Title

Estimating Pedestrian Accident Exposure: Approaches to a Statewide Pedestrian Exposure Database

Authors
Greene-Roesel, Ryan
Diogenes, Mara Chagas
Ragland, David R

Publication Date
2007-03-01

ESTIMATING PEDESTRIAN ACCIDENT EXPOSURE

Approaches to a Statewide Pedestrian Exposure Database

March 2007

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

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).

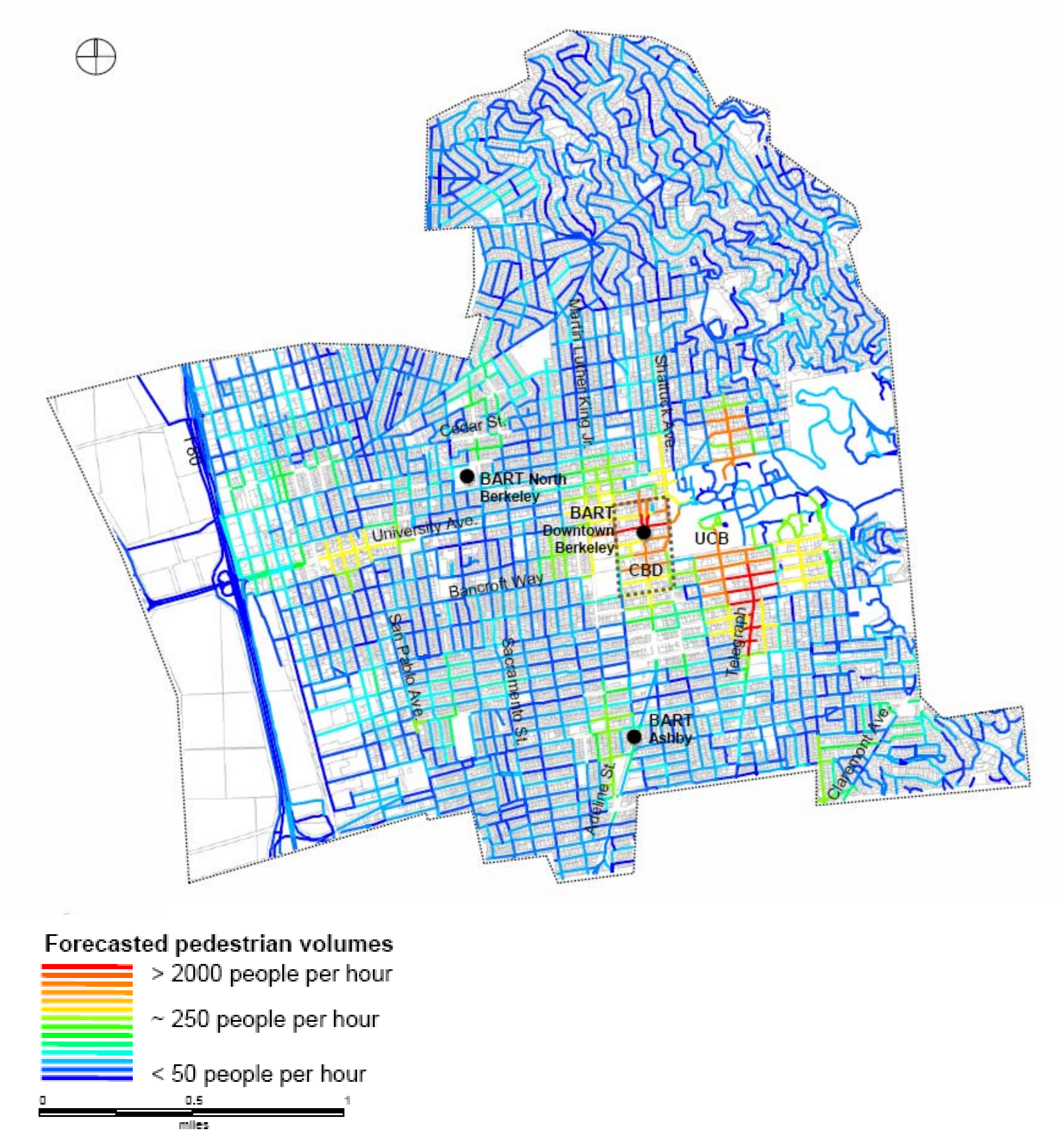

Figure 5.1: Forecasted pedestrian volumes in the city of Berkeley

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.