iCSC International Cycling Safety Conference

6th Annual International Cycling Safety Conference
21-22 September 2017, Davis, California, USA

Identification of California Bicycle Factor Groups

F.R. Proulx*, H. Louch#, S. Forsythe†

*Senior Planner/Researcher
Alta Planning + Design
233 A Street, Suite 703, San Diego, CA 92101
email: frankproulx@altaplanning.com

#Principal
Alta Planning + Design
100 Webster Street, Suite 300, Oakland, CA 94607
email: hughlouch@altaplanning.com

†Office of Multimodal System Planning
California Department of Transportation
Division of Transportation Planning, MS #32, P.O. Box 942874
Sacramento, CA 94274-0001
Email: scott.forsythe@dot.ca.gov

Keywords: exposure estimation, bicycle counts, clustering, California.

1 INTRODUCTION

A critical task in understanding the state of bicycling, including bicyclist safety, is developing an estimate of bicycle traffic volumes and bicycle miles traveled. California's recently drafted first statewide bicycle and pedestrian plan, *Toward an Active California*, placed a major emphasis on improving data collection practices and data systems for bicycle and pedestrian activity. We developed a baseline data methodology for the state to follow, which is organized around three broad categories of data: activity/demand, network/infrastructure, and safety. For each of these categories, an overview of existing data collection practices is provided, best practices are reviewed, and specific recommendations are made on how to improve the state of data collection.

One of the main foci of the activity/demand data section is a recommendation to develop a statewide non-motorized traffic monitoring program. Bicycle traffic counts conducted within this program will be critical for understanding bicyclists' exposure to crashes, which is a widely acknowledged difficulty in studying bicycle crash risk factors. Because these recommendations are made at the state level, the largest hurdle is in coordinating data storage and standardizing data collection methods. A statewide traffic count database, similar to the current motorized traffic count Performance Measurement System (PeMS), should be developed and maintained to house data collected by the state Department of Transportation (DOT) as well as local partners [1]. Short-duration and continuous counting programs also need to be established to populate this database.

This abstract presents a preliminary analysis of existing continuous bicycle counter data collected in California in 2015-2016. These counters are grouped using a clustering algorithm to establish factor groups. Findings presented here can help agencies within California and elsewhere to extrapolate their short-duration counts, providing a stronger empirical basis for the estimation of exposure and hence informing evaluations of bicyclist crash risk.

2 METHODOLOGY

The bicycle count data used here was collected using induction loops installed in a combination of on-street locations and off-road paths. All of the data was collected in the San Francisco and San Diego regions, which are two major population centers within the state.

Our primary goal in this analysis is to understand the temporal traffic patterns occurring at these count locations. That is, we seek to separate them into *factor groups*. Factor groups are collections of monitoring locations with similar peaking patterns [2]. The goal of developing factor groups is to provide a basis for standardizing short-duration counts, such as PM peak weekday counts which continue to be widely collected for bicycle and pedestrian traffic [3]. In order to understand how much of the total traffic at a site this short-duration count represents, and therefore what the overall exposure is, we require an understanding of the underlying temporal patterns.

Previous work has established groupings based on both a site's hour-of-day patterns and day-of-week patterns [4]. However, we suspect that this might be overly restrictive – a site with heavy AM/PM peaking patterns on weekdays may not necessarily have higher weekday traffic than weekend traffic.

To identify bicycle factor groups, counters were paired so that every observed site consists of bidirectional traffic, either by using both directions of counts along a given street/path segment, or by pairing counts from sets of one-way couplets. Within each site, normalized average hour-of-weekday patterns and day-of-week patterns were calculated. The hour-of-day pattern is composed of 24 hourly observations per site, and similarly the day-of-week pattern is composed of 7 daily observations per site, each of which sum to 1 within each location.

These features were then grouped using k-means clustering [5]. The number of groups was identified using the "elbow technique" heuristic, where the sum-of-squared deviations from cluster centers is plotted as a function of the number of clusters. An "elbow" often occurs where the addition of more clusters has a diminished effectiveness in summarizing clusters.

3 RESULTS

For the hour-of-weekday patterns, we identify 4 clusters as the optimum. As can be seen in Figure 1, we see a dominant AM/PM peak at locations in downtown San Francisco while locations on bike paths outside of the central cities tend to have less peaking activity on weekdays. These groups can be summarized as:

- **Hourly group 0:** These locations also have pronounced commute peaks, although in this group the A.M. peak dominates. There is little immediate explanation for this patterning, but it bears noting that these locations are also primarily in San Francisco on the outer edge of the downtown area.
- **Hourly group 1:** These locations have minimal discernible A.M./P.M. peaking activity, suggesting that these locations are not heavily used for commuting and may have more weekday recreational activity. These sites are in locations that would likely be characterized as recreational, primarily near the ocean in both San Diego and San Francisco.
- **Hourly group 2:** Sites in this group are characterized by both high A.M. and P.M. peaks on weekdays. This type of pattern would suggest heavy rates of commuting activity. These sites are all located within the downtown area of San Francisco.
- **Hourly group 3:** These sites have a pattern similar to group 1 with no discernible peaking.

For the day-of-week grouping, we identify 3 clusters, which are summarized as:

- **Daily Group 0:** Higher activity levels on the weekends than the weekdays. These appear to be primarily recreational sites, which is supported by their locations around the Golden Gate Bridge and the San Diego waterfront.
- **Daily Group 1:** Consistent volumes between weekdays and weekends. These sites, which could be categorized as having mixed daily patterns, are generally located in locations with moderate employment density as well as recreation potential, such as the western half of San Francisco, downtown San Diego, and some inland locations in San Diego county.
- **Daily Group 2:** Higher activity levels on weekdays than on weekends, suggesting primarily utilitarian travel. The majority of count sites in downtown San Francisco fall into this grouping.

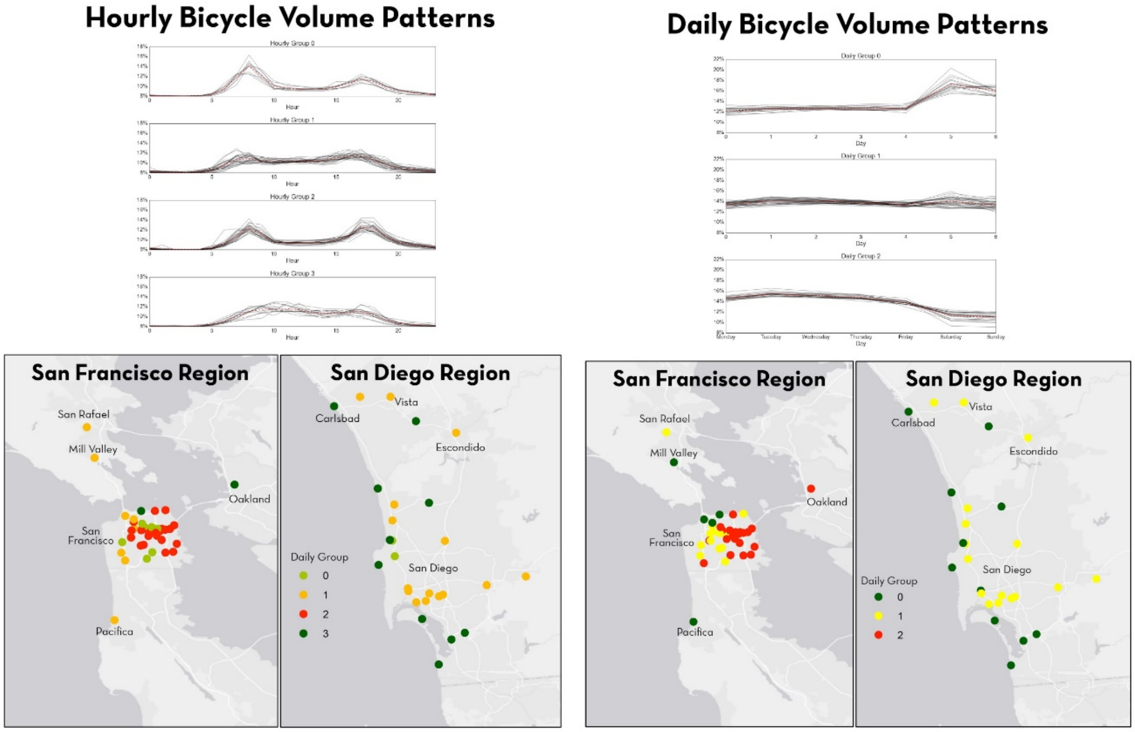

Figure 1. Observed bicycle traffic patterns in San Francisco and San Diego, CA.

The temporal patterns observed here provide a basis for extrapolating short-duration peak period counts conducted in these regions.

4 CONCLUSIONS

This abstract presents the results of a study of continuous bicycle counts conducted in California. These findings are intended to help inform the design of a statewide non-motorized traffic monitoring programs, which can in turn be used to estimate bicyclist exposure to crashes. California has a strong start in collecting continuous bicycle counts, thanks to the efforts of local and regional agencies, but these programs do not necessarily represent all of the traffic patterns that might exist across the state. The momentum generated from analyzing data from these existing programs can be used to push forward with expanding monitoring across the state.

5 REFERENCES

[1] California Department of Transportation, "Performance Measurement System". http://pems.dot.ca.gov/.

[2] Federal Highway Administration. *Traffic Monitoring Guide*, 2016.

[3] Alta Planning + Design. "National Bicycle and Pedestrian Documentation Project." http://bikepeddocumentation.org/.

[4] Miranda-Moreno, L., Nosal, T., Schneider, R., & Proulx, F. (2013). Classification of bicycle traffic patterns in five North American Cities. Transportation Research Record: Journal of the Transportation Research Board, (2339), 68-79.

[5] MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In Proceedings of the fifth Berkeley symposium on mathematical statistics and probability (Vol. 1, No. 14, pp. 281-297).