difficult to accomplish for pedestrian volumes without a major commitment of resources. Absent such a commitment, pedestrian demand estimation techniques may be used to approximate pedestrian volumes with limited resources. Chapter 5 describes these techniques in more detail.

4.4. Collection of Additional Data Points

Along with collection of pedestrian volumes, it may be desirable to collect additional data points for inclusion in the database, either as part of the pedestrian facility inventory or as part of the volume sampling process. The need for additional data will depend on the purposes for which the database is used, such as measurement of pedestrian risk; advocacy for additional pedestrian facilities; tracking of utilitarian physical activity, and so on. The number of additional data points that can be collected will depend on the available resources and the ease of data collection.

In addition to the possible candidate data points that could be collected as part of the facility inventory process, there are several variables that could be collected simultaneously with pedestrian volume samples. For example, estimation of pedestrian risk would be made more precise if information on pedestrian age, gender, and time of day were collected along with pedestrian volumes. These variables are known to be associated with pedestrian risk (Keall, 1995).

5. PEDESTRIAN DEMAND ESTIMATION

There are several possible approaches that could be used to estimate pedestrian volumes for elements of the inventory that were not sampled directly. These approaches include (i) representative sampling; (ii) utility / route choice modeling; and (iii) multiple regression techniques. As mentioned previously, the first strategy would likely be very costly, so is not discussed in detail here. The second strategy, utility/route choice modeling, is described in a paper presented as part of this study (Radford and Ragland, 2006). This strategy is similar to trip generation techniques commonly used to model vehicle volume flows. It relies on the notion that pedestrians seek to maximize their utility when choosing routes in a transportation network. This strategy is not considered in detail in this report because it is very computationally intensive and requires a high volume of input data (Radford and Ragland, 2006).

The final approach, multiple regression/ configurational modeling, was judged to be most appropriate for extended discussion in this report, because it would allow reasonable estimation of pedestrian flows using a small number of input samples.

This report does not aim to describe a specific modeling technique, but rather describes a family of techniques that could be used to estimate pedestrian volume throughout the state using a sample of pedestrian volumes, an inventory of the pedestrian network, and a limited amount of additional data, such as population and employment densities, land uses, and so on. It also lists a series of variables that would be good candidates for inclusion in the data collection and modeling process.

5.1. Sketch Plan and Configurational Models

Simple multiple regression models, also referred to as “sketch plan” models, are commonly used in planning applications to estimate the number of pedestrians using a facility based on easily accessed data such as population and land use. The advantage of these models is that they are relatively simple to understand, are easy to apply, and yield rough estimates of pedestrian volume (FHWA, 1999).

Several examples of pedestrian sketch-plan modeling are described in the Guidebook on Methods to Estimate Non-Motorized Travel (FHWA, 1999). Some attempt to estimate the aggregate number of trips generated in an area, while others focus on estimating the flow in a specific corridor. One study, for example, used data on household population, National Household Travel Survey mode split, and the location of activity centers to estimate the number of walking trips in a specific corridor (FHWA, 1999).

The disadvantage of sketch plan methods is that they rely on assumptions about travel behavior that may not be applicable to all locations. In other words, sketch plan models applied over very broad areas (e.g. the state), would not account well for idiosyncratic local conditions that may influence walking behavior (FHWA, 1999). This issue could be dealt with by breaking the database into parts, such as counties or Caltrans districts, and using slightly different modeling techniques for each.

Another way to improve the sophistication of sketch plan modeling is to take the spatial characteristics of the travel network into account in the modeling process. For example, configurational models, such as the Space Syntax model, use travel network connectivity as a model parameter. Because it is a promising method of estimating pedestrian volumes over wide areas, Space Syntax is described in more detail below. It is also described in a paper prepared for this project and presented at the Transportation Research Board conference (Radford and Ragland, 2004).

5.2. Space Syntax Example

Used widely in Europe, Space Syntax is a modeling tool that uses multiple regression techniques to estimate pedestrian flows based on the connectivity of the pedestrian network and a limited number of additional parameters, such as population and employment density. The model analyzes the connectivity of the pedestrian network, which is input in GIS format, and develops pedestrian “movement potentials”. It then compares these potentials to a small number of samples of pedestrian volume taken at different locations throughout the network, and computes volumes for the remainder of unsampled network segments. Space Syntax is capable of accounting for up to 80 percent of the variation of pedestrian flows in urban areas (Radford and Ragland, 2006).

The city of Berkeley, California, recently used Space Syntax modeling to estimate midday pedestrian flows on every city block. Information on land uses and street network characteristics, as well as 64 pedestrian volume samples were used to predict pedestrian flows. Figure 5.1 illustrates the model output. These volumes were then combined with SWITRS data to calculate collision rates for intersections throughout the city (City of Berkeley, 2006).

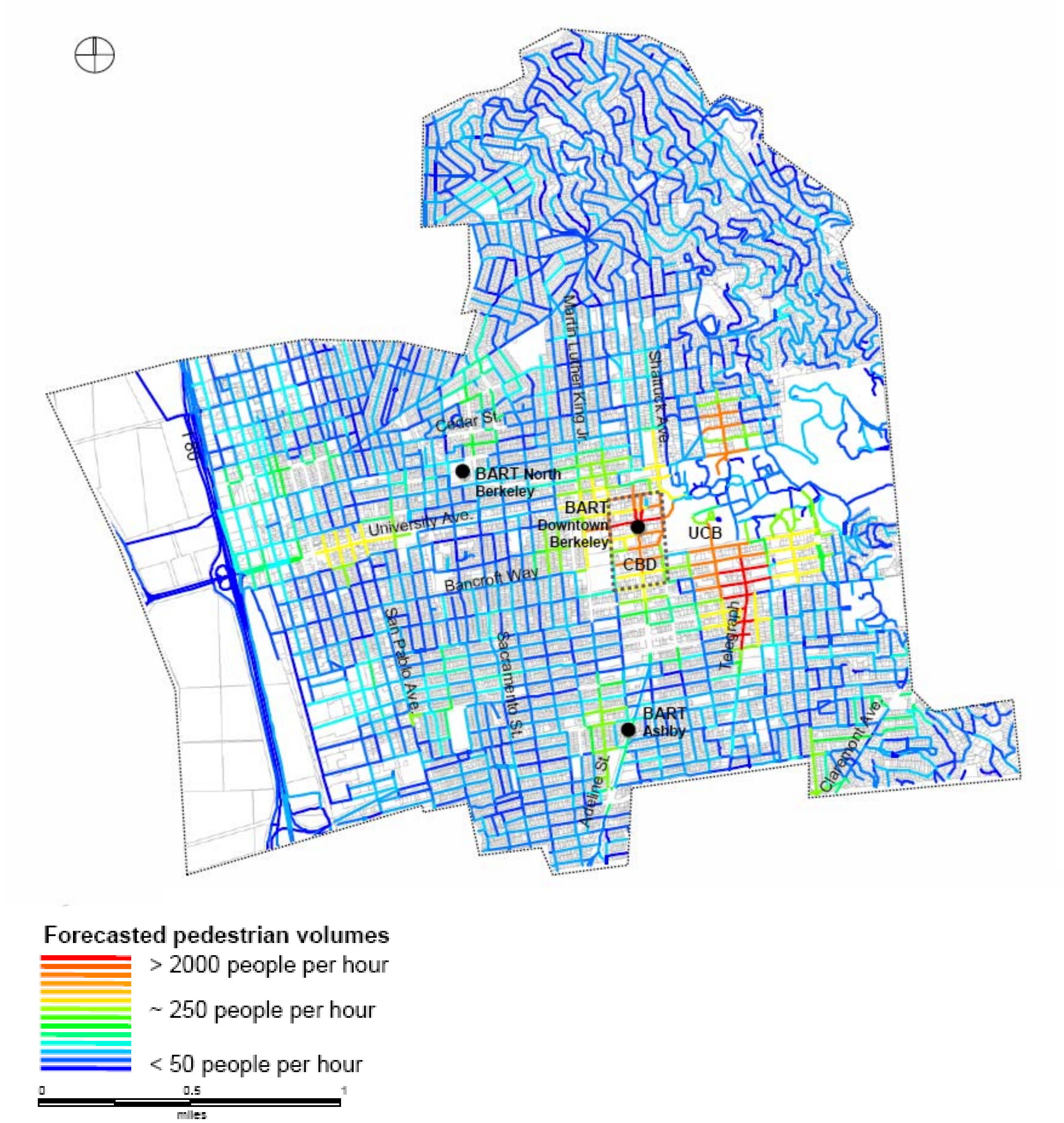


Figure 5.1: Forecasted pedestrian volumes in the city of Berkeley

6. SUMMARY AND RECOMMENDATIONS

6.1. Summary

Pedestrian volume data is an invaluable tool for safety analysts, researchers, advocates, and government agencies. In spite of this, very little systematically collected pedestrian volume data is publicly available in California or elsewhere in the United States.

This report investigates the possibility of creating a pedestrian volume database for the state of California. It identifies major technical and institutional challenges to database creation, and explores the steps that would be necessary to begin database development. These steps include selection or creation of an inventory of the pedestrian network; selection and implementation of a data collection strategy; and estimation of pedestrian volumes.

Beyond these basic steps, there are a wide range of approaches to database development. The selection of an approach depends on the purpose of the database; the available resources; and the level of data quality desired. This report presents several possible alternatives, including:

- ✓ a low-cost “data repository,” in which data is submitted on a voluntary basis by local organizations or agencies;
- ✓ a middle-cost alternative, in which data collection is institutionalized but the number of samples are limited, and modeling is used to estimate volumes at the remaining sites;
- ✓ a high-cost alternative, in which data collection is institutionalized and a large, statistically representative sample of the pedestrian network is gathered on a regular basis.

Any of these alternatives may provide useful information to pedestrian stakeholder groups. However, the higher cost alternatives will likely provide more meaningful, usable data than the lowest cost alternative. The ideal data collection strategy would be backed by long-term institutional commitment and resources.

6.2. Recommendations

Development of a statewide database is a major task that should proceed in steps, so that the form and content of the database can be revised and improved before it is fully implemented. The following steps are recommended to move the database concept to the next stage of development:

- ✓ Refine database goals and institutional responsibilities;
- ✓ Consider how database creation could connect to state policies and objectives. For example, could the pedestrian data collection program be linked to routine calculation of pedestrian risk statistics or to the allocation of funds for new pedestrian facilities?
- ✓ Select a sub-state area, such as a Caltrans district or county, in which to develop and test a pilot database. As described in this report, the database could consist of a sample of pedestrian volumes from intersections listed in the TSN-TASAS roadway inventory, or could sample portions of a GIS-based road network;
- ✓ Use the pilot project as a means to develop and test a predictive model of pedestrian volumes;
- ✓ Install automated counting devices at a small number of locations in the pilot area to collect data on temporal variation in pedestrian volumes.

REFERENCES

Alta Planning and Design. **National Bicycle and Pedestrian Document Project**. Alta Planning and Design, Inc. www.altaplanning.com/. Accessed May 21, 2006.

Boehm, B. **[Personal communication].** March 2007. Caltrans Automated Traffic Accident Surveillance and Analysis System, Division of Traffic Operations.

California Department of Transportation (Caltrans) Traffic Manual Chapter 3: Accident and Roadway Records. In: **Traffic Manual** (1996 Metric Version with updates), California, 2004. http://www.dot.ca.gov/hq/traffops/signtech/signdel/trafficmanual.htm. Accessed March 10, 2007.

California Department of Transportation (Caltrans). **California Motor Vehicle Stock, Travel and Fuel Forecast**. California Department of Transportation, Division of Transportation System Information, 2005.

California Department of Transportation (Caltrans) **Geographic Information Systems Data Library**, Office of Geographic Information Systems. http://www.dot.ca.gov/hq/tsip/TSIPGSC/library/libdatalist.htm. Accessed March 10, 2007.

Cervero, R. and C.Radisch. Travel Choices in Pedestrian Versus Automobile-Oriented Neighborhoods. **Transport Policy**. Vol 3, No. 3., 1996, pp. 127-141.

City of Berkeley. **Pedestrian Master Plan.** Draft Final Report, 2006.

Federal Highway Administration (FHWA). **Financing Federal-Aid Highways**. Publication No. FHWA-PL-99-015. Office of Legislation and Strategic Planning, 1999. http://www.fhwa.dot.gov/reports/fifahiwy/finfahwy.pdf. Accessed February 10th, 2007.

Federal Highway Administration (FHWA). **Traffic Monitoring Guide**. Publication FHWA-PL-01-021. U.S. Department of Transportation, Federal Highway Administration, Office of Highway Policy Information, 2001.

Federal Transit Administration. **National Transit Database.** http://www.ntdprogram.gov/ntdprogram/ntd.htm. Accessed March 12, 2007.

Gosling, G. and M. Hansen. **Implementation of a National Database of Air Passenger Survey Data**. Institute of Transportation Studies, University of California, Berkeley, 2006.

Hu, P. and T. Reuscher, T. **Summary of Travel Trends: 2001 National Household Travel Survey**. Federal Highway Administration, 2004.

Keall, M.D. Pedestrian Exposure to Risk of Road Accident in New Zealand. **Accident Analysis and Prevention**, Vol. 27, No. 3, 1995, pp. 729-740

Kerridge, J., J. Hine, and M. Wigan. Agent-based Modeling of Pedestrian Movements: the Questions that Need to be Asked and Answered. **Environment and Planning B: Planning and Design**. Vol. 28, 2001, pp. 327-341.

New Jersey Department of Transportation (NJDOT). **New Jersey Bicycle and Pedestrian Project Database.** 2005. http://www.state.nj.us/transportation/commuter/bike/projdatabase.shtml. Accessed March 5, 2007.

Prevost, M. **[Personal communication]**. March 2007. Caltrans Traffic and Vehicle Data Systems Unit, Division of Traffic Operations.

Raford, N. and D.R. Ragland. Space Syntax: Innovative Pedestrian Volume Modeling Tool for Pedestrian Safety. In **Transportation Research Record: Journal of the Transportation Research Board**, No. 1878, TRB, National Research Council, Washington, D.C., 2004, pp. 66-74.

Raford, N. and D.R. Ragland. Pedestrian Volume Modeling for Traffic Safety and Exposure Analysis: Case of Boston, Massachusetts. In **Transportation Research Board 85th Annual Meeting**. CD-ROM. Transportation Research Board, Washington , D.C., 2006

Rails-to-Trails Conservancy (RTC). **National Transportation Enhancements Clearinghouse Project Lists**. www.enhancements.org/projeclist.asp. Accessed February 27, 2007.

Schneider, R., R. Patton, J. Toole and C. Raborn. **Pedestrian and Bicycle Data Collection in United States Communities: Quantifying Use, Surveying Users, and Documenting Facility Extent**. FHWA, Office of Natural and Human Environment, U.S. Department of Transportation, Washington, D.C., 2005.

Schwartz, W. and C. Porter. **Bicycle and Pedestrian Data: Sources, Needs, and Gaps.** Publication BTS00-02. Bureau of Transportation Statistics. U.S. Department of Transportation, Washington, D.C., 2000.

United States Census Bureau. **Topologically Integrated Geographic Encoding and Referencing System (TIGER)**: **TIGER®, TIGER/Line® and TIGER®-Related Products.** 2007. http://www.census.gov/geo/www/tiger/. Accessed March 5, 2007.

Pedestrian Volume Modeling for Traffic Safety and Exposure Analysis: The Case of Boston, Massachusetts

Noah Raford
Space Syntax Limited
11 Riverside Studios
28 Park Street
London
SE1 9EQ
United Kingdom
n.raford@spacesyntax.com

David R. Ragland (corresponding author)
Traffic Safety Center
University of California, Berkeley
140 Warren Hall #7360
Berkeley, CA 94709
510 642 0655
davidr@berkeley.edu

Submission Date: *July 29, 2005*

Word Count: *5,812 (7,312 with tables and figures)*

Abstract Word Count: *139*

This paper examines three types of pedestrian volume models in light of their usefulness for estimating pedestrian exposure for pedestrian safety research. The need for pedestrian flow data as part of pedestrian exposure and safety analysis is outlined, and the background of each type of model is discussed. It then selects the space syntax network analysis model to estimate pedestrian volumes for the city of Boston, Massachusetts. It was found that the model was able to accurately predict pedestrian flows (r-squared 0.81, p-value < 0.0001) after incorporating distance to transit stops and major tourist attractions. These findings suggest that in addition to estimating pedestrian volumes in geographic locations where data is not available, pedestrian volume modeling can also be useful for estimating pedestrian volumes in future conditions. Planning and policy implications are discussed, as are directions for future research.

INTRODUCTION

Transportation demand modeling has a long history and a complex heritage *(1,2)*. The need to estimate the amount, type, and distribution of vehicular traffic in cities is well recognized and traffic models have played an important part in the planning of modern urban growth since the late 1950's *(3,4)*. The need and ability to model pedestrian movement is a more recent development, however, resulting from an increased interest in the public health, environmental, economic, and social benefits of walking. New advances in computational power and understanding have made such modeling approaches feasible, giving rise to the emerging field of pedestrian volume modeling and simulation.

One important field where this research is being applied is in the field pedestrian safety and exposure analysis. The Federal Highway Administration (FHWA) and the National Highway Traffic Safety Administration (NHTSA) have identified accurate pedestrian exposure as one of the least understood and most important areas of research for pedestrian planners and decision makers *(5)*. The term "exposure" originates from the field of epidemiology and is defined as the rate of contact with a potentially harmful agent or event *(6)*. Pedestrian exposure is therefore defined as a pedestrian's rate of contact with potentially harmful vehicular traffic *(7)*. In practical terms, this can be measured by pedestrian volume (as expressed in units of pedestrians per hour or per year). Pedestrian risk can therefore be defined as the annual number of vehicle pedestrian collisions divided by the annual estimated pedestrian volume at a given intersection (*7,8,9)*.

While many American cities have access to pedestrian crash data through police reports, relatively few cities have accurate estimations of pedestrian volume. Without pedestrian volume counts to determine walking rates, however, cities may be left with an incomplete picture of actual pedestrian risk. High volume intersections may experience a large number of collisions per year, for example, but they may be relatively safer per pedestrian than intersections which experience less annual collisions but also less pedestrian usage. This data mismatch often results in policy prioritization based on the "squeaky wheel" principle instead of on objective data analysis (i.e., intersections with the highest rates of collision are given attention instead of those that experience the greatest risk).

Pedestrian volume modeling offers a potential solution to this challenge. A great deal of recent literature has explored aspects of the physical and social environment that encourage or stimulate walking *(10, 11)*. Physical factors such as residential population density, mixed land use, street connectivity, and adequate pedestrian facilities have been identified as key physical variables that influence the number and types of walking trips. *(12, 13, 14, 15)*. Unfortunately little of this has been translated into practical solutions for planners in need of pedestrian volume estimation.

This paper compares three different types of pedestrian volume models available to transportation and pedestrian planners, discusses possible criteria for evaluation, and then uses one of these approaches (space syntax) to estimate pedestrian volumes for the city of Boston, Massachusetts. Space syntax was chosen because this approach builds upon previous research estimating exposure for geographic areas for which data were not available *(7)*. It attempts to extend this research by estimating future pedestrian volumes which result from changes in the urban environment. It is hoped that this might allow planners to estimate future pedestrian exposure resulting from proposed or on-going development projects in their city.

The case of the civil engineering project known as “the Big Dig” is used to test the space syntax approach for this utility. The results of this approach are then analyzed and compared to the other modeling approaches. The paper closes with a discussion of the role of pedestrian volume modeling in the transportation planning process and their potential utility for further pedestrian safety research.

OVERVIEW OF PEDESTRIAN VOLUME MODELING TOOLS

Pedestrian volume modeling has several fundamental differences from vehicle modeling. These differences pose significant challenges to traditional traffic modeling approaches and require new methods for estimating pedestrian trip behavior *(16)*. Kerridge and colleagues, note that pedestrian trips are less homogenous than vehicle trips in terms of journey purpose and their route choices are less well defined *(17)*. Pedestrian trips are also often parts of larger trips or tours of connected trips which use other modes, such as walking to or from a bus or subway stop. The pedestrian network can also be much harder to define than vehicular networks because cities and buildings have numerous pathways available to pedestrians that are not available to vehicles. Finally, many unrecorded intermediate stops or pauses can be made when traveling through urban environments on foot that cannot be made in automobiles.

The goal of a pedestrian volume model is to predict or estimate pedestrian volume based on certain assumptions about pedestrian travel including trip generation levels, mode choice, trip distribution and route choice. More generally, pedestrian volume models are mathematical models that combine existing data with key assumptions to estimate volumes in existing conditions where data are unavailable or to estimate future conditions when key variables in the model change.

A large number of simulation models have been proposed that could be useful for pedestrian exposure analysis. Different models often use different inputs and outputs and knowledge of these models are often difficult to obtain *(18,19,20,21,22,23)*. Many of these models have been developed in Europe, Japan, and the United Kingdom in journals which are less accessible for American researchers, or exist as project reports in the grey literature of government agencies and private firms.

This paper attempts to classify major developments in pedestrian volume modeling research and to discuss how they might be of use for American researchers interested in pedestrian safety, exposure, and volume modeling.

State of the Practice

Attempts to understand pedestrian movement dynamics date back nearly four decades. Most early studies focused on the behavior of pedestrians in confined circumstances such as subways, airports, or building entrances because these were easier to understand *(24,25,26)*, but others sought a broader understanding of pedestrians in urban environments such as central shopping districts *(10,11)*..

In recent years measurement tools have become more powerful and sophisticated, resulting in more useful and complex models of pedestrian movement prediction. Many of these models have been developed for specific purposes, but all share the goal of helping planners and architects create efficient, comfortable, and safe pedestrian facilities *(27)*.

For the sake of clarity, this paper divides these models into three approaches and discusses each in turn. The approaches are:

1. Sketch plan models
2. Network analysis models
3. Microsimulation (or agent based) models

The difference between each approach is their scale of application, their necessary inputs, and their most frequent outcomes. Sketch plan models focus on regional demand estimation, network analysis models focus on city-wide and neighborhood levels, and microsimulation focuses on single or a small number of streets, intersections, open spaces, or building interiors.

Although individual models often differ in their assumptions and techniques, most fall within this general typology of pedestrian models. The following section discusses each in turn relative to these criteria, citing relevant examples for each where possible.

Sketch plan models estimate pedestrian volume at the statewide or regional levels. These models were among the first models attempted by planners and researchers and use simple planning guidelines and to produce “rules of thumb” estimates of pedestrian volume based on key indicators such as square footage of office space, parking capacity, vehicular traffic movements, and movement levels in similar environments *(28)*. These models have been applied in large regional and multi zone urban environments where estimates of pedestrian volumes are desirable, but where high accuracy or more detailed estimations are not required *(29,30,31)*. Pushkarev and Zupan and Behnam and Patel were among the first researchers to attempt to forecast pedestrian volumes using this approach. They used commercial land use space and observed counts to estimate sidewalk volume levels in Manhattan and Milwaukee, respectively *(10,11)*. Swords and colleagues used population and employment density plus transit access at the statewide level to create a Pedestrian Potential Index (but not pedestrian volumes) for the State of New Jersey (*32)*. Finally, the regional land use growth model INDEX has been used in many cases to estimate regional pedestrian suitability (but not pedestrian volumes) using indicators such as population density, parcel side, and network continuity *(33).*

The second category is network analysis models. These models are more detailed than sketch plans models and can estimate volumes for street segments and intersections over an entire city or neighborhood. Although the models vary in technique, most use a variation on the four step modeling approach to generate and distribute trips based upon assumptions about the amount of walking trips in a study area and various route choice algorithms *(34,35,36)*. Ness and colleagues used this approach in their analysis of the city of Toronto, where they divided the area into traffic analysis zones and then code the links between these zones based on the street network and various “friction factors”. Trip generation and distribution was then measured to create an origin-destination matrix, and trips were then distributed using a gravity-based model *(37).* Ercolano and colleagues used traffic analysis zones, mode split assumptions based on peak vehicle volumes, and a network assignment model *(38).*

The space syntax approach uses a network modeling technique to estimate pedestrian movement potentials based on a graph “nearness” algorithm that measures route directness *(39, 40).* It then uses pedestrian counts instead of a generation and distribution phase to calibrate these relative values and convert them into actual pedestrian per hour estimates. Hillier et al. and Penn et al. found that this approach estimated pedestrian volumes in central London with an r-squared of 0.77 *(41, 42).* Stonor et al. combined distance to transit, land use composition, pedestrian crossing design, and signal phase information in a multivariate space syntax regression model of south London with 80% predictive accuracy *(43).* Raford and Ragland incorporated residential and employment densities into their space syntax model of Oakland, California, yielding city wide pedestrian volume predictions with an r squared of 0.72 when compared to observed pedestrian traffic *(7).* This approach offers a more economical way of network calibration then origin-destination surveys and has been used with relatively accuracy in hundreds of large scale real world projects in Europe and the United Kingdom.

The third and final type of approach uses microsimulation and agent based models. These models offer highly realistic simulations of small areas such as individual streets or intersections and enclosed spaces such as transit centers, airports, and malls. Microsimulation models use detailed virtual representations of their study area, either pre-determined or random origins and destinations, and specific rules for pedestrian navigation and movement to simulate thousands of individual pedestrians (or agents) in high volume conditions. Simulated pedestrians seek their destinations based on rules of movement such as avoiding collision with walls and other pedestrians and seeking the shortest route to their destination. The output of these individual interactions can then be analyzed and visualized.

Microsimulation draws heavily from the physical sciences for their rules about pedestrian behavior and are often based upon observations that crowds of people behave similar to flowing liquid in confined situations *(44,45,46).* The emphasis on confined situations and high density flows has resulted in successful application for environments such as corridors and bottlenecks *(47,48,49),* places free of automobile traffic such as subway and metro stations *(50,51,52),* and for bridges and pedestrian walkways such as those used by pilgrims to Mecca *(53).* Microsimulation and agent based models are often at the root of popular evacuation software packages such as SimWalk, Legion, and Exodus, but other models have begun to explore more complex origin destination matrices and performance in open ended urban environments *(54,55,56).*

These approaches are summarized in **Table 1.** It can be seen that each of the three approaches are applicable for different scales of analysis. Each approach also produces different outputs, with vary degrees of accuracy and utility for pedestrian safety and exposure research. The following section describes the use of one approach using a space syntax model to estimate pedestrian volumes for the city of Boston, Massachusetts.

METHODOLOGY

To evaluate the utility of pedestrian volume modeling, a space syntax approach was chosen. This model was chosen because of its utility for estimating pedestrian volumes for each street segment at the urban scale. Another factor was its success in estimating exposure in past research.

The city of Boston, Massachusetts was chosen as a case to estimate future pedestrian volumes using the space syntax model. This city was chosen becaue it is currently nearing completion of a large civil engineering project in the downtown area, colloquially named "the Big Dig". This project will effect major changes in both pedestrian and vehicular circulation, and thus offered a good "natural experiment" to test the applicability of pedestrian volume modeling for future exposure conditions. **Figure 1** displays a map of downtown Boston with the location of the "Big Dig" highlighted in grey.

To estimate the changes involved with the Big Dig construction project, a space syntax model for existing and future conditions was created. The first modeled the time period of existing conditions, based upon data surveyed in August, 2004. For this model, the majority of important highway infrastructure had been completed but the pedestrian park was still under construction and thus unopened. The second time period modeled the final conditions of the area, based upon plans for the most recent designs of the Rose Kennedy Greenway park. This reflected the conditions of what the area will be like upon its completion.

The creation of the space syntax predicative model comprised six steps:

1) Base data collection
2) Pedestrian route network modeling
3) Processing for movement potentials
4) Collection of pedestrian counts to calibrate the model
5) Addition of land use, transit, and other variables
6) Testing the accuracy of the model
7) Forecasting future pedestrian volumes based upon network change

The first phase comprised base data collection. Geographic Information Systems (GIS) data for the downtown Boston area were procured from the Boston Redevelopment Agency (BRA) as part of the first step. This included the street and sidewalk network, building outlines, aerial photos, land uses, tourist trails, and underground public transportation stops. These data were freely available and easily accessed via the BRA's website.

After collecting and compiling data, the pedestrian route network was created using a GIS. TIGER road centerline, street network data from the BRA, and aerial photographs were used to trace every possible pedestrian path and open space in the downtown area. This included each block and street segment in the downtown accessible to pedestrians, including passageways through buildings, pedestrian malls, and existing parks and public spaces. A total of 468 elements were included in the pedestrian network.

Then the pedestrian network was then processed using space syntax software to determine the relative movement potentials of each street. The MapInfo GIS software Confeego was used, which was freely available for academic use. This software converted the Boston pedestrian network into a link and node graph in order to perform a topological analysis of the mathematical nearness of each node in the network. This was then used to estimate the movement potential of each street based upon a route choice algorithm specifying pedestrian preference for the most direct pathways with the least change of direction from all origins to all other destinations. The output of this stage was a quantitative measurement of movement potentials (called "integration" in space syntax parlance) based upon the relative accessibility of each street segment in the system.

Next, a detailed field survey of pedestrian flows was conducted to determine the relationship between the relative movement potentials calculated by the space syntax software and actual pedestrian volume. Observations were made at 82 locations through out the city by a team of 9 researchers, conducted in 5 minute segments every hour between 8 AM to 8 PM. Two days of observation were made, one on a weekday and one on a weekend during the first week in August. This allowed direct measurement of morning, lunchtime, and evening peak movement, as well as general movement during other periods of the day. Pedestrian movement was found to range from 0 pedestrians per hour to over 2,000 pedestrians per hour during the lunchtime peak. The pattern of peak pedestrian movement was found to be non-normally distributed, so it was transformed using the square root of observed values to in order to create a normal distribution for further statistical analysis.

Initial correlation of the model found that movement potentials correlated relatively poorly with observed movement in some areas (r-squared = 0.55, $p < 0.0001$). In order to account for various other influences, additional variables were added to each street segment in the GIS. These variables included land use type and square footage and are presented in **Table 2.**

In order to check the accuracy of the model including relevant additional variables, the study area was divided into four neighborhoods covering the entire downtown area. A step wise multiple regression analysis (MRA) was then conducted to measure the influence of each variable on existing observed pedestrian volumes. All variables were correlated individually and then step wise in groups to determine the optimal correlation combinations, given adequate p values, t rations, and statistical validity. The accuracy of the model was found to vary between neighborhoods and was mapped to visualize the changes in correlation over the geographic area of coverage.

The last step involved converting the calibrated model into future movement forecasts. After testing the accuracy of the model, the MRA equation for each neighborhood was used to estimate the influence of each input variable (such as accessibility or land use) on the output variable of pedestrian movement. These equations were then used to change the value of variables that would change after the completion of the infrastructure project, in particular pedestrian network accessibility and land uses. The final pedestrian movement forecasts for the entire city were then estimated using the equations derived from the calibration and multiple regression stage.

FINDINGS

Base Model

The quantification of the accessibility of the pedestrian network allowed for precise measurement of the changes for pedestrian resulting from the "Big Dig". The inclusion of other variables in the MRA allowed for additional layers to be added in the space syntax model and more accurate estimates of average pedestrian flows to be derived. Finally, the use of different phases of construction in the modeling process allowed for estimates of future pedestrian volumes to be derived, which were used for pedestrian exposure estimates in future conditions.

The space syntax model was found to correlate differently for each of the city's four main neighborhoods. Pedestrian movement in the city center, including the city's financial district, was found to correlate with pedestrian movement potential, distance to the regional rail stations, and distance to underground transit stops (r squared=0.86; p<0.0001). Of these, pedestrian accessibility was found to be the most important variable (t ratio of 8.3 versus 2.5 for each of the other variables).

Pedestrian movement in the area around the proposed park itself was also found to correlated well with pedestrian movement potential (r squared=0.81; p value < 0.001.) The distance to two major tourist attractions were also found to be the explanatory variables, but pedestrian movement potential was found to be the most important variable (Although not ignorantly so, a t ratio of 6.87 was found for movement potential vs. 5.62 for distance major tourist attractions).

The space syntax model was also found to correlate well with pedestrian movement in the North End area (r squared=0.79; p< 0.01). As in the other neighborhoods, movement potential as a function of network accessibility was found to be the most important variable

Finally, the Bulfinch Triangle neighborhood was found to correlate well, but with statistically less significant results (r squared=0.85; p < 0.09). A possible cause for this lack of statistically significant correlation may be the fact that that fewer pedestrian samples were conducted in this area. This would have resulting in greater volatility in the measurement process and less accurate output. **Table 3** presents the results of all four correlation tests in each neighborhood, and **Figure 2** maps the accuracy of the model for the entire downtown Boston area, with the numbers representing the correlation co-efficient for each neighborhood.

Projected Changes in Pedestrian Flow

After testing the association between observed pedestrian volume and each variable in the space syntax model, the conditions of the model were changed to simulate the effect of the pedestrian park after its completion. **Figure 3** displays the estimated volume forecasts for each street in the downtown study area after completion.

It can be seen that several streets were found to experience major increases in pedestrian traffic, while others were found to experience less. The model found a major new axis of east west movement was likely to emerge along the State Street and the Quincy Market area, with peak movement rates up to 1,900 pedestrians per hour. The street running northwest from the regional rail station was also found to experience increased use, which was estimated to be approximately 1,450 pedestrian per hour. This was found to have a secondary impact on neighboring Washington Street, which was found to experience similar movement levels. The major north – south axis through the central business district (Federal Street) was also found to weaken slightly. This could be explained by the fact that movement was predicted to funnel east along Atlantic Avenue towards the wharf, or along Franklin Street in the same direction.

DISCUSSION

A space syntax pedestrian volume model was used to forecast future pedestrian flows in changed conditions in Boston, Massachusetts. It was found that model accurately described changes in pedestrian movement for each street and intersection in the downtown area. This suggests that such a model would be useful for providing the necessary input for a city-wide pedestrian exposure analysis, as discussed in the introduction.

Three categories of models were also discussed, of which the space syntax model was one example. Sketch plans models were found to be useful for large scale, statewide and regional estimations of pedestrian volume. Such models have the benefit of requiring little data collection and no prior training in mathematical simulation or computer modeling. They are able to offer quick estimations of pedestrian volume, but only at the aggregate level and often with questionable accuracy. Such models are also not able to assign realistic pedestrian volumes to specific streets or intersections. Because the level of detail is necessary for pedestrian exposure analysis it is argued that these types of models may be less applicable for urban pedestrian volume estimation.

Microsimulation models such as VISSIM or Legion are often extremely accurate at the site-specific level, with excellent levels of detail. Most are able to output convincing animations and graphics that allow planner a more intuitive understanding of proposed scenarios. Their major limitations are in their scale and complexity, however. Such models are currently only appropriate for site-specific simulations, covering at most an area of a few blocks or a large internal building. They also require advanced knowledge to operate, detailed data on environmental conditions, and can require significant effort to prepare and calibrate. Such models have been effectively used for simulating detailed interactions at specific intersections, such as crossing behavior between a given level of pedestrian movement and a given level of vehicle movement. But their level of detail and lack of pedestrian assignment capabilities suggests that they are less applicable for urban pedestrian exposure analysis.

Network analysis approaches such as space syntax may offer a balance between these strengths and weaknesses. Such models do not require as much data collection as microsimulation models, but also lack the level of individual detail and accuracy which they can provide. They may be more appropriate for urban exposure analysis, however, because network analysis models are able to assign pedestrian volumes to each street in large urban systems, something which is beyond the capability of most microsimulations. These reasons, and their success in past exposure analysis, lent themselves well to network analysis to the purposes of this paper.

Despite the findings of this and other research, additional research is necessary before pedestrian volume models become a widely accepted and practical solution. In particular a more extensive and rigorous inventory of pedestrian volume modeling methods and packages should be conducted. Testing of each model’s accuracy should be conducted, preferably upon a shared data set. Sensitivity analysis and a systematic application against a wider variety of cases would also improve the utility of such models. Finally, the financial and information management requirements of each model should be evaluated above and beyond their technical applicability if such models are to be widely accepted by the urban planning research and practitioner communities.

CONCLUSION

The need for pedestrian flow data as part of pedestrian exposure and safety analysis was outlined, a model was presented to address this issue, and the background of the model's use for such a purpose in existing conditions was discussed. Findings suggested that in addition to estimate pedestrian volumes in geographic locations where data is not available, pedestrian volume modeling may also be useful for estimating pedestrian volumes in future conditions. This suggests that pedestrian exposure analysis could be used as part of a city's ongoing planning process by evaluating the effects of proposed changes.

Pedestrian volume as input to pedestrian facility planning is receiving increased recognition in policy and planning circles. Interest in and understanding of pedestrian models is increasing as well. Pedestrian modeling as a field is developing past the initial stages of development and is finding practical applications in industries around the world. Although no single solution exists, practitioners are nearing the point where they will be able to select from a wide variety of modeling tools to suite any given problem.

Significantly more research is necessary before pedestrian volume modeling becomes a standard, easily available, and cheaply executed practice. In the future, hybrid models combining several approaches are likely to develop with increased flexibility and power. As this occurs, the planning, engineering, and architecture professions will likely see increased benefits from pedestrian modeling, and demand may grow for its application to a wide range of issues and challenges. If the modeling process becomes more accessible and less expensive then the true value of pedestrian simulations as a decision support system and scenario planning tool for urban planning may be realized.

ACKNOWLEDGEMENTS

The authors would like to thank the University of California Traffic Safety Center for their ongoing support and contributions. "Estimating Pedestrian Accident Exposure" funded by the California Department of Transportation under Master Agreement 65A0161. Funding for the Traffic Safety Center was provided by a grant from the California Office of Traffic Safety through the Business, Transportation and Housing Agency.

Bill Hillier, Alan Penn, and Mike Batty at the University College London also provided valuable support and feedback during this research. Space Syntax Limited provided generous support for the pedestrian observation component of this study and ongoing professional expertise in its writing.

REFERENCES

1. Bates, J. History of Demand Modeling. In *Handbook of Transport Modeling*, ed. Hensher, D. and Button, K., Pergamon, New York, 2000.

2. Newell, G. *Traffic Flow on Transportation Networks*. MIT Press, Cambridge, Massachusetts, 1980.

3. Hensher, D. and Button, K., eds. *Handbook of Transport Modeling*. Pergamon, New York, 2000.

4. Ortúzar, J. *Modelling Transport*. Wiley, New York, 1994.

5. NHTSA/FHWA Pedestrian and Bicycle Strategic Planning Research Workshops, Final Report, April, 2000.

6. Rothman K J. *Modern Epidemiology*. Little, Brown, Boston, 1986.

7. Raford, N. and Ragland, D. *Space Syntax: An innovative pedestrian volume modeling tool for pedestrian safety*. Paper 04-2977, Annual Meeting of the Transportation Review Board, 2003.

8. Leden, L. *Pedestrian risk decrease with pedestrian flow. A case study based on data from signalized intersections in Hamilton, Ontario*. Accident Analysis and Prevention, Volume 34, pp. 457-464, 2002.

9. Jacobsen, PL. *Safety in numbers: more walkers and bicyclists, safer walking and bicycling.* Injury Prevention, Volume 9, pp. 205-209. 2003.

10. Navin, P. and Wheeler, R. *Pedestrian Flow Characteristics.* Traffic Engineering. Vol. 39, 1969, pp. 31 – 36.

11. Pushkarev, B. and Zupan, J. *Pedestrian Travel Demand.* Highway Research Record. Vol. 355, 1971, Washington, D.C.

12. Saelens, B., Sallis, J., Frank, L. Environmental Correlates of Walking and Cycling: Findings from the Transportation, Urban Design, and Planning Literatures. *Annals of Behavioral Medicine*, Vol. 25, 2003, pp. 80-91.

13. Frank, L. and Pivo, G. Impacts of Mixed Use and Density on Utilization of Three Modes of Travel: Single-Occupant Vehicle, Transit, and Walking. *Transportation Research Record 1466*, TRB, National Research Council, Washington, D.C., 1994, pp. 44-52.

14. Moudon, A.V., Hess, P, Snyder, M.C., Stanilov, K. Effects of Site Design on Pedestrian Travel in Mixed-Use, Medium-Density Environments. *Transportation*

Research Record 1578, TRB, National Research Council, Washington, D.C., 1997, pp. 48-55.

15. Shriver, K. Influence of Environmental Design on Pedestrian Travel Behavior in Four Austin Neighborhoods. *Transportation Research Record 1578*, TRB, National Research Council, Washington, D.C., 1997, pp. 64-75.

16. Clifton, K. and Krizek, K. The Utility of the NHTS in Understanding Bicycle and Pedestrian Travel, *National Household Travel Survey Conference*, Washington, D.C., November, 2004

17. Kerridge, J., Hine, J., Wigan, M. Agent-Based Modelling Of Pedestrian Movements: The Questions That Need To Be Asked and Answered. *Environment and Planning B: Planning and Design*. Vol. 28, 2001, pp. 327 – 341.

18. Helbing, D., Molnar, P., Farkas, I., and Bolay, K. Self-Organizing Pedestrian Movement. *Environment and Planning B: Planning and Design.* Vol. 28, 2001, pp. 361 – 383.

19. Fruin, J. Designing For Pedestrians: A Level-Of-Service Concept. In *Highway Research Record Number 355: Pedestrians.* Highway Research Board, Washington, DC, 1971, pp. 1 – 15.

20. Moıri, M. and Tsukaguchi, H. A New Method for Evaluation of Level of Service in Pedestrian Facilities. *Transportation Research A.* Vol. 21, 1987, pp. 223 – 234.

21. Pauls, J. The Movement of People in Buildings and Design Solutions for Means of Egress. *Fire Technology*. Vol. 20, 1984, pp. 27 – 47.

22. Davis, D. and Braaksma, J. Adjusting For Luggage-Laden Pedestrians in Airport Terminals. *Transportation Research A.* Vol. 22, 1988, pp. 375 – 388.

23. Transportation Research Board. Pedestrians. In *Highway Capacity Manual Special Report 209.* Transportation Research Board, Washington, DC, Chapter 13, 1984.

24. Carstens, R. and Ring, S. Pedestrian Capacities of Shelter Entrances. *Traffic Engineering.* Vol. 41, 1970, pp. 38 – 43.

25. Hankin, B. and Wright, R. Passenger Flow in Subways. *Operational Research Quarterly*. Vol. 9, 1958, pp. 81 – 88.

26. Navin, P. and Wheeler, R. Pedestrian Flow Characteristics. *Traffic Engineering.* Vol. 39, 1969, pp. 31 – 36.

27. Hoogendoorn, S. “Pedestrian Travel Behavior Modeling”, Proceedings of the 10th International Conference on Travel Behavior Research, Lucerne, 2003.

28. Federal Highway Administration, 1999, “Guidebook on Methods to Estimate Non-Motorized Travel: Overview of Methods”, United States Department of Transportation, Publication No. FHWA-RD-98-165, Virginia.

29. 1,000 Friends of Oregon. 1993, Making the Land Use Transportation Air Quality Connection: Volume 4A, The Pedestrian Environment. Portland, OR. Available at http://www.teleport.com/~friends/Lutraq2/Docs.htm

30. Cambridge Systematics, Inc., 1994, “Short-Term Travel Model Improvements, Travel Model Improvement Program”, U.S. Department of Transportation; DOT-T-95-05, pp. 2-1 to 2-7

31. Rossi T, Lawton T, Kim K. 1993, “Revision of Travel Demand Models to Enable Analysis of Atypical Land Use Patterns”, Cambridge Systematics, Inc. and Metropolitan Service District.

32. Swords, A., Goldman, L., Feldman, W., Erlich, T., Bird, W., Analytical Framework for Prioritizing Bicycle and Pedestrian Investments: New Jersey’s Statewide Master Plan Update, Phase 2, Presentation 04-3905, Annual *Meeting of the Transportation Review Board*, 2003

33. Allen, E., “INDEX: Software for Community Indicators”, *in Planning Support Systems: Integrating Geographic Information Systems, Models, and Visualization Tools*; Richard Brail, editor; Rutgers University Center for Urban Policy Research and ESRI Press, 2001

34. Ben-Akiva M, 1985, *Discrete choice analysis: theory and application to travel demand*, MIT Press, Cambridge, Massachusetts

35. McNally M, 2000a, “The Four-step Model” in Hensher and Button, eds., *Handbook of Transport Modelling*, Pergamon, New York, New York

36. Senevarante P and Morall J, “Analysis of factors affecting the choice of route of pedestrians”, *Transportation Planning and Technology*, Volume 10, p. 147 – 159, 1986.

37. Ness, M.P., J.F. Morrall, and B.G. Hutchinson. *An Analysis of Central Business District Pedestrian Circulation Patterns*. Highway Research Record 283, 1969.

38. Ercolano J., Olson J, Spring D, 1997, Sketch-Plan Method for Estimating Pedestrian Traffic for Central Business Districts and Suburban Growth Corridors, Transportation Research Record 1578, Washington, DC

39. Hillier, B. *Space is the Machine*. Cambridge University Press, Cambridge, England, 1996.

40. Teklenburg, J A F, Timmermans H J P, and van Wagenberg, A, 1993, “Space Synax: Standardised Integration Measures and Some Simulations” Environment and Planning B 20(3) 347–357

41. Hillier B, Penn A, Hanson J, Grajewski T, Xu J, 1993, “Natural movement: or configuration and attraction in urban pedestrian movement”, *Environment & Planning B: Planning & Design*, volume 19, 29-66

42. Penn, A., B. Hillier, D. Banister, and Xu, J. Configurational modeling of urban movement networks. *Environment and Planning B-Planning & Design*, Vol. 25, No. 1, 1998, pp. 59-84.

43. Stonor, T., Arruda-Campos, M.B., Smith, A. Towards A Walkability Index. Presented at Walk21 3rd AnnualInternational Conference, Donostia – San Sebastian, Spain, 2002.

44. Schadschneider, A. Traffic Flow: A Statistical Physics Point of View. *Physica A*, Vol. 313, 2002, pp. 153 – 187.

45. Henderson, L. The Statistics of Crowd Fluids. *Nature*. Vol. 229, 1971, pp. 381 – 383.

46. Prigonine, I. and Herman, R. *Kinetic Theory of Vehicular Traffic*, American Elsevier, New York, 1971.

47. Lovas,G. Modelling And Simulation Of Pedestrian Traffic Flow. *Transportation Research B*. Vol. 28, 1994, pp. 429 – 443.

48. Timms, P. Putting Pedestrians into Network Planning Models. Paper presented at the 6th World Conference on Transport Research, Lyon, 1992.

49. Timms, P. and Cavalho, S. Inclusion of Pedestrians and Cyclists In Network Planning Models. In Proceedings of the DRIVE Conference Advanced Telematics in Road Transport, Amsterdam, 1991.

50. Annesley, T., Dix, M., Beswick, A., and Buchanan, P. Development and Application of Pedestrian Assignment Models In London Railway Station Studies. *Traffic Engineering and Control*. Vol. 30, , 1989, pp. 345 – 352.

51. Daly, P., McGrath, F. and Annesley, T. Pedestrian Speed/Flow Relationships for Underground Stations. *Traffic Engineering and Control*, Vol. 32, 1991, pp. 75 – 78.

52. Harris N G, 1991, ``Modelling walk link congestion and the prioritisation of congestion relief " Traffic Engineering and Control, volume 32(2), pp. 78 – 80

53. Selim S Z, Al-Rubeh A H, 1991, ``The modelling of pedestrian flow on the Jamarat Bridge'' Transportation Science 24 257 – 263

54. Turner A and Penn A, 2002, “Encoding natural movement as an agent-based system: an investigation into human pedestrian behaviour in the built environment”, Environment and Planning B: Planning and Design 2002, volume 29, pp. 473 – 490

55. Batty M, 2001, ``Agent-based pedestrian modeling'' Environment and Planning B: Planning and Design 28 321 – 326

56. Haklay M, O’Sullivan D, Thurstain-Goodwin M, Schelhorn T, 2001, “``So go downtown'': simulating pedestrian movement in town centres”, Environment and Planning B: Planning and Design, volume 28, pp. 343 - 359

Table 1 – Pedestrian Volume Modeling Approaches

Modeling Approach	**Description**	**Primary Uses**	**Scale of Application**	**Data Sources**	**Strengths**	**Weaknesses**
Sketch plans	Estimation of aggregate pedestrian volumes based on population level statistics	Providing regional and city wide estimates of pedestrian movement for large scale planning studies	State and regional level	Census statistics, land uses, movement samples	Simple, easily available data	Inaccurate, insufficient detail
Network analysis	Large scale estimates of pedestrian volume based on route choice assumptions and medium level urban modeling	Urban pedestrian volume modeling, exposure analysis	Urban and neighborhood level	Road and pedestrian network, Census statistics, land uses, movement samples	Large geographic coverage, good detail, reasonable accuracy, limited data requirements	Less complex then microsimulation models
Microsimulation or agent based	Simulation of individual pedestrian movement in crowds (“agents”) based on complex behavioral rules and environmental modeling	Evacuation simulation, movement in confined environments (train stations, airports, malls)	Site specific level (individual streets, intersections, and enclosed environments)	GIS and CAD boundary layers for buildings and streets, origin and destination matrices, movement samples, rule based movement algorithms	Highly accurate, detailed, and visually communicative	Complex, steep learning curve, significant initial data requirements, smaller geographic coverage

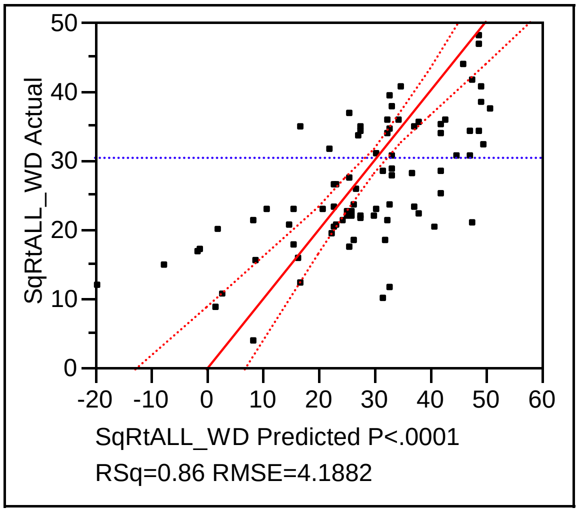
Figure 1 – Location of the "Big Dig" Project and Study Neighborhoods in Downtown Boston, Massachusetts

Table 2 – Additional Variables Included in the Space Syntax Model

Variable	Variable Type
Floor space per street	Continuous
Land use	Ordinal based on use (residential, retail, commercial, government, mixed)
Distance to transit	Ordinal (1 through 4 based on walking distances of 50, 250, 500 or 1,000 yards)
Distance to regional rail	Ordinal (1 through 4 based on walking distances of 50, 250, 500 or 1,000 yards)
Distance to tourist attractions	Ordinal (1 through 4 based on walking distances of 50, 250, 500 or 1,000 yards)
Pedestrianization	Dummy (0 or 1)
Tourist Trail	Dummy (0 or 1)

Table 3 – MRA Correlations between Observed Pedestrian Movement and Estimated Volume for Four Neighborhoods in Downtown Boston

Southern Area

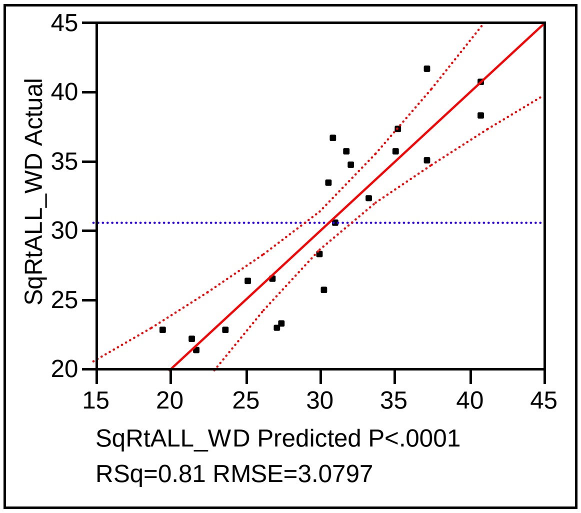


Scaled Estimates

Continuous factors centered by mean, scaled by range/2

Term	Scaled Estimate	Std Error	t Ratio	Prob>|t|
Intercept	30.368126	0.892923	34.01	<.0001
MovPotential	17.313265	2.08278	8.31	<.0001
TransitDist	5.8015945	2.500096	2.32	0.0323
Rail Dist	-6.697017	2.438154	-2.75	0.0133

Wharf / City Area

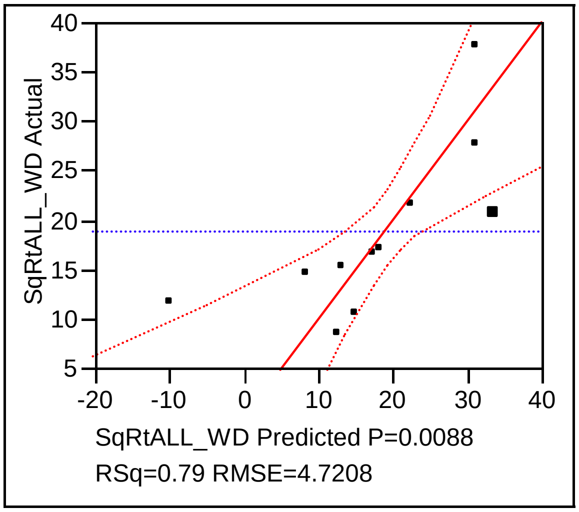


Scaled Estimates

Continuous factors centered by mean, scaled by range/2

Term	Scaled Estimate	Std Error	t Ratio	Prob>|t|
Intercept	30.501294	0.656603	46.45	<.0001
MovPotential	7.3106479	1.063979	6.87	<.0001
AttractDist	-6.843801	1.217462	-5.62	<.0001
AttractDist	-3.062365	1.577548	-1.94	0.0681

North End

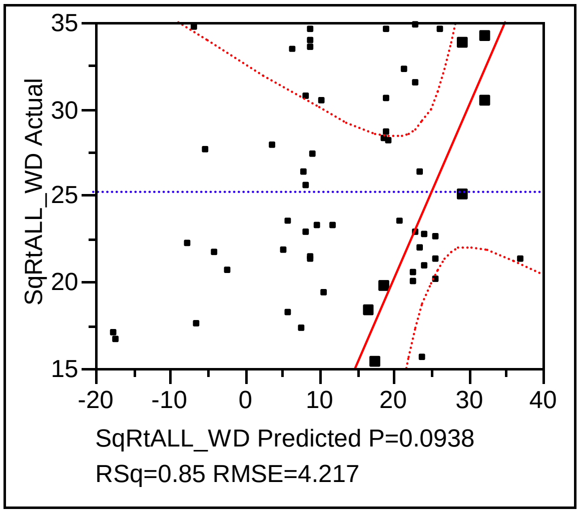


Scaled Estimates

Continuous factors centered by mean, scaled by range/2

Term	Scaled Estimate	Std Error	t Ratio	Prob>|t|
Intercept	18.846429	1.573602	11.98	<.0001
MovPotential	11.606795	3.842703	3.02	0.0234
AttractDist	0.2665158	2.585756	0.10	0.9213

Bulfinch Triangle Area



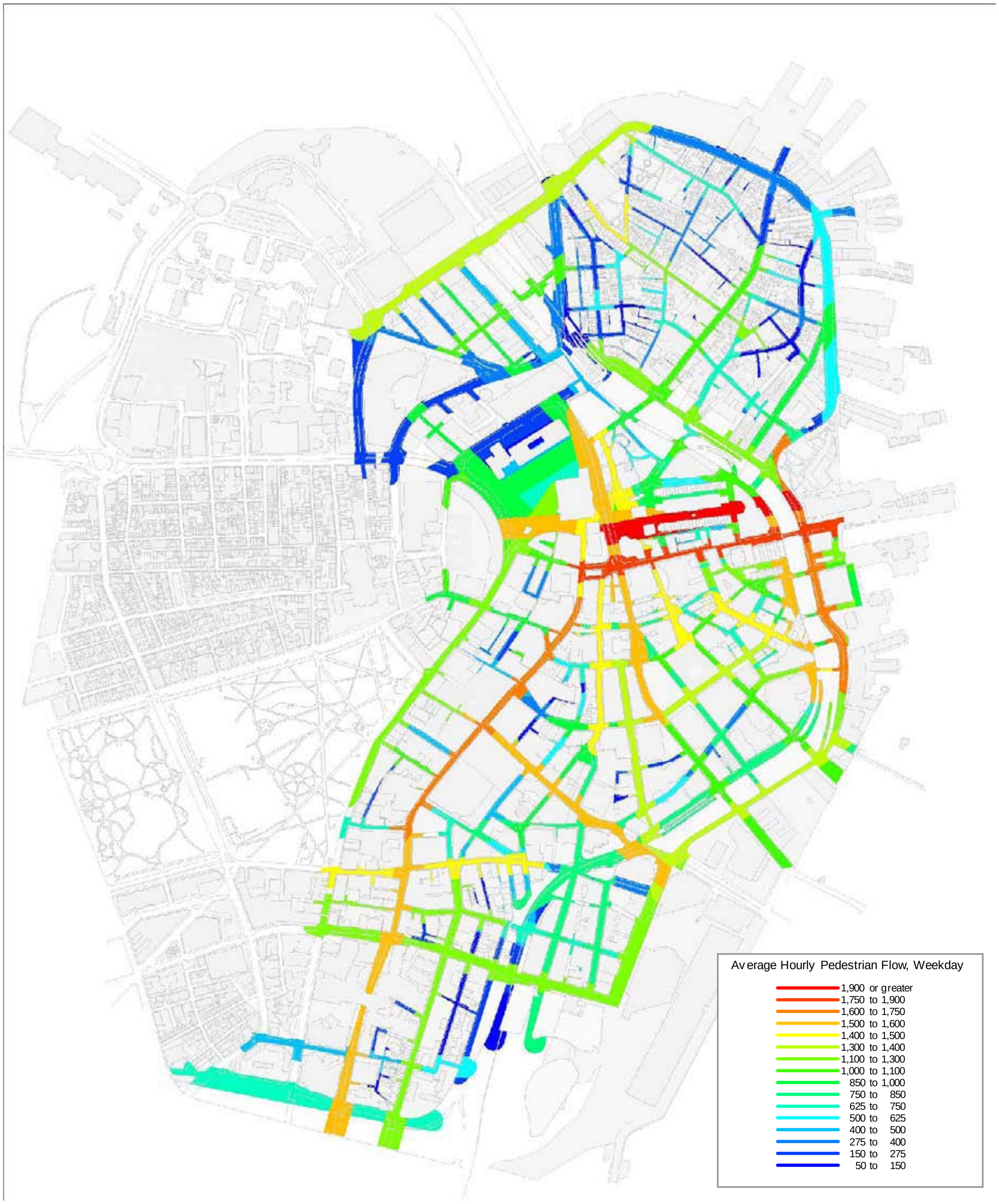
Scaled Estimates

Continuous factors centered by mean, scaled by range/2

Term	Scaled Estimate	Std Error	t Ratio	Prob>|t|
Intercept	25.258017	1.593871	15.85	0.0005
MovPotential	-1.870684	5.419837	-0.35	0.7528
Rail Dist	-1.607209	2.158239	-0.74	0.5105
Transit Dist	-6.965615	3.872895	-1.80	0.1699

Figure 2 – Map of the Accuracy of the Space Syntax Model for Downtown Boston and with Detail for Four Neighborhoods

Figure 3 – Pedestrian Volume Forecasts for Future Conditions in the Study Area



Alameda County Pedestrian and Bicycle Counting Protocol

Robert Schneider, Lindsay Arnold, Noah Raford, David Ragland
University of California Traffic Safety Center
DRAFT—April 2008

Background

This document describes the methods that will be used to collect pedestrian and bicycle counts at a sample of roadway intersections in Alameda County. The theoretical basis for this research design is described in a white paper: *Draft Strategy for Countywide Pedestrian Volume Modeling: CalTrans State Highway System for San Diego County*[1]. Guidance for selecting appropriate methods of data collection and representative counting locations is also drawn from Chapters 4, 5, and 6 of *Estimating Pedestrian Accident Exposure: Protocol Report*[2].

There are two immediate purposes of this counting effort: 1) obtain a sample of counts that can be used as a basis for predicting the number of pedestrians and bicyclists at all 531 intersections of CalTrans roadways in the county, and 2) demonstrate that the data collection and modeling methods used in this pilot study have the potential to be applied to CalTrans roadways statewide. Ultimately, the predicted pedestrian and bicycle volumes can be used to represent exposure in a crash risk analysis. This will allow CalTrans and Alameda County to evaluate and prioritize pedestrian and bicycle safety needs more accurately at each intersection.

A longer-term purpose of this project is for the Alameda County Transportation Improvement Authority (ACTIA) to establish the baseline counts and methodology for a pedestrian and bicycle counting program in the County. The methods used in this effort can be repeated by the County at regular intervals to track changes in pedestrian and bicycle activity over time.

The budget available for this counting effort is approximately $40,000 ($15,000 from CalTrans and $25,000 from ACTIA). Therefore, the research design is intended to be the most efficient method for gathering a sample of counts and estimating pedestrian and bicycle volumes throughout Alameda County within this budget (see Appendix A).

Research Design

Pedestrian and bicycle counts will be taken at a sample of approximately 50 intersections throughout Alameda County[3]. Thirty of these intersections will be at CalTrans roadway intersections (intersection of a CalTrans roadway with any other roadway or trail), and 20 will

---

[1] Raford, Noah. *Draft Strategy for Countywide Pedestrian Volume Modeling: CalTrans State Highway System for San Diego County*, Drafted for the University of California Traffic Safety Center, August 25, 2007.

[2] Greene-Roesel, R., M.C. Diogenes, and D.R. Ragland. *Estimating Pedestrian Accident Exposure: Protocol Report,* University of California Traffic Safety Center, Prepared for the California Department of Transportation, Task Order 6211, March 2007.

[3] Additional partnerships with the Alameda County Transportation Authority or other organizations may increase the number of locations and time periods that can be counted in separate, follow-up studies.

be at other intersections along major (arterial and collector) roadways in the county[4]. Field data collectors will take manual counts at 50 locations, while five infrared sensors will be used to count pedestrians continuously near five to 15 of these locations and two in-pavement loop sensors will be used to count bicyclists continuously near two of these locations. Manual counts will be taken during specific observation periods during March, April, and May 2008. Observations will be made on one weekday (Tuesday, Wednesday, or Thursday) and on one Saturday for each location. Each observation period will be from either 12:00 p.m. to 2:00 p.m. or 3:00 p.m. to 5:00 p.m. The infrared sensor counts will gather 24-hour data continuously for over an entire year (March 2008 to March 2009) to develop factors for estimating full-day volumes from the sample period manual counts.

The sections below provide more detailed descriptions of the users that will be counted and how and when the counts will be taken. They also include reasons for selecting the particular methods for this study.

Types of Users

The counts will focus on pedestrians and bicyclists. There may also be Segway users and in-line skaters that pass a counting location, but these users will not be documented. Automobile counts will not be taken in the field. Average Daily Traffic (ADT) values from the CalTrans TASAS database will be used to account for motor vehicle exposure at each intersection. While it would be ideal to count the automobiles crossing each intersection during the pedestrian count period to have a more direct value for comparison, additional data collectors and budget would be needed to collect these data. ADT values for non-CalTrans roadways will be collected from local jurisdictions.

User Characteristics

Because data collectors will be in the field, there is also potential to collect data about pedestrian and bicyclist characteristics. However, any additional observations will require additional effort, which may decrease the accuracy of the counts. Therefore, only three characteristics will be observed for each user: type (pedestrian and bicyclist), gender (female or male), and location within the intersection (e.g., specific leg of the intersection the user is crossing—north, east, south, or west)[5]. While these three characteristics will be observed by field data collectors, only the type of user will be collected from the infrared sensors.

Observations of pedestrian, bicycle, and driver behaviors can be useful for identifying possible safety problems at particular intersections. For example, pedestrians crossing against a traffic signal, bicyclists riding in the wrong direction, or drivers not yielding to pedestrians in a crosswalk can indicate the need for particular engineering, education, and enforcement treatments. However, making accurate assessments of behaviors requires very specific definitions and well-trained observers. While the scope of this project does not allow for behavioral observations, they could be considered at another point in the safety evaluation process.

---

[4] A sample size of 30 out of 531 CalTrans intersections will yield a 17% margin of error in the volumes estimated from the final model, assuming a confidence level of 95% and response distribution of 50%.

[5] Age category (under age 18, 18 to 64, and 65 and older) will not be collected due to the complexity of this observation. It would require additional time and concentration for data collectors to determine and record this information.

Count Sites

Manual counts will be taken when pedestrians and bicyclists cross the roadway at or near an intersection location (see Appendix B: Data Collection Sheet). While pedestrian counts along sidewalk segments are important for planning and prioritization, this study is evaluating intersection exposure to vehicle-pedestrian collisions, so it focuses on roadway crossings at intersections. Any pedestrian or bicyclist crossing within a crosswalk or within 50 feet of either side of the crosswalk will be counted (this includes bicyclists in the roadway making "vehicle-style" crossings). Each leg of the intersection will be counted separately and summed to derive the total pedestrian or bicycle volume. At "T-intersections", there are only three roadway crossings. However, pedestrians using the sidewalk on the fourth side of the intersection will still be counted like the other three legs. This will make it possible to make direct comparisons between the total intersection volumes at 3- and 4-way intersections[6]. Note that a single user could be counted multiple times at the same intersection if he or she crosses more than one leg of the intersection. In addition, right-turning pedestrians and bicyclists on the sidewalk will not be counted because they do not cross the roadway. It is important to count crossings of all intersection legs separately because each time a pedestrian or bicyclist makes a crossing, he or she is exposed to crash risk. "Vehicle-style" bicycle turn movements will be classified differently than pedestrians. This is because bicyclists turning right and left from travel lanes do not cross either roadway directly. Twelve possible "vehicle-style" bicyclist movements at four-way intersections will be recorded, including north leg to south leg (through), north leg to west leg (right), north leg to east leg (left), east leg to west leg (through), and so on.

Midblock crossings (more than 50 feet from the intersection crosswalk) will not be observed during this analysis. Taking midblock counts would require data collectors to focus simultaneously at the intersection and further down all approaching roadways. This is extremely difficult to do accurately without additional data collectors at midblock locations. Future studies should examine the question of midblock pedestrian crash risk using data collectors who focus only on midblock crossing counts.

Automated 24-hour pedestrian counts will be taken by sensors at sidewalk locations near 13 of the study intersections (for pedestrians) and on approaching roadways near two of the study intersections (for bicycles). While it would be ideal to parallel the manual counting effort by taking counts at roadway crossings, the sensor technology used for this study is not capable of doing this[7]. However, the continuous automated counts will be used to identify daily and weekly variations, so the adjustment factors that are developed from them are assumed to be accurate for adjusting the counts at nearby the roadway crossings. The automated counters should not be placed adjacent to bus stops or building entrances that may have peaking patterns that are significantly different than the intersection itself (representative locations will be determined by field visits). Note that the automated counts are expected to undercount

---

[6] A separate "T-intersection" variable will be included in the analysis to test if the intersection configuration has an impact on the overall count, all else equal.

[7] Since the automated counts are not directly comparable to the manual counts, the study team may do a brief analysis to compare the manual and automated counts at several different locations to see if the difference between the two types of counts varies between sites.

pedestrians and bicyclists by approximately 10 to 15 percent, a factor that will be accounted for in the analysis[8].

Extensive counts of pedestrians along sidewalks are important for urban planning applications, but they are not included in this study because they are not a direct measure of exposure for vehicle-pedestrian collision analysis. One exception is where sidewalks cross driveways or alleys. These are conflict points where pedestrian crashes occur. However, these locations are not included in this study because traffic volumes for driveway entrances and exits are not available in CalTrans or other databases. Future studies could collect counts for vehicles, pedestrians, and bicyclists at driveway conflict points.

Counting Methods

Teams of data collectors from a project consultant (Population Research Systems—PRS) will take manual counts at the 50 sample intersections. Between one and four data collectors will be used at each intersection, depending on expected levels of pedestrian and bicycle use. Five active infrared sensors (EcoCounter Pyroelectric Sensor) will be used to take continuous, automated pedestrian counts at sidewalk locations near sample count locations[9]. An automated sensor (EcoCounter Zelt Inductive Loop) will also be installed in the pavement to take continuous 24-hour bicycle counts at two intersection approaches.

It will be necessary to do an initial pilot test of the automated counters to ensure that they operate and count properly. The automated count portion of the study is critical because it will allow the research team to develop factors for converting hourly manual counts to 24-hour, weekly, monthly, or annual volume estimates.

For both manual and automated counts, raw data will need to be entered into electronic spreadsheets and tabulated. It is assumed that approximately one hour of data entry will be needed for every ten hours of manual counts (this task could be done by either PRS or TSC). Time will also be needed to install the counters and convert raw data from automated sensors to spreadsheets (this task will be done by TSC).

Observation Periods

Manual counts will be taken either between 12 p.m. and 2 p.m. or between 3 p.m. and 5 p.m. Two back-to-back time periods were chosen so that the data collectors' could do two sites on the same day and use their time most efficiently. The early afternoon observation period corresponds with lunchtime peak pedestrian activity and the late afternoon observation period corresponds closely to the weekday pedestrian peak travel period for aggregate national data. Within each counting period, data collectors will note 15-minute time increments. This will make it possible to identify any peaking trends within the two-hour observation period. While the two-hour counts should capture some fluctuations in pedestrian and bicycle activity, they may not necessarily capture the peak hour for pedestrian or bicycle travel in each location. Factors developed from automated count data will be necessary to convert the two-hour counts

---

[8] Since the automated counters are known to undercount pedestrians, the study team may do several additional manual counts at counter locations with different characteristics to establish an accurate undercounting adjustment factor.

[9] If a separate follow-up study is conducted, a competing technology (Trailmaster) may be installed at two of the same infrared sensor pedestrian count locations.

to 24-hour counts. A two-hour counting period was chosen rather than a longer counting period (e.g., 10-hour day count or morning, mid-day, and evening counts) because of the limited budget available for field data collection.

Because the manual counts will capture only two hours of activity, daily, weekly, monthly, and seasonal variations will be noted through the automated counting process. These variations will be captured by taking automated counts throughout the one-year period between December 2007 and December 2008.

Observation Schedule

To capture daily variation in pedestrian and bicycle activity during the week, manual counts will be taken on one weekday (Tuesday, Wednesday, or Thursday) and one Saturday at each site. The counts will be taken in March, April, and May 2008. There are 51 possible weekdays and 13 possible Saturdays for counting during this period. In order to complete the counts at all 50 sites within the three-month timeframe, data collectors must take counts at two locations on each weekday and four locations on each Saturday. Because of the intensive data collection needs on Saturdays, these counts may need to continue for several additional weeks.

This counting strategy assumes that travel patterns on Tuesday, Wednesday, and Thursday will be similar. Saturday and Sunday were assumed to have different travel patterns that would affect the analysis if half of the locations were observed on Saturday and other half of the locations were observed on Sunday. Saturday was chosen as the designated day for weekend data collection.

Optimal deployment of the automated counters is critical. The available budget includes only five automated pedestrian counters. However, the hourly, daily, weekly, and seasonal variation from these sensors will be used to represent the pedestrian activity variations at all 50 sample sites. While there will be differences in peaking patterns between sites, the study will assume that the hourly variations in pedestrian activity can be classified into several distinct categories[10]. In order to represent the different categories of peaking patterns accurately, it would be beneficial to use the five automated sensors in as many different locations as possible.

Several options for deploying the five infrared sensors were considered, and the preferred strategy is to move four of the sensors between different locations on a regular basis and to keep one of the sensors in the same location for the entire year. Four of the sensors will be moved in a circuit between a first set of locations, a second set of locations, and a third set of locations on a monthly basis (i.e., four locations will be counted in March, June, September, and December; four locations would be counted in April, July, October, and January; and four locations would

---

[10] Several previous studies have categorized hourly variations in pedestrian activity, including:
1) Cameron, R.M. Pedestrian Volume Characteristics. Institute of Transportation Engineers Compendium of Technical Papers, 1976.
2) Davis, S.E., L.E. King, and D.H. Robertson. Predicting Pedestrian Crosswalk Volumes. Transportation Research Record 1168, Transportation Research Board, 1988, pp 25-30.
3) Hocherman, I., A.S. Hakkert and J. Bar-Ziv. Estimating the Daily Volume of Crossing Pedestrians from Short-Counts. Transportation Research Record 1168, Transportation Research Board, 1988, pp. 31-38.
4) Zeeger, C.V., R. Stewart, H. Huang, P.A. Lagerwey, J. Feaganes and B.J. Campbell. Safety Effects of Marked versus Unmarked Crosswalks at Uncontrolled Locations: Final Report and Recommended Guidelines. Publication FHWA–HRT–04–100. Office of Safety Research and Development, Federal Highway Administration, 2005.

be counted in May, August, and November, and February). A total of 12 sites will be counted by these four sensors. Each site will have continuous count data for one whole month, every third month. The count data will document hourly, daily, and weekly volume fluctuations. In addition, seasonal trends for each location will be apparent from the volumes observed every third month, though there will be less certainty in these trends than if the counts were taken continuously for the entire year. In particular, there is a reasonable chance that the peak month will not be captured for particular sites. This strategy also involves the challenge of moving the sensors on a monthly basis.

Keeping the final counter at the same location will provide a constant count that represents every hour of the year. This will make calculating the hourly, daily, weekly, and seasonal variation at that site relatively simple. It will also provide a continuous count for comparison with the monthly counts from the movable sensors.

Number of Intersections

Fifty intersections will be counted manually by data collection teams, 13 sidewalk locations will be observed using automated technology, and two roadway locations will be counted using in-pavement automated technology. This number of intersections is the maximum number that can be counted based on the budget available for hours needed to travel to and count at sites, set up and test equipment, and enter raw data into spreadsheets (see Appendix A: Project Budget Summary).

Selection of 30 Sample CalTrans Intersections

The primary objective of the data collection effort is to be able to estimate pedestrian and bicycle volumes at a target population of 528 CalTrans highway intersections in Alameda County[11]. To develop an accurate model for estimation, the sample of 30 CalTrans intersections must be selected strategically (note that the method for selecting 20 additional non-CalTrans intersections is described in the section below)[12].

A stratified random sampling technique will ensure that the 30 CalTrans intersections will have a wide range of variation within three key variables: gross population density (within a ¼-mile radius), median income (of population within a ¼-mile radius)[13], and commercial retail land use intensity (of nearby properties)[14]. The three variables chosen for sampling have been shown to be correlated with levels of non-motorized transportation activity in previous

---

[11] There are a total of 528 CalTrans intersections in Alameda County. Originally, a set of 534 intersection points was reviewed. Three of the points were within ¼-mile of the country boundary. Since the sampling scheme uses GIS data within a ¼-mile radius to categorize the built environment characteristics surrounding each intersection, they will not be considered in the analysis (GIS data were not gathered for areas outside the county due to budget constraints). In addition, three other intersections were represented by duplicate GIS points (within 20 meters of the same point). Therefore, 528 intersections were used in the analysis.

[12] Intersections with slight offsets (intersections of the roadway centerlines are less than 20 meters apart) are counted as a single intersection for selection and analysis purposes.

[13] Data on median household income in 1999 is provided by the US Census 2000. The measure used in this analysis is the average of all median household income values (weighted by the proportion of each census block group within the ¼-mile buffer).

[14] Commercial retail land use intensity will consider parcel land uses adjacent to the intersection (within 1/10-mile) and in the neighborhood of the intersection (within ¼-mile). The retail intensity score will be calculated as the sum of the total number of retail uses within a 1/10-mile radius plus the total number of retail uses within a ¼-mile radius of the site. Note that this gives twice as much weight to intersections within 1/10 mile of retail uses.

studies[15]. The three categories are also specific enough to select a wide range of sample intersections, but remain relatively broad so that the selection process can be done with the available resources and budget. More factors will be considered when developing models to estimate the actual pedestrian volumes, including proximity to transit (within a ¼-mile radius)[16], proximity to multi-use trails[17], and street network density[18].

Spatial autocorrelation between the count locations is another important issue to consider. To apply the predictive modeling methods appropriately, each of the count locations should be independent of all of the other locations. In other words, the pedestrian counts at two adjacent intersections in a commercial corridor may not be independent if some of the people who walk through one intersection also walk through the other. While there is no way to know the amount of interdependence between the sample locations, autocorrelation can be minimized by ensuring that the random sample locations are not clustered geographically in certain parts of Alameda County. Because of the constraints placed on the three key variables, the sample of locations is likely to be spread through sample selection and clustering is unlikely to occur. Therefore, an additional test of the sample will be to ensure that no intersection in the sample is within ¼-mile of any other intersection in the sample.

In order to obtain an adequate range for each variable, ratings of “high”, “medium”, and “low” will be given to each of the 528 locations for each of the three key variables. The “high” category will represent the top third of locations, “medium” will represent the medium third, and “low” will represent the lowest third. For example, the 176 intersections with the densest populations within a ¼-mile radius will be given a rating of “high” for the population density variable (see Table 1. Key Variables and Sample Strata for CalTrans Intersection Selection).

Table 1. Key Variables and Sample Strata for CalTrans Intersection Selection

---

[15] Several papers have reviewed the literature on factors associated with pedestrian and bicycle activity, including:
1) Ewing, R. and R. Cervero. “Travel and the Built Environment: A Synthesis,” Transportation Research Record 1780, pp. 87-113, 2001.
2) Handy, S. Critical Assessment of the Literature on the Relationships Among Transportation, Land Use, and Physical Activity, Transportation Research Board Special Report 282, Available online: http://trb.org/downloads/sr282papers/sr282Handy.pdf, 2005.
3) Krizek, K. “Operationalizing Neighborhood Accessibility for Land Use-Travel Behavior Research and Regional Modeling,” Journal of Planning Education and Research, Volume 22, pp. 270-287, 2003.
4) Shriver, K. Influence of Environmental Design on Pedestrian Travel Behavior in Four Austin Neighborhoods. Transportation Research Record 1578, TRB, National Research Council, Washington, D.C., 1997, pp. 64-75.

[16] Proximity to transit was not included in the selection process. This factor will consider nearby rail stations and bus stops. Rail stations will be given a subjective weight of 10 points, and bus stops will be given a weight of 1 point for each route served. The proximity to transit score will be the sum of all transit access points within a ¼-mile radius of each intersection.

[17] Proximity to multi-use trails was not included in the selection process. The multi-use trail proximity score will be calculated as the sum of the total number of retail uses within a 1/10-mile radius plus the total number of retail uses within a ¼-mile radius of the site. This will give twice as much weight to intersections within 1/10 mile of a trail.

[18] Street network density was not included in the selection process. The coding of street types in the Alameda County streets GIS layer did not have separate attributes for freeway ramps. Therefore, they could not be removed from consideration in the analysis. In addition, since many arterial boulevards were represented by two centerlines, they would have been double-counted. This would give an inaccurate representation of street network density. Space syntax variables, such as node connectivity, sight distance along the roadway, or number of turns required to reach other destinations on the roadway network, are not included in the sampling scheme due to the available budget. However, these variables may also be tested during the analysis process.

In order to ensure adequate variation in the characteristics of the count locations, the CalTrans intersections were stratified into the categories in the table below:

	Population Density (persons/sq. mi.)	**Median Income (1999 Dollars)**	**Commercial Retail Land Use Intensity**
High (highest third of locations)	11,258 to 19,545	53,444 to 166,900	38 to 138
Medium (middle third of locations)	7,689 to 11,257	39,854 to 53,443	14 to 37
Low (lowest third of locations)	50 to 7,688	9,000 to 39,853	0 to 13

Since the selection will use three categories for each of three variables, 27 different strata will be used for selection (e.g., Population Density = High, Median Income = High, Commercial Retail = High; Population Density = High, Median Income = Medium, and Commercial Retail = Low; etc.). One intersection was to be chosen from each of these strata (representing each of the 27 combinations of characteristics). However, two of the strata did not contain any intersections (Population Density = High, Median Income = Low, Commercial Retail = Low and Population Density = Medium, Median Income = Low, Commercial Retail = Low). Therefore, 25 intersections were chosen from the 25 strata that were represented. Five additional intersections were chosen randomly from the remaining set of intersections to complete the sample of 30 locations for analysis (each of these five intersections will be taken from different strata). The primary sample of 30 intersections was reviewed to find any locations that were within ¼-mile of another selected location. A backup set of 30 intersections was also selected using the same stratified-random sampling criteria. The backup intersection for the corresponding category was used as a substitute if there was a problem with the initial selection. Problems encountered included intersections located within ¼-mile of another selected intersection, being under construction, or having a configuration that was not conducive to pedestrian counting, such as being a grade separated intersection.

A number of other sampling methods were considered. If a simple random method were used, strata that contain a larger number of intersections would also have a greater chance of being selected than strata with fewer intersections. In addition, a small random sample of 30 locations could miss representing locations with the highest pedestrian volumes. Alternatively, selecting the 30 locations by convenience (such as locations suggested by local experts or community members) would introduce bias into the method.

Selection of 20 Non-CalTrans Intersections

A deliberate method was also be used to select 20 additional intersections that are on other major (arterial and collector) roadways in Alameda County[19]. Alameda County has 7,488 intersections along major roadways. However, several criteria were used to narrow the number of intersections that could be selected for analysis. The roadway intersections were required to have a population density of at least 50 residents per square mile (485 intersections in low-density areas were removed). This criterion was used because low-density areas are likely to

---

[19] The research team will also select a set of 20 backup non-CalTrans locations that meet the same criteria. If any of the 20 primary locations are unsuitable for counting (such as being under construction or being within ¼-mile of another selected intersection) a replacement location with identical characteristics (except location) from the backup set will be used.

have very sparse, variable pedestrian activity, which is difficult to model. In addition, intersections were required to be more than ¼-mile from an adjacent county (101 intersections close to the county border were removed). After establishing these criteria, 6,902 roadway intersections were considered for selection. At the request of ACTIA, 36 major roadway/trail intersections were also included in the sampling frame. Therefore, a total of 6,938 intersections were considered for selecting the 20 non-CalTrans intersections.

The 20 points were sampled using a series of random selections according to specific criteria. These criteria were established in coordination with ACTIA. First, four central business district (CBD) intersections were chosen randomly from all possible CBD locations. Second, two roadway/trail intersections were chosen. Finally, 14 additional intersections were chosen randomly from the remaining locations. This set of locations was reviewed to ensure that at least three of the 50 total intersections sampled were in each of the four county planning areas (North, Central, South, and East). In addition, the locations were checked to make sure that none of the 20 intersections could be within ¼-mile of any other intersection being counted (including the 30 CalTrans intersections). Selecting the 20 non-CalTrans intersections increased the variation in roadway and surrounding land use types being sampled and provided important contextual information for pedestrian exposure modeling.

Characteristics of the 50 Selected Intersections

The 50 selected intersections have a wide variety of characteristics. The selection process ensured that the intersections were in areas with a variety of population densities, income levels, and access to commercial retail. While there is large variation in these characteristics, the average values are similar to the county as a whole (see Table 2). Other characteristics of the selected intersections include:

- 9 intersections within ½ mile of a Bay Area Rapid Transit (BART) station
- 4 trail/roadway intersections
- 13 intersections including bicycle lanes on at least one approach
- 6 central business district intersections
    - Oakland (4)
    - Hayward
    - Fremont
- A variety of other characteristics, including number of travel lanes, traffic volumes, speed limits, median islands, curb radii, traffic signals, pedestrian signals, on-street parking, nearby land uses

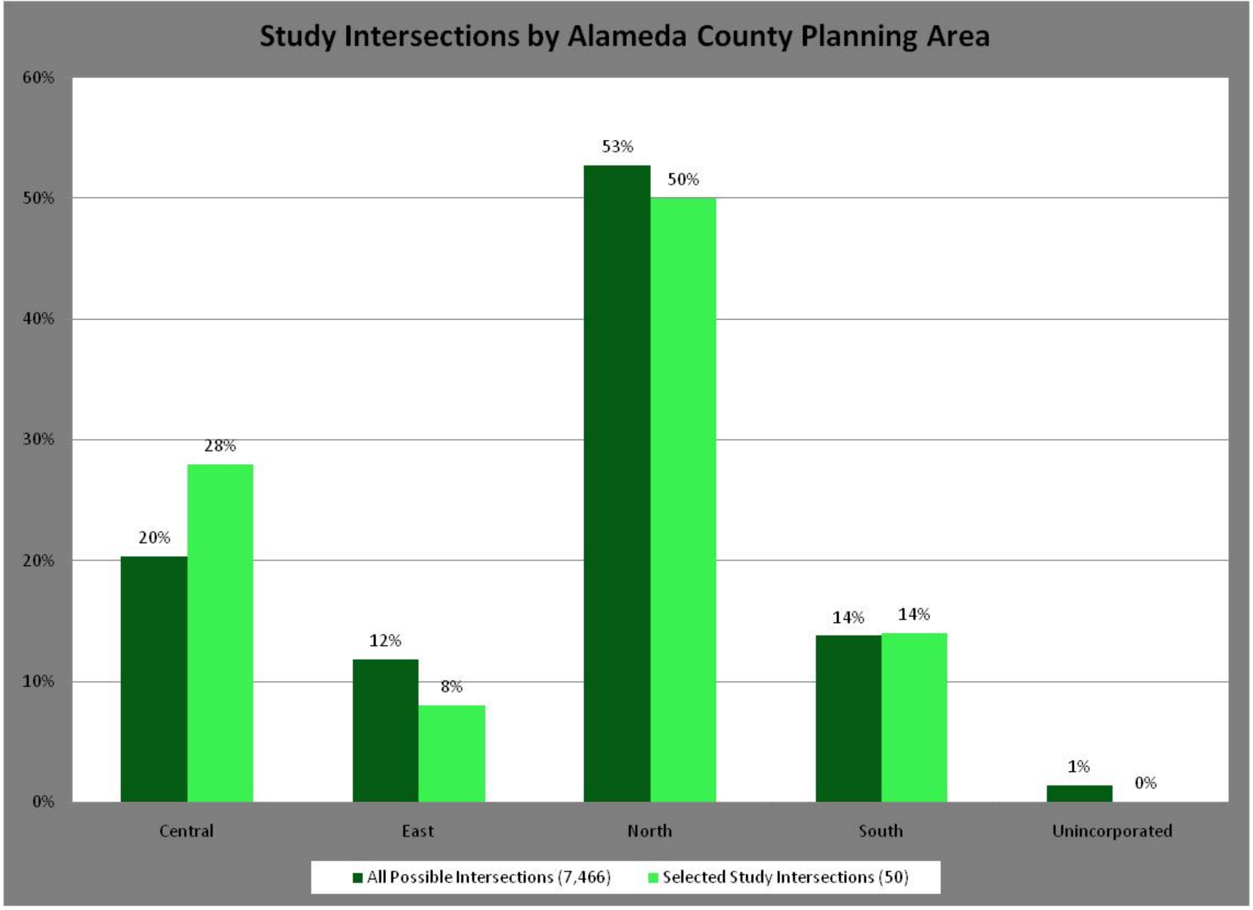
All four Alameda County Planning Areas are represented in the sample (see Figure 1).

Table 2. Characteristics of CalTrans and Non-CalTrans Intersections

	CalTrans Intersections		Non-CalTrans Intersections	
	Total	Selected	Total	Selected
Number*	528 (100%)	30 (100%)	6,938 (100%)	20 (100%)
Planning Area*				
North	232 (43.9%)	13 (43.3%)	3,703 (53.4%)	12 (60.0%)
Central	168 (34.1%)	12 (40.0%)	1,350 (19.5%)	2 (10.0%)
South	105 (19.9%)	5 (16.7%)	922 (13.3%)	2 (10.0%)
East	18 (3.4%)	0 (0.0%)	865 (12.5%)	4 (20.0%)
Unincorporated	0 (0.0%)	0 (0.0%)	98 (1.4%)	0 (0.0%)
Average population density within 1/4-mile	9,250	9,070	8,930	10,300
Average median income within 1/4-mile	$52,100	$47,800	$58,250	$51,800
Average retail land use intensity within 1/4-mile	30	34	19	28

*Table shows column percentages

Figure 1. Study Intersections by Alameda County Planning Area



Selection of Intersections for Automated Counters

The automated counters will be installed at a subset of the 50 intersection locations. The choice of locations for automated counters was based more on site conditions than random selection. Each automated pedestrian counter site needed to have safe and effective places to mount the infrared sensors. Each automated bicycle counter location needed a place to cut the pavement to install the inductive loop sensors and counter box.

Several selection requirements were used to increase variation between the automated count locations. The 50 intersections were classified into five general land use categories based on the characteristics of their surrounding areas: 1) Mixed Residential/Commercial (Small Scale), 2) Commercial Retail Strip, 3) Residential, 4) Mixed Residential/Commercial (Large Scale), and 5) Central Business District (CBD). It was assumed that each of these five categories would have different pedestrian daily and weekly peaking patterns.

Both bicycle sensors were installed on the far side of intersections within a bicycle lane. This site location was used because it was assumed to have the least variation in lateral riding location for each bicyclist (intersection approaches would have more mixing of bicycles and turning vehicles, increasing the variability in lateral bicycling position and reducing the number of bicyclists that would ride over the sensor). One of the bicycle sensors was installed in a roadway approaching a trail and the other was on an arterial leading to a major employment center.

Modeling Process

After the count data are collected, they will be compiled in a database for analysis. Initial analysis will include descriptive statistics and correlations. Further analysis will involve testing a variety of models for predicting pedestrian and bicycle volumes. For these models, the dependent variable will be a pedestrian or bicycle volume during a specific time period (e.g., number of pedestrians crossing an intersection during the peak hour on a Saturday or number of bicyclists crossing an intersection during a 24-hour weekday). In some cases, the volumes used as the dependent variable will be projected from the two-hour sample count to a full day, month, or year estimate using the adjustment factors developed from the continuous infrared sensor counts. One of the key model outputs will be estimated annual pedestrian and bicycle volumes. These annual estimates will be compared to annual pedestrian and bicycle crash totals to account for exposure.

At a minimum, two different sets of data will be used to generate models. One model will be based on the sample of 30 CalTrans roadway intersections. A second model will be based on all 50 intersection count locations. The first model will apply to the 528 CalTrans intersections and the second model will apply to all intersections in Alameda County.

Independent variables will be land use characteristics, roadway and motor vehicle traffic characteristics, pedestrian and bicycle infrastructure/facility characteristics, socioeconomic characteristics, and weather/topographic characteristics. The independent variables that will be considered for the analysis will come from the data sources listed in Appendix D: Data Sources Needed for Analysis.

Techniques for estimating the models may include simple linear regression or more advanced logistic regression techniques.

The models that are produced should be evaluated for goodness of fit between observed and predicted values. Effectiveness can be tested in the following ways:

- Comparing the predicted volumes from the model to the sample counts at the set of 30 or 50 locations (R-squared value)
- Calculating the GEH statistic to compare forecasted values against observed values at the same locations[20]
- Taking additional pedestrian and bicycle counts at a different sample of intersections (selected from the remaining 498 CalTrans intersections and/or remaining 6,918 non-CalTrans intersections) and calculating the percentage difference between the counts and the predicted model volumes

An additional possibility for this research project could be to compare the predictive accuracy of the models developed through this counting effort with the predictive accuracy of other pedestrian and bicycling modeling methods, such as Space Syntax.

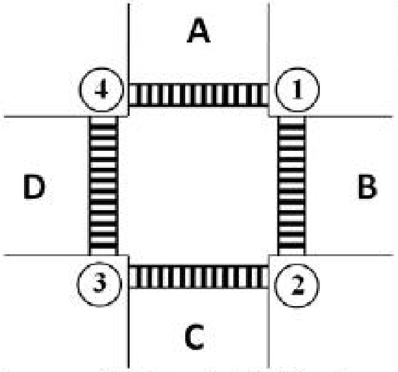
---

[20] Raford, Noah. *Draft Strategy for Countywide Pedestrian Volume Modeling: CalTrans State Highway System for San Diego County*, Drafted for the University of California Traffic Safety Center, August 25, 2007.

Appendix B: Data Collection Sheet

Intersection Pedestrian and Bicycle Count Sheet

Mainline Roadway: ____
Intersecting Roadway: ____
Observer Name(s): ____
Date: ____
Observation Time: (Start) ________________(End)________________
Temp. (°F): __________ Sunny, cloudy, rainy, etc.: __________________
Description of Specific Observation Location: ____________________
____
____



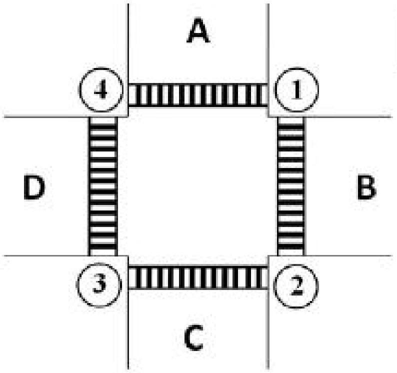
Write the compass direction next to each intersection leg (A, B, C, D) and each corner (1, 2, 3, 4) on the diagram. Also label road names.

Tally each time a pedestrian crosses in either direction on each leg of the intersection (count all crossings within 50 ft. of the crosswalk). If the pedestrian is female, mark an "O"; if male, mark an "X".

Time Period #	Pedestrian Counts							
	Crossing Leg A		Crossing Leg B		Crossing Leg C		Crossing Leg D	
	From 1 to 2	From 2 to 1	From 2 to 3	From 3 to 2	From 3 to 4	From 4 to 3	From 4 to 1	From 1 to 4
(0-15 min)								
(15-30 min)								
(30-45 min)								
(45-60 min)								
(60-75 min)								
(75-90 min)								
(90-105 min)								
(105-120 min)								
TOTAL	Female: Male:	Female: Male:	Female: Male:	Female: Male:	Female: Male:	Female: Male:	Female: Male:	Female: Male:

Intersection Pedestrian and Bicycle Count Sheet

Mainline Roadway: ______________________________
Intersecting Roadway: ______________________________
Observer Name(s): ______________________________
Date: ______________________________
Observation Time: (Start) ______________(End)______________
Temp. (°F): __________ Sunny, cloudy, rainy, etc.: ______________
Description of Specific Observation Location: ______________________
____
____



Write the compass direction next to each intersection leg (A, B, C, D) and each corner (1, 2, 3, 4) on the diagram. Also label road names.

Tally each time a bicyclist leaves each leg of the intersection. If the bicyclist is female, mark an "O"; if the bicyclist is male, mark an "X".

Time Period #	Bicycle Counts											
	Leaving Leg A			Leaving Leg B			Leaving Leg C			Leaving Leg D		
	(Turning Left) A to B	(Going Straight) A to C	(Turning Right) A to D	(Turning Left) B to C	(Going Straight) B to D	Turning Right) B to A	(Turning Left) C to D	(Going Straight) C to A	Turning Right) C to B	(Turning Left) D to A	(Going Straight) D to B	Turning Right) D to C
(0-15 min)												
(15-30 min)												
(30-45 min)												
(45-60 min)												
(60-75 min)												
(75-90 min)												
(90-105 min)												
(105-120 min)												
TOTAL	Female: Male:	Female: Male:	Female: Male:	Female: Male:	Female: Male:	Female: Male:	Female: Male:	Female: Male:	Female: Male:	Female: Male:	Female: Male:	Female: Male:

Appendix C: Intersections Selected for Analysis

Alameda County Pedestrian and Bicycle Count Study

Draft List of Intersections for Counts

UC-Berkeley Traffic Safety Center--February 2008

CalTrans Intersections (30)

Intersection	ID #	City	ACTIA Planning Area	Population per Sq. Mi. within 1/4-mile	Population Density Category	Average 1999 Median Income within 1/4-mile	Median Income Category	Commercial Score	Commercial Score Category	Transit Access Score	Transit Access Category	Trail/Road Intersection	Central Business District	Mainline ADT	Cross-Street ADT	Possible Counter Locations	Countywide Bicycle Plan	Countywide Pedestrian Plan
Ashby Avenue (CA 13) & Benvenue Avenue	19	Berkeley	North	11770	High	64159	High	50	High	64	High	No	Non-CBD	17565	501	PED 1.2		
Ashby Avenue (CA 13) & Telegraph Avenue	24	Berkeley	North	12711	High	53453	High	27	Medium	50	High	No	Non-CBD	20400	26700	BIKE 2; PED 2.3		
Ashby Avenue (CA 13) & Acton Street	42	Berkeley	North	14828	High	36083	Low	24	Medium	28	Medium	No	Non-CBD	23855	800	PED 3.1		
Doolittle Drive (CA 61) & Airport Access Road	50	Oakland	North	1295	Low	16029	Low	2	Low	21	Low	Yes	Non-CBD	22301	14210			SF Bay Trail
Broadway (CA 61) & Calhoun Street	66	Alameda	North	14005	High	56037	High	0	Low	28	Medium	No	Non-CBD	10390	401	PED 3.2	Class II Bicycle Lane	AC Transit
Encinal Avenue (CA 61) & Oak Street	74	Alameda	North	12564	High	48531	Medium	93	High	89	High	No	Non-CBD	9258	2601			
Encinal Avenue (CA 61) & Benton Street	82	Alameda	North	13492	High	53845	High	11	Low	33	Medium	No	Non-CBD	7475	201			
Ardenwood Boulevard (CA 84) & Newark Boulevard (E side interchange ramp)	102	Newark	South	3092	Low	28670	Low	31	Medium	20	Low	No	Non-CBD	XXXX	XXXX		Class II Bicycle Lane	Dumb. Express/SF Bay Trail
Thornton Avenue (CA 84) & Oak Street	141	Fremont	South	6477	Low	68050	High	14	Medium	20	Low	No	Non-CBD	29580	1501			Major Commercial
Fremont Boulevard (CA 84) & Peralta Boulevard	148	Fremont	South	7301	Low	58525	High	75	High	22	Medium	No	Non-CBD	19850	33001		Class II Bicycle Lane	AC Transit/Amtrak
Mowry Avenue (CA 84) & Cherry Lane	167	Fremont	South	8172	Medium	73529	High	10	Low	15	Low	No	Non-CBD	28298	501		Class II Bicycle Lane	
Davis Street (CA 61) & Warden Avenue	195	San Leandro	Central	3146	Low	58455	High	9	Low	4	Low	No	Non-CBD	46079	3001	PED 4.1	Class III Bicycle Route	
Davis Street (CA 61) & Pierce Avenue	197	San Leandro	Central	12066	High	50928	Medium	5	Low	10	Low	No	Non-CBD	37494	2401		Class II Bicycle Lane	
San Pablo Avenue (CA 123) & Ward Street	244	Berkeley	North	6663	Low	35483	Low	54	High	36	Medium	No	Non-CBD	26931	501	PED 2.2		AC Transit
San Pablo Avenue (CA 123) & Harrison Street	267	Berkeley	North	7680	Low	44186	Medium	46	High	92	High	No	Non-CBD	27356	501			AC Transit
Mission Boulevard (CA 185) & Grove Way	292	Cherryland	Central	9987	Medium	37706	Low	27	Medium	30	Medium	No	Non-CBD	29500	6800		Class II Bicycle Lane	AC Transit
East 14th Street (CA 185) & Hasperian Boulevard	326	San Leandro	Central	7548	Low	47281	Medium	52	High	22	Medium	No	Non-CBD	25450	20010		Class III Bicycle Route	AC Transit
East 14th Street (CA 185) & Maud Avenue	347	San Leandro	Central	10790	Medium	41419	Medium	69	High	41	Medium	No	Non-CBD	19969	1501			AC Transit
East 14th Street (CA 185) & Belleview Drive	363	San Leandro	Central	9484	Medium	54417	High	38	High	40	Medium	No	Non-CBD	22590	401			AC Transit
International Boulevard (CA 185) & 107th Avenue	370	Oakland	North	13146	High	41390	Medium	24	Medium	51	High	No	Non-CBD	23813	2001			AC Transit
International Boulevard (CA 185) & 99th Avenue	379	Oakland	North	18241	High	33655	Low	47	High	66	High	No	Non-CBD	25302	501			AC Transit
International Boulevard (CA 185) & 46th Avenue	434	Oakland	North	11644	High	28313	Low	90	High	56	High	No	Non-CBD	28355	9300			AC Transit
High Street (CA 185) & E 12th Street	438	Oakland	North	7763	Medium	29617	Low	48	High	44	Medium	No	Non-CBD	3470	12900		Class III Bicycle Route	
Mission Boulevard (CA 238) & Nichols Avenue	458	Fremont	South	2362	Low	82795	High	23	Medium	18	Low	No	Non-CBD	29250	2501			
Mission Boulevard (CA 238) & Overhill Drive	487	Hayward	Central	8036	Medium	59612	High	14	Medium	9	Low	No	Non-CBD	34000	251			AC Transit
Mission Boulevard (CA 238) & Valle Vista Avenue	489	Hayward	Central	8031	Medium	47969	Medium	10	Low	9	Low	No	Non-CBD	34000	1001			AC Transit
Mission Boulevard (CA 238) & Jefferson Street	499	Hayward	Central	6579	Low	50906	Medium	19	Medium	9	Low	No	Non-CBD	41675	601	PED 4.3		AC Transit
Mission Boulevard (CA 238) & Torrano Avenue	506	Hayward	Central	7531	Low	48338	Medium	11	Low	29	Medium	No	Non-CBD	39842	1101			AC Transit
Foothill Boulevard (CA 238) & D Street	516	Hayward	Central	7006	Low	38072	Low	88	High	150	High	No	Hayward	55000	4000	PED 4.2		AC Transit
Foothill Boulevard (CA 238) & Cotter Way	526	Hayward	Central	8564	Medium	45525	Medium	23	Medium	8	Low	No	Non-CBD	50507	151			AC Transit

Other Alameda County Intersections (20)

Intersection	ID #	City	ACTIA Planning Area	Population per Sq. Mi. within 1/4-mile	Population Density Category	Average 1999 Median Income within 1/4-mile	Median Income Category	Commercial Score	Commercial Score Category	Transit Access Score	Transit Access Category	Trail/Road Intersection	Central Business District	Mainline ADT	Cross-Street ADT	Possible Counter Locations	Countywide Bicycle Plan	Countywide Pedestrian Plan
Santa Clara Street & Ocie Way	910	Hayward	Central	9835	Medium	55358	Medium	0	Low	22	Medium	No	Non-CBD	(Unknown)	(Unknown)			
Alvarado Niles Road & Western Avenue	1141	Union City	South	6109	Medium	73232	High	3	Low	9	Low	Yes	Non-CBD	(Unknown)	(Unknown)			AC Transit
W Harder Road & Tarman Avenue	2240	Hayward	Central	13354	High	48633	Medium	9	Medium	4	Low	No	Non-CBD	(Unknown)	(Unknown)			
Bancroft Avenue & Auseon Avenue	2460	Oakland	North	16240	High	29675	Low	12	Medium	42	High	No	Non-CBD	(Unknown)	(Unknown)		Class II Bicycle Lane	AC Transit
Daugherty Road & Scarlett Drive (Iron Horse Trail)	7219	Dublin	East	2333	Low	69304	Medium	1	Low	2	Low	Yes	Non-CBD	(Unknown)	(Unknown)		Class I Bicycle Trail	
University Avenue & Bonar Street	2875	Berkeley	North	13789	High	31532	Low	37	High	48	High	No	Non-CBD	(Unknown)	(Unknown)	PED 2.1		AC Transit
College Avenue & Derby Street	2973	Berkeley	North	17537	High	46120	Medium	12	Medium	54	High	No	Non-CBD	(Unknown)	(Unknown)			AC Transit
Mandana Boulevard & Carlston Avenue	3734	Oakland	North	8336	Medium	113800	High	0	Low	16	Medium	No	Non-CBD	(Unknown)	(Unknown)			
Webster Street & 21st Street	3786	Oakland	North	8159	Medium	16279	Low	52	High	202	High	No	Oakland	(Unknown)	(Unknown)	PED 5.3		Downtown
Martin Luther King Jr. Way & 17th Street	3836	Oakland	North	11804	High	18042	Low	41	High	83	High	No	Oakland	(Unknown)	(Unknown)	PED 1.1		Downtown
Paseo Padre Parkway & Mowry Avenue	9179	Fremont	South	8760	Medium	26243	Low	10	Medium	48	High	No	Fremont	(Unknown)	(Unknown)			Downtown
Foothill Boulevard & 15th Avenue	4342	Oakland	North	18720	High	28589	Low	58	High	73	High	No	Non-CBD	(Unknown)	(Unknown)	PED 1.3		AC Transit
Solano Avenue & Masonic Avenue (Ohlone Trail)	2650	Albany	North	11732	High	61405	Medium	49	High	31	Medium	Yes	Non-CBD	(Unknown)	(Unknown)		Class I Bicycle Trail	Interj. Trail; Major Comm.
Moraga Avenue & Masonic Avenue	6703	Oakland	North	3608	Low	101200	High	0	Low	36	Medium	No	Non-CBD	(Unknown)	(Unknown)			
Owens Drive & Andrews Drive	7195	Pleasanton	East	2631	Low	71089	High	0	Low	10	Low	No	Non-CBD	(Unknown)	(Unknown)			AC Transit
Amador Valley Boulevard & Stagecoach Road	8422	Dublin	East	6396	Medium	71071	High	0	Low	6	Low	No	Non-CBD	(Unknown)	(Unknown)	BIKE 1, PED 3.3		
Stoneridge Drive & Hacienda Drive	8717	Pleasanton	East	2631	Low	71089	High	1	Low	25	Medium	No	Non-CBD	(Unknown)	(Unknown)		Class II Bicycle Lane	
Broadway & 12th Street	9436	Oakland	North	16757	High	14571	Low	118	High	345	High	No	Oakland	(Unknown)	(Unknown)	PED 5.1		AC Transit; Downtown
Webster Street & 7th Street	9471	Oakland	North	13612	High	39732	Low	149	High	114	High	No	Oakland	(Unknown)	(Unknown)	PED 5.2		Downtown
Chatham Road & 13th Avenue	9881	Oakland	North	13489	High	49481	Medium	12	Medium	107	High	No	Non-CBD	(Unknown)	(Unknown)			

Appendix D: Data Sources

Below is a list of data that have been used (or may be desirable) for selecting the sample of intersections and modeling the effects of different factors on pedestrian and bicycle counts. While it would be ideal to have all data on the list, budget limitations did not allow for primary data collection. Therefore, the list below indicates project data that have been gathered from existing national, state, regional, and local sources. Items that have not yet been gathered are in *italics*.

Land Use	**Data Source (Status)**
Population density (neighborhood level)	Census block groups (obtained from MTC and US Census)
Employment density (neighborhood level)	Traffic analysis zones (obtained from MTC)
Land use type/code (parcel level) • Commercial Retail • Commercial Office • Industrial • Residential (number of housing units) • Specific land use codes	Property parcels (obtained from Alameda County Assessor's Office)
Park locations	*Parks (obtained regional parks from MTC; still need data on local parks)*
School locations • *Grade levels* • *Number of students*	*Schools (obtained from MTC, but still need to create unique GIS layer)*
Building footprints • *Height* • *Number of stories* • *Setback from street*	*Building footprints (still need data from local municipalities)*
Other land use and landscape features	*Aerial photography (still need data from MTC)*
Transportation Infrastructure	**Data Source**
Sidewalk locations • *Width of sidewalk* • *Surface type/quality* • *Width of buffer between sidewalk and street*	*Sidewalks (still need data from local municipalities)*
Multi-use trail locations • Width • Surface type/quality	Multi-use trails (obtained from MTC)
Street tree locations	*Street trees (still need data from aerial photographs)*
Bicycle lane locations	*Bicycle lanes (obtained from MTC, but may require updating in City of Oakland)*
Bicycle route locations (signed routes)	Bicycle routes (obtained from MTC)
Roadway crossing characteristics • *Traffic volume (AADT)* • *Functional classification* • *Speed limit* • *Number of motorized travel and turning lanes* • *Roadway width* • *Locations with traffic signals (including*	*Alameda County roadway centerlines; CalTrans roadway centerlines; all roadway intersection points (including all points where roads intersect trails) (obtained centerlines from MTC and CalTrans intersections from CalTrans, but still need traffic volume, speed limit, number of lanes, and traffic signal data for non-CalTrans arterial and collector roadways)*

standard and mid-block pedestrian/bicycle crossing signals)	
Roadway intersection density	All roadway intersection points (including all points where roads intersect trails) (Created intersection layer from existing roadway centerline data)
Driveway intersection/crossing density	*Driveway points (not available)*
Adjacent roadway characteristics • *CalTrans vs. local ownership* • *Functional classification* • *Traffic volume (AADT)* • *Speed limit* • *Number of motorized travel lanes* • *Roadway width* • *Presence of on-street parking*	*Alameda County roadway centerlines; CalTrans roadway centerlines (obtained centerlines from MTC and CalTrans intersections from CalTrans, but still need traffic volume, speed limit, number of lanes, and traffic signal data for non-CalTrans arterial and collector roadways)*
Freeway locations	Freeways (obtained from MTC)
Bus route locations • *Headways*	*Bus routes (will use bus stop location data instead)*
Bus stop locations (by route) • Service frequency	Bus stops by route (obtained from MTC)
Transit station locations • Transit access volumes and modes	Transit stations (obtained from MTC)
Public and private parking lots • *Number of parking spaces*	*Parking lots (may be available from parcel data or aerial photo data)*
Railroad locations	Rail lines (obtained from MTC)
Socioeconomic Characteristics	**Data Source**
Percent of households with 0 motor vehicles (neighborhood)	Census block groups (obtained from MTC and US Census)
Median household income (neighborhood)	Census block groups (obtained from MTC and US Census)
Population under age 16 (neighborhood)	Census block groups (obtained from MTC and US Census)
Population over age 64 (neighborhood)	Census block groups (obtained from MTC and US Census)
Minority population (neighborhood)	Census block groups (obtained from MTC and US Census)
Percent of workers commuting by transit (neighborhood)	Census block groups (obtained from MTC and US Census)

Alameda County Pedestrian and Bicycle Counting Project Summary

UC-Berkeley Traffic Safety Center
Robert Schneider, Lindsay Arnold, and David Ragland
September 2008

With an increasing number of California communities adopting pedestrian and bicycle plans, conducting walking and bicycling safety audits, and encouraging cost-effective, low-emission transportation options, there is a greater need for pedestrian and bicycle data. In response to this need, the UC-Berkeley Traffic Safety Center gathered counts at 50 intersection locations throughout Alameda County to quantify pedestrian and bicycle activity and gain a more accurate understanding of pedestrian and bicycle crash risk. This pilot study demonstrates data collection and modeling methods that could be applied to CalTrans roadways statewide.

The project followed a rigorous scientific process, including selecting representative locations for sample counts; collecting data through manual and automated methods; developing adjustment factors for time of day, day of week, location, and weather; gathering fine-grained land use and transportation infrastructure data in GIS; and estimating a basic statistical model. Both pedestrians and bicyclists were counted, but the detailed analyses focus on pedestrians.

Two products of the study are summarized on the following pages. The first, “Extrapolating Weekly Pedestrian Intersection Crossing Volumes from 2-Hour Manual Counts,” describes a methodology that can be used to collect pedestrian counts and derive adjustment factors for estimating weekly pedestrian volumes from short manual counts. The second part of the study, “A Pilot Model for Estimating Pedestrian Intersection Crossing Volumes,” presents a simple mathematical formula for estimating the total number of pedestrians crossing an intersection during a typical week.

This research project is being conducted by the UC-Berkeley Traffic Safety Center for the California Department of Transportation and Alameda County Transportation Improvement Authority.

Extrapolating Weekly Pedestrian Intersection Crossing Volumes from 2-Hour Manual Counts

Purpose

Estimate weekly pedestrian volumes from 2-hour manual counts by accounting for time of day, location, and weather.

Abstract

Manual counts were conducted at a set of 50 intersections in Alameda County, CA. Automated counts from sidewalk locations near a subset of 11 intersections were used to adjust these counts for time, surrounding land use characteristics, and weather.

Key Points

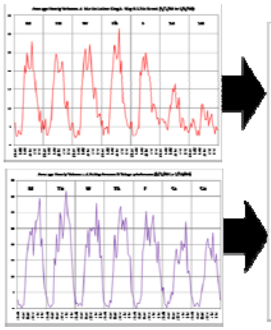
- Short pedestrian counts can be used to estimate weekly pedestrian volumes
- Manual and automated count methods can be used together to collect pedestrian data
- When developing weekly pedestrian volume estimates, it is critical to account for
  - Time of day and day of week
  - Surrounding land use characteristics
  - Weather conditions
- Estimated weekly volumes can be used to analyze pedestrian risk and develop pedestrian volume models
- Over 690,000 pedestrians were counted during the 13-week data collection period

Manual Counts

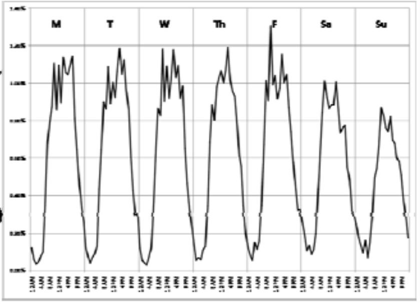
- Two 2-hour count periods at each location
  - Tue., Wed., or Thu. from 12-2 p.m. or 3-5 p.m.
  - Sat. from 9-11 a.m., 12-2 p.m., or 3-5 p.m.
- Counted people crossing each crosswalk
- Captured 15-minute intervals & gender

Automated Counts

- Used EcoCounter Pyroelectric Dual Infrared Sensors
- 5 counters were rotated between sidewalk locations within 100 feet of 11 of the 50 manual count intersections



Percent of Weekly Volume by Hour (Composite of 11 Sites)



Manual Pedestrian Count Results

Two-hour pedestrian counts were taken on weekdays and weekends at 50 intersections. Below are the counts and estimated total weekly pedestrian intersection crossing volumes at a sample of the intersections.

		Manual Count Data				*Calculated Results*
Mainline Roadway	Intersecting Roadway	Weekday Count[1]	Saturday Count[2]	*% Female*	*% Crossing Mainline Roadway*[3]	Estimated Total Weekly Pedestrian Crossings[4]
Broadway	12th Street	3577	1374	*45.3%*	*27.5%*	112,896
Solano Avenue	Masonic Avenue	514	397	*52.4%*	*48.2%*	22,203
International Boulevard	46th Avenue	287	286	*40.3%*	*15.9%*	12,303
University Avenue	Bonar Street	229	225	*39.2%*	*27.3%*	11,175
Mission Boulevard	Jefferson Street	171	77	*45.5%*	*46.0%*	5,736
San Pablo Avenue	Harrison Street	99	114	*35.7%*	*39.4%*	4,930
International Boulevard	107th Avenue	89	69	*34.2%*	*27.2%*	3,985
East 14th Street	Hasperian Boulevard	78	69	*45.6%*	*24.5%*	3,777
Ardenwood Boulevard	Newark Boulevard	55	15	*27.1%*	*77.1%*	1,635
Davis Street	Pierce Avenue	28	33	*45.9%*	*54.1%*	1,570
Mission Boulevard	Torrano Avenue	16	28	*38.6%*	*11.4%*	1,169
Foothill Boulevard	D Street	20	4	*33.3%*	*54.2%*	632
Total (for all 50 intersections)		12786	7248			462,841
Mean (for all 50 intersections)		256	145			9,257
Standard Deviation (for all 50 intersections)		579	261			17,960

1) Counts taken during one of the following time periods: Tuesday 12-2 p.m., Tuesday 3-5 p.m., Wednesday 12-2 p.m., Wednesday 3-5 p.m., Thursday 12-2 p.m., and Thursday 3-5 p.m. (April through June 2008)

2) Counts taken during one of the following time periods: Saturday 9-11 a.m., Saturday 12-2 p.m., and Saturday 3-5 p.m. (April through June 2008)

3) Mainline roadway is intersecting roadway with highest motor vehicle volume.

4) Total Estimated Weekly Volume is adjusted for time of day, day of week, land use type, and weather.

Pedestrian Crash Risk Analysis

Pedestrian volume estimates were used for a preliminary analysis of pedestrian risk at the 50 intersections. Below are the calculations for a sample of the intersections using reported crashes between 1996 and 2005.

Mainline Roadway	Intersecting Roadway	Estimated Total Weekly Pedestrian Crossings[1]	Annual Pedestrian Volume Estimate[2]	Ten-Year Pedestrian Volume Estimate[2]	Reported Pedestrian Crashes (1996-2005)[3]	Pedestrian Risk (Crashes per 10,000,000 crossings)
Mission Boulevard	Torrano Avenue	1,169	60,796	607,964	5	82.24
Davis Street	Pierce Avenue	1,570	81,619	816,187	4	49.01
Foothill Boulevard	D Street	632	32,862	328,624	1	30.43
Mission Boulevard	Jefferson Street	5,236	272,246	2,722,464	5	18.37
University Avenue	Bonar Street	11,175	581,113	5,811,127	7	12.05
Ardenwood Boulevard	Newark Boulevard	1,635	85,038	850,382	1	11.76
International Boulevard	107th Avenue	3,985	207,243	2,072,429	2	9.65
San Pablo Avenue	Harrison Street	4,930	256,357	2,563,572	2	7.80
East 14th Street	Hasperian Boulevard	3,777	196,410	1,964,102	1	5.09
International Boulevard	46th Avenue	12,303	639,752	6,397,522	3	4.69
Solano Avenue	Masonic Avenue	22,203	1,154,559	11,545,589	2	1.73
Broadway	12th Street	112,896	5,870,590	58,705,898	5	0.85
Total (for all 50 intersections)		462,841	24,067,724	240,677,241	89	443.59
Mean (for all 50 intersections)		9,257	481,354	4,813,545	1.78	13.86
Standard Deviation (for all 50 intersections)		17,960	933,976	9,339,260	1.95	16.78

1) Total Estimated Weekly Volume is adjusted for time of day, day of week, land use type, and weather.

2) Annual and ten year pedestrian volume estimates do not account for potential seasonal variations.

3) Police-reported intersection pedestrian crashes were compiled by the Alameda County Public Health Department.

A Pilot Model for Estimating Pedestrian Intersection Crossing Volumes

Purpose

Develop a simple, user-friendly model to estimate the total number of pedestrians crossing collector and arterial roadway intersections during a typical week.

Abstract

The pilot model of pedestrian intersection crossing volumes is based on pedestrian volume counts at a sample of 50 intersections in Alameda County, CA with a wide variety of surrounding land uses, transportation system attributes, and neighborhood socioeconomic characteristics. The final recommended model has a good overall fit (adjusted-R^2=0.897). Statistically-significant factors in the model include the population within a 0.5-mile radius, employment within a 0.25-mile radius, number of commercial retail properties within a 0.25-mile radius, and the presence of a regional transit station within a 0.1-mile radius of an intersection. The model has a simple structure, and it can be implemented by practitioners using geographic information systems and a basic spreadsheet program. Since the study is based on a relatively small number of intersections in one urban area, additional research is needed to refine the model and determine its applicability in other areas.

Key Points

- Pilot model is easy to estimate: data can be gathered in GIS, and estimates can be calculated with a simple spreadsheet
- Model has a good overall fit
  - Adjusted-R^2 = 0.897
  - F-test is significant at >99.9% confidence level
  - Predictive variables are sig. at >95% confidence level
- Estimated pedestrian volumes depend on:
  - Total population within 0.5-mile radius
  - Total employment within 0.25-mile radius
  - Number of commercial properties within 0.25-mile radius
  - Presence of transit station within 0.1-mile radius
- Model is based on a relatively small sample of intersection counts (N=50), so additional research is needed to refine the model and test its applicability in other areas

Strategic Sampling Process

- Sample = 50 intersections on arterial & collector roadways
- 30 intersections on state roadways (stratified random selection process). Stratification was based on:
  - Population density within 0.25 miles of the intersection (low, med., high)
  - Median income of households within 0.25 miles of the intersection (low, med., high)
  - Number of commercial properties within 0.10 miles of the intersection (low, med., high)
- 20 intersections on other roadways (constrained random selection process). Constraints included:
  - At least 4 intersections in central business districts
  - At least 2 intersections with major multi-use trail crossings
  - At least 3 intersections in each of the four county planning areas
- Other rules:
  - No intersection within 0.25 miles of any other intersection
  - Min. population density within 0.25 miles = 50 residents per square mile
  - No intersections under construction or grade-separated intersections

Key Factors Used in Model


Total employment within 0.25 miles


Number of commercial retail properties within 0.25 miles


Total population within 0.5 miles

Transit station within 0.1 miles

3 Model Alternatives

Dependent Variable = Total Weekly Pedestrian Intersection Crossings

	Model A		Model B		Model C	
Model Variables	Coeff.	(Std. Err.)^1	Coeff.	(Std. Err.)^1	Coeff.	(Std. Err.)^1
CONSTANT	4170	(4270)	-4910	(2050)**	-5790	(1990)***
TOTPOP_T					14.5	(7.19)**
TOTPOP_H	0.884	(0.254)***	0.928	(0.266)***		
TOTEMP_Q	1.72	(0.400)***	2.19	(0.367)***		
NCOMPROP_T					456	(118)***
NCOMPROP_Q	106	(39.0)***	98.4	(40.8)**		
NBARTSTA_T	56,800	(7810)***	54,600	(8160)***		
NBARTSTA_Q					44,800	(6280)***
NBUSSTOP_T					465	(81.4)***
PCTU18_Q	-36,400	(15,200)**				
Overall Model						
Sample Size (N)	50		50		50	
Adjusted R^2	0.907		0.897		0.870	
F-Test	96.6***		108***		83.2***	

1) Significance is indicated by asterisks: "***" indicates significant at 99% (p<0.01), "**" indicates significant at 95% (p<0.05), "*" indicates significant at 90% (p<0.10)

Pilot Model Formula

Total pedestrian intersection crossings per week =

0.928	x	Total population within 0.5-miles of the intersection
+2.19	x	Total employment within 0.25-miles of the intersection
+98.4	x	Number of commercial retail properties within 0.25-miles of the intersection
+54,600	x	Number of regional transit stations within 0.1-miles of the intersection
- 4,910		(constant)

Land Use and Transportation System Variables Considered for the Pedestrian Volume Model

LAND USE VARIABLES

		50 Study Intersections (N = 50)				All Major Street Intersections (N = 8055)				
Variable Name	**Description**	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**Data Source (Year)**
TOTPOP_T	Total population within 1/10-mile (161 m)	291	156	7.03	614	265	208	0.0658	1700	U.S. Census (2000)
TOTPOP_Q	Total population within 1/4-mile (402 m)	1880	869	254	3670	1640	1180	0.390	7430	U.S. Census (2000)
TOTPOP_H	Total population within 1/2-mile (805 m)	7500	3290	798	15100	6410	4140	1.96	21700	U.S. Census (2000)
TOTEMP_T	Total employment within 1/10-mile (161 m)	315	764	1.54	4170	151	350	0.896	4190	SF MTC[6](2005)
TOTEMP_Q	Total employment within 1/4-mile (402 m)	1660	3510	9.60	18900	930	1930	5.60	19600	SF MTC[6](2005)
PCTVAC_T	Proportion of housing units within 1/10-mile (161 m) that are vacant	0.0398	0.028	0.00673	0.122	0.0373	0.0349	0.00	0.371	U.S. Census (2000)
PCTVAC_Q	Proportion of housing units within 1/4-mile (402 m) that are vacant	0.0385	0.026	0.00849	0.106	0.0366	0.0325	0.00	0.290	U.S. Census (2000)
TOTVAC_T	Number of housing units within 1/10-mile (161 m) that are vacant	5.22	5.66	0.127	29.7	4.13	5.51	0.00	75.9	U.S. Census (2000)
TOTVAC_Q	Number of housing units within 1/4-mile (402 m) that are vacant	32.1	28.5	1.58	124	25.4	30.2	0.00	316	U.S. Census (2000)
PCTRENT_T	Proportion of housing units within 1/10-mile (161 m) that are rented	0.549	0.198	0.0560	0.923	0.449	0.254	0.00	1.00	U.S. Census (2000)
PCTRENT_Q	Proportion of housing units within 1/4-mile (402 m) that are rented	0.544	0.186	0.0555	0.912	0.45	0.244	0.00	1.00	U.S. Census (2000)
TOTRENT_T	Number of housing units within 1/10-mile (161 m) that are rented	69.4	52.1	3.39	230	60.0	76.8	0.00	970	U.S. Census (2000)
TOTRENT_Q	Number of housing units within 1/4-mile (402 m) that are rented	453	310	22.0	1240	369	427	0.00	2850	U.S. Census (2000)
NCOMPROP_T	Number of commercial properties within 1/10-mile (161 m)	6.66	8.11	0.00	40.0	3.48	6.04	0.00	48.0	Alameda Co. Assessor (2007)
NCOMPROP_Q	Number of commercial properties within 1/4-mile (402 m)	25.3	26.5	0.00	50.0	15.3	20.6	0.00	134	Alameda Co. Assessor (2007)
NESCH_T	Number of elementary schools within 1/10-mile (161 m)	0.0400	0.196	0.00	1.00	0.049	0.22	0.00	2.00	Alameda Co. Assessor (2007)
NESCH_Q	Number of elementary schools within 1/4-mile (402 m)	0.320	0.508	0.00	2.00	0.307	0.53	0.00	3.00	Alameda Co. Assessor (2007)
NMSCH_T	Number of middle schools within 1/10-mile (161 m)	0.00	0.00	0.00	0.00	0.00857	0.0922	0.00	1.00	Alameda Co. Assessor (2007)
NMSCH_Q	Number of middle schools within 1/4-mile (402 m)	0.08	0.337	0.00	2.00	0.0567	0.233	0.00	2.00	Alameda Co. Assessor (2007)
NHSCH_T	Number of high schools within 1/10-mile (161 m)	0.00	0.00	0.00	0.00	0.00372	0.0609	0.00	1.00	Alameda Co. Assessor (2007)
NHSCH_Q	Number of high schools within 1/4-mile (402 m)	0.0400	0.196	0.00	1.00	0.0539	0.232	0.00	2.00	Alameda Co. Assessor (2007)
NTSCH_T	Number of elem., middle, high, and other schools within 1/10-mile (161 m) [1]	0.060	0.237	0.00	1.00	0.0683	0.260	0.00	2.00	Alameda Co. Assessor (2007)
NTSCH_Q	Number of elem., middle, high, and other schools within 1/4-mile (402 m)[1]	0.480	0.700	0.00	3.00	0.458	0.669	0.00	4.00	Alameda Co. Assessor (2007)
COLDUM_T	Presence of college campus within 1/10-mile (161 m) (Yes=1, No=0)	0.00	0.00	0.00	0.00	0.0283	0.166	0.00	1.00	Alameda Co. Assessor (2007)
COLDUM_Q	Presence of college campus within 1/4-mile (402 m) (Yes=1, No=0)	0.00	0.00	0.00	0.00	0.0622	0.242	0.00	1.00	Alameda Co. Assessor (2007)

TRANSPORTATION SYSTEM VARIABLES

		50 Study Intersections (N = 50)				All Major Street Intersections (N = 8055)				
Variable Name	**Description**	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**Data Source (Year)**
NBARTSTA_T	Number of regional rail transit stations within 1/10-mile (161 m)	0.0200	0.140	0.00	1.00	0.00993	0.0992	0.00	1.00	SF MTC[6](2007)
NBARTSTA_Q	Number of regional rail transit stations within 1/4-mile (402 m)	0.0400	0.196	0.00	1.00	0.0467	0.212	0.00	2.00	SF MTC[6](2007)
NBUSSTOP_T	Number of bus route stops within 1/10-mile (161 m)[2]	12.9	17.1	0.00	118	8.64	11.4	0.00	135	SF MTC[6](2007)
NBUSSTOP_Q	Number of bus route stops within 1/4-mile (402 m)[2]	47.4	55.9	2.00	335	36.1	39.8	0.00	337	SF MTC[6](2007)
TRAILMI_T	Total multi-use trail centerline distance (miles) within 1/10-mile (161 m)	0.0258	0.0765	0.00	0.365	0.014	0.0558	0.00	0.609	SF MTC[6](2007)
TRAILMI_Q	Total multi-use trail centerline distance (miles) within 1/4-mile (402 m)	0.0916	0.262	0.00	1.32	0.0719	0.207	0.00	1.78	SF MTC[6](2007)
STREETMI_T	Total street centerline distance (miles) within 1/10-mile (161 m)	0.939	0.287	0.227	1.53	0.758	0.254	0.00	2.34	SF MTC[6](2007)
STREETMI_Q	Total street centerline distance (miles) within 1/4-mile (402 m)	5.64	1.43	2.23	9.40	4.70	1.57	0.278	10.5	SF MTC[6](2007)
BL_MI_T	Total centerline (miles) of streets with bicycle lanes within 1/10-mile (161 m)	0.101	0.157	0.00	0.471	0.086	0.153	0.00	0.936	SF MTC[6](2007)
BL_MI_Q	Total centerline (miles) of streets with bicycle lanes within 1/4-mile (402 m)	0.321	0.365	0.00	1.10	0.327	0.443	0.00	2.20	SF MTC[6](2007)
FWY_DUM_T	Freeway presence within 1/10-mile (161 m) (Yes = 1, No = 0)	0.120	0.325	0.00	1.00	0.168	0.374	0.00	1.00	SF MTC[6](2007)
FWY_DUM_Q	Freeway presence within 1/4-mile (402 m) (Yes = 1, No = 0)	0.180	0.384	0.00	1.00	0.302	0.459	0.00	1.00	SF MTC[6](2007)
SWCOV_Q	Est. sidewalk coverage (0.00,0.25,0.50,0.75,1.00) within 1/4-mile (402 m)[3]	0.875	0.195	0.25	1.00		Not calculated[5]			Google Earth® (2008)
SWBUF_Q	Est. prop. of sidewalks with buffer (0.00,0.25,0.50,0.75,1.00) within 1/4-mile[4]	0.525	0.288	0.00	1.00		Not calculated[5]			Google Earth® (2008)

1) Total schools does not include colleges. Colleges are included in a separate variable.

2) The number of "bus route stops" is the sum of the number of different bus routes servicing each bus stop within a given distance of the intersection (e.g., if 4 routes service a single bus stop, that particular bus stop will be counted 4 times).

3) Sidewalk coverage is estimated from aerial photography. 100% coverage (1.00) is sidewalks on both sides of all surface streets within 1/4-mile of the intersection. Sidewalks on only one side of all streets would be considered 50% coverage (0.50).

4) Sidewalk buffer is estimated from aerial photography. 100% buffer (1.00) indicates that the sidewalks on both sides of all surface streets are separated from the edge of the roadway by a grass, tree, shrub, or other type of buffer. If

5) Detailed intersection characteristics were not gathered for all roadways in Alameda County. Because of cost, these characteristics would only be collected if they were significant in the final regression model.

6) SF MTC = San Francisco Bay Area Metropolitan Transportation Commission.

Neighborhood Socioeconomic and Intersection Site Variables Considered for the Pedestrian Volume Model

NEIGHBORHOOD SOCIOECONOMIC VARIABLES

		50 Study Intersections (N = 50)				All Major Street Intersections (N = 8055)				
Variable Name	**Description**	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**Data Source (Year)**
PCTWHITE_T	Proportion of population within 1/10-mile (161 m) that is white	0.461	0.202	0.0401	0.822	0.495	0.234	0.0233	1.00	U.S. Census (2000)
PCTWHITE_Q	Proportion of population within 1/4-mile (402 m) that is white	0.463	0.196	0.0655	0.822	0.492	0.232	0.0364	0.920	U.S. Census (2000)
PCTMALE_T	Proportion of population within 1/10-mile (161 m) that is male	0.489	0.0286	0.421	0.556	0.492	0.0319	0.231	0.742	U.S. Census (2000)
PCTMALE_Q	Proportion of population within 1/4-mile (402 m) that is male	0.492	0.0222	0.455	0.560	0.491	0.0271	0.356	0.679	U.S. Census (2000)
PCT0VEH_T	Proportion of households within 1/10-mile (161 m) that have no automobile	0.168	0.168	0.0138	0.769	0.124	0.13	0.00	0.802	U.S. Census (2000)
PCT0VEH_Q	Proportion of households within 1/4-mile (402 m) that have no automobile	0.159	0.150	0.0150	0.638	0.126	0.127	0.00	0.738	U.S. Census (2000)
TOT0VEH_T	Total households within 1/10-mile (161 m) that have no automobile	23.2	36.2	0.299	182	18.3	31.3	0.00	483	U.S. Census (2000)
TOT0VEH_Q	Total households within 1/4-mile (402 m) that have no automobile	148	200	2.06	964	112	172	0.00	1600	U.S. Census (2000)
MEDINC_T	Median income (1999 dollars) of households within 1/10-mile (161 m)[1,2]	47800	21000	122	107500	59700	27900	122	167000	U.S. Census (2000)
MEDINC_Q	Median income (1999 dollars) of households within 1/4-mile (402 m)[1,2]	49400	20300	14600	114000	59600	27200	1051	169400	U.S. Census (2000)
PCTU18_T	Proportion of population within 1/10-mile (161 m) that is under 18 years old	0.223	0.0675	0.0563	0.372	0.234	0.0728	0.00872	0.626	U.S. Census (2000)
PCTU18_Q	Proportion of population within 1/4-mile (402 m) that is under 18 years old	0.223	0.0633	0.0742	0.364	0.236	0.0694	0.0112	0.625	U.S. Census (2000)
PCTO64_T	Proportion of population within 1/10-mile (161 m) that is over 64 years old	0.117	0.0776	0.0245	0.423	0.108	0.0573	0.00	0.502	U.S. Census (2000)
PCTO64_Q	Proportion of population within 1/4-mile (402 m) that is over 64 years old	0.114	0.0631	0.0245	0.340	0.108	0.0508	0.00	0.394	U.S. Census (2000)

INTERSECTION SITE VARIABLES

		50 Study Intersections (N = 50)				All Major Street Intersections (N = 8055)				
Variable Name	**Description**	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**Data Source (Year)**
CONTROLDUM	Either traffic signal or stop sign controling mainline roadway (Yes=1, No=0)[3]	0.560	0.496	0.00	1.00	Not calculated[7]				Field observation (2008)
MAXADT	Max. average daily traffic volume on a roadway passing through intersection	24000	12200	3000	55000	Not calculated[7]				CA DOT (2007); local municipalities[8]
MAIN_WIDTH	Average curb-to-curb length (feet) of the 2 crosswalks across the mainline ro	79.2	27.8	29.0	163	Not calculated[7]				Field obs., Google Maps® (2008)
MAIN_LANES	Average number of lanes on mainline approaches to the intersection[3,4]	4.44	1.47	2.00	8.50	Not calculated[7]				Field obs., Google Maps® (2008)
MAIN_XW	Number of marked crosswalks across the mainline roadway[3]	1.30	0.755	0.00	2.00	Not calculated[7]				Field obs., Google Maps® (2008)
MAIN_MED	Median refuge area present for at least one mainline roadway crosswalk[3]	0.580	0.494	0.00	1.00	Not calculated[7]				Field obs., Google Maps® (2008)
MAIN_BL	Bicycle lanes on at least one mainline approach to intersection[3]	0.260	0.439	0.00	1.00	Not calculated[7]				Field obs., Google Maps® (2008)
CURBRADCAT	Curb radius category (<15 feet (<4.57 m)=1, 15-25 feet=2, >25 feet (>7.62 m)	1.90	0.806	1.00	3.00	Not calculated[7]				Field obs., Google Maps® (2008)
TINTER	Intersection is a "T" intersection (Yes=1, No=0)[6]	0.240	0.427	0.00	1.00	Not calculated[7]				Field obs., Google Maps® (2008)

1) Median income is calculated as the weighted average of median incomes reported for the census block groups surrounding the intersection. Weights are assigned based on the proportion of the census block group within the specific buffer diatance from the intersection.

2) Several census block groups did not have data for median income. Intersections with a median income of 0 within the given buffer distance considered in this statistical summary.

3) Mainline roadway is the intersecting roadway with the higher traffic volume.

4) Average number of lanes on each mainline approach includes all through-, left-, and right-turn lanes.

5) Curb radius category reflects the average estimated curb radius of all corners at the intersection.

6) "T" intersections are 3-way intersections. Intersections were not considered to be "T" intersections if the fourth approach was a commercial driveway.

7) Detailed intersection characteristics were not gathered for all roadways in Alameda County. Because of cost, these characteristics would only be collected if they were significant in the final regression model.

8) Traffic volume data were gathered from Alameda (2004), Berkeley (2000-2007), Dublin (2000-2007), Fremont (2005), Hayward (2003-2008), Livermore (2007), Pleasanton (2007), and Oakland (2007).