Towards More Robust Spatial Sampling Strategies for Non-motorized Traffic

Chen Zhang, Ph.D., Research Analyst
Lance Jennings, Research Assistant
Lisa Aultman-Hall, Ph.D., Professor and Director

UVM Transportation Research Center
Farrell Hall
210 Colchester Avenue
Burlington, VT 05405
Phone: (802) 656-1312
Fax: (802) 656-9892
chen.zhang.1@uvm.edu
Website: www.uvm.edu/transportationcenter

Daryl Benoit, Transportation Planner
Chittenden County Metropolitan Planning Organization (CCMPO)
Phone: (802) 660-4071 extension 12
dbenoit@ccmpo.org

Total word count: 4,650 + 6* 250 = 6,150

Submitted to 89th Annual Transportation Research Board meeting for presentation and publication

Towards More Robust Spatial Sampling Strategies for Non-motorized Traffic

ABSTRACT

With the widespread promotion of New Urbanism and Smart Growth there is an assumption that levels of non-motorized traffic are increased. However, planners and analysts for non-motorized transportation modes still rely on very limited data resources and are therefore limited in identifying demand patterns and moving forward with more productive management and planning schemes. In this study, we utilized continuous non-motorized traffic counts collected along four shared use paths in Chittenden County, Vermont and analyzed the association between hourly (volume percentages of daily total) distribution patterns at each count station and land uses in the adjacent areas. Our findings show the linkage is not as evident as expected between surrounding land use and the hourly patterns of the counts gathered, likely due to the insufficient diversity of the land use patterns around the count stations. Therefore, there is a dire need for a more robust sampling strategy to be developed to obtain counts efficiently that are able to extrapolate short period counts into region-wide travel estimates. In this study, we propose a spatial-based clustering analysis which identifies five land use categories to assist planning practitioners in selecting sampling locations that are representative for generating consistent non-motorized traffic counts for entire network.

INTRODUCTION AND BACKGROUND

Increasing levels of non-motorized traffic is an important goal in research studies related to New Urbanism and Smart Growth comprising of traditional (neo) neighborhood design, transit-oriented development, new pedestrianism and more (1). In addition to serving as an alternative to replace motorized traffic thus reducing congestion and the related environmental, energy and social concerns, non-motorized transportation has also been at the center stage for promoting healthy living. In a recent cycling and walking study (2) utilizing nationwide county-based data, the authors pointed out the spatial distribution of cycling and walking commuting is positively associated with population density, natural amenities, education, wealth and estimates of local civic concerns. By focusing on the non-motorized mode choice of individuals, Rodriguez and Joo (3) revealed that objectively measured physical environment features such as travel time, access time and out-of-pocket costs have significant contributions to attractiveness of non-motorized mode choices. Many reports using large study areas similar to the above have identified important linkages between non-motorized travel behavior and the physical and social environment features, fewer researchers have been able to use location specific non-motorized traffic counts to yield progress in forecasting future volumes at a microscopic scale. Pulugurtha and Pepaka (4) studied the pedestrian counts collected at 176 intersections in the City of Charlotte, North Carolina and developed models predicting pedestrian activity using factors ranged from demographic characteristics, such as population and household units, to land use characteristics, including residential, commercial, industrial, etc. Their study results showed that urban residential density has the most significant impact on pedestrian activity at intersections.

Other studies have also identified plausible association between land use, including both zonal structure and physical characteristics, and the tendency for generating non-motorized traffic, again usually at the larger spatial scale or study area. Of them, Guo et al (5) conducted research to assess the effects of the built environment on motorized and non-motorized trip making, they discovered that few built environment factors would successfully lead to substitution of motorized traffic by non-motorized traffic. However, they argued that the increase in bikeway density or the connectivity in street network would have better potential in supplementing the existing motorized traffic with non-motorized traffic. It is suggested that land use tends to have high correlation with the level of non-motorized traffic volume because in most cases it is highly associated with the presence of supporting infrastructure and the transportation network features encouraging pedestrian and bicyclist activities. However, while it is generally accepted that non-motorized traffic levels vary from location to location with different land use patterns, by hour throughout the day and with weather by month of the year, the robust hourly patterns documented for vehicular traffic have not been defined comprehensively for non-motorized transportation modes. Moreover, the range of land use or spatial characteristics was often times overlooked when selecting locations for limited non-motorized traffic counts. Instead, researchers and planners often elect to collect data in the most traveled locations. The intersection of how these temporal patterns change with both the availability of dedicated facilities as well as land use patterns has not fully been studied and the full range of data to consider these relationships has not been collected.

In this study, we examined the relative hourly distribution of non-motorized traffic data throughout a day at numerous locations along several shared use paths in Chittenden County, Vermont and investigate the probable linkage between the daily volume distribution and surrounding land use patterns. This study also serves as part of an investigation to understand how to identify better sets of locations for future non-motorized traffic counts that would provide a better representative sample for extrapolation of region-wide bicycle and pedestrian volumes as well as travel exposure. Currently, agencies tend to conduct counts at the suspected highest volume areas. This large number of homogenous locations comprises the most heavily traveled non-motorized areas and cannot be used for total travel or exposure estimates. They are also of limited value for fully understanding what factors affect biking and walking levels because we have not studied the low or no volume areas. While the need for more random representative sampling applies to all facilities including roads and shared use paths, having more comprehensive knowledge of non-motorized traffic volume levels on shared used paths, the focus of this paper, is important to assist planning/operating agencies in mainstreaming these facilities as a legitimate component of the overall multimodal transportation system.

DATA SOURCES

Three primary data sources were used in this study: (1) geo-coded land use data and street network for Chittenden County provided by Vermont Center for Geographic Information (VCGI); (2) a geo-coded Champlain Valley pedestrian/bikeways network from Local Motion, a member supported non-profit organization in northwestern Vermont; and (3) multiple-day continuous pedestrian and bicyclist counts collected in between 2007 and 2009 at multiple locations along four shared use paths in Chittenden County, Vermont provided by Chittenden County Metropolitan Planning Council (CCMPO). Chittenden County is the most populated county of all fourteen in State of Vermont, with a population of nearly 150,000 and an area of 620 square miles. Figure 1 shows the nine locations where the existing counts were collected. In the same figure, the access points along the shared use paths and the sections of the paths providing open access from adjacent neighborhoods to the paths were also highlighted illustrating the level of accessibility to these facilities. Note that one path in particular, the one along the lake shore, has more limited access to adjacent land use. Others along roadways such as the Kennedy drive path have access at effectively every intersection.

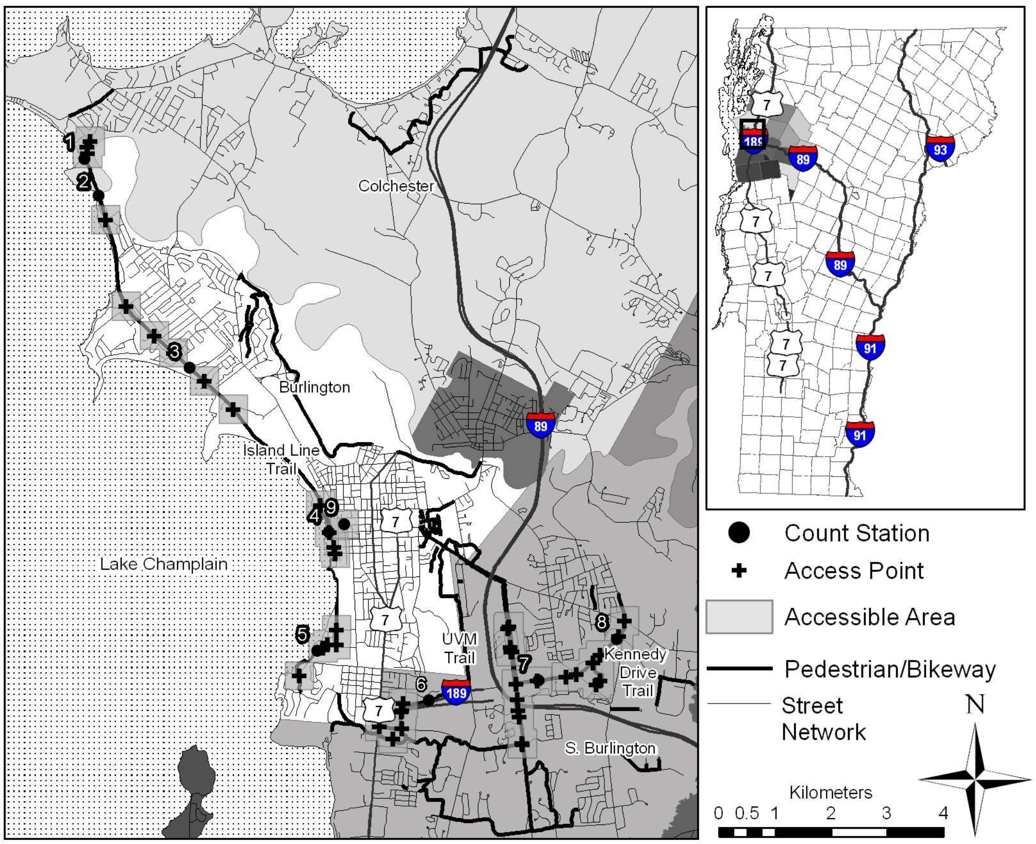


Figure 1 Multi-day continuous count stations along shared use paths and their access points

Table 1 lists for all nine stations the duration of the counts employed for this study and the number of weekdays, Saturdays, and Sundays during which hourly pedestrian and bike volumes were counted. In total 265 days of non-motorized traffic volumes were counted. The count stations were mostly placed in arbitrary locations by CCMPO while taking into consideration of (1) locations with higher non-motorized traffic volumes of interest for safety or management reasons, and (2) locations requiring higher maintenance costs such as pedestrian/bike bridges. The infrared pedestrian and bicycle counter (Eco-counter) has been employed at all stations to collect the combined pedestrian and bicycle volume data. This device is capable of collecting bi-directional bicycle and pedestrian traffic although total volume is used in this paper. By reading body temperature, the device's sensor detects the infrared radiation emitted by each person who passes by it, and the sensor's narrow profile further enables it to count two or more people following closely to one another (6). One of the previous studies (Aultman-Hall et al) utilizing counts collected by the same type of counter indicated an accuracy level of 98% when compared to manual counts.

Table 1 List of count stations and their duration

Shared use path	Station #	Count Duration	# of Weekdays	# of Saturdays	# of Sundays	# of Holidays
Island Line Trail	1	June 11 – August 18, 2008	36	7	7	1
	2	July 3 - July 31, 2007	20	4	4	1
	3	August 20 - September 23, 2008	24	5	5	1
	4	May 3 - May 20, 2007	12	3	3	0
	5	August 5 - September 1, 2008	19	4	4	1
UVM Trail	6	July 26 – August 3, 2008	5	2	2	0
Kennedy Drive Trail	7	Sep 2 – Sep 30, 2008 May 1 - May 26, 2009	38	8	8	1
	8	July 12 - July 24, 2008	9	2	2	0
Downtown Burlington	9	April 26 - May 27, 2007	18	4	4	1

The nine locations in Table 1 are located along shared use paths within the urban and suburban portions of Burlington. As Figure 1 shows, Island Line Trail runs through Burlington both north and south from downtown near the Lake Champlain shore going by mainly residential and recreational land use areas. The UVM Trail at its west end connects to US 7, a major multilane arterial which serves as a critical link for north-south motorized traffic especially in northwestern Vermont. The path runs partially parallel to Interstate 189, and winds to go up north passing by a series of University of Vermont properties. While within the urban area, after leaving the end proximate to the arterial highway, the path is surrounded by treed and farm areas. The Kennedy Drive Trail runs parallel and adjacent the entire length of the 4-lane arterial Kennedy Drive. It is separated from the road by a narrow green barrier (about 5 feet). Also included in the analysis were multi-day continuous counts from a downtown Burlington location (location 9 in Table 1) that is along a short pedestrian and bike path which connects large residential and hotel buildings to the city’s commercial center.

Of primary interest in this study is not only the total volume or peak volume of non-motorized users as is often studied, but rather the relative distribution or pattern of hourly volume throughout the day. We hypothesize that the daily pattern varies as a function of surrounding land use. Multi-day continuous hourly non-motorized traffic volume (pedestrians and bicyclists) collected during 2007-2009 were examined at nine count locations along the four shared use paths and the large dataset allows consideration of weekdays, Saturdays, and Sundays. In a study assessing impact of weather and season on pedestrian volume conducted by Aultman-Hall et al (7) utilizing year round continuous hourly pedestrian counts at a sidewalk in Downtown Montpelier, VT, the authors found consistency in day types such as weekdays/Saturdays vs. Holidays/Sundays. In this study, considering we were examining multiple locations across various types of surroundings of which some are deemed popular tourism and recreation spots, we decided to remove holidays from the data and to separate Saturdays from weekdays. Aultman-Hall et al. also found in their study that during winter time the overall pedestrian volume reduced by 16%. To avoid any discrepancy caused by season-related impact, we based our analysis solely on “summer months” which correspond to May through September. Holidays removed from the data included Memorial Day, Independence Day and Labor Day that did not occur on weekends.

Around every count location along the paths, we also examined the accessibility to the paths from the proximate trail-side areas. With the assistance of aerial photos and on-site visits, we found the majority of the access points along the paths are at the intersections of those trails with local roads. In addition to characterizing the land use immediately surrounding the count locations, we also considered the land use surrounding the proximate access points to the paths.

For this study, the VCGI Chittenden County land use data (8) were aggregated into seven different categories:

- residential (residence or accommodation);
- agricultural (agriculture, forestry, fishing and hunting);
- recreational (arts, entertainment, recreation);
- commercial (general sales or services);
- public institutional (public administration, education, other institution);
- transportation (transportation, communication, information, and utilities); and
- all others

Of the entire Chittenden County, residential and agricultural land uses represent the highest proportion of area both exceeding 30 percent. The next highest was public institutional and recreation land uses, together comprising 20 percent of the total. Commercial land use types were at the low end and mostly concentrated in the Great Burlington area which consists of the City of Burlington, the largest city in Vermont (with a population of 40, 000) also home of the University of Vermont campus, and several surrounding towns. The geo-coded Champlain Valley pedestrian/bikeways network was then combined with the street network data in order to identify the length of both roads (from the street network data) as well as trails in the study area as two variables used to classify surrounding land use type.

METHODOLOGY AND RESULTS

Land Use along Shared use Paths

In this study we were interested in the land use around those shared use path count stations, how the land use type affected the daily pattern in hourly non-motorized traffic volumes and whether the land use type was distinct from the range of land use combinations found in the whole County. Assessing the land use pattern that is possibly relevant to non-motorized traffic on shared use paths is not straightforward. First, many shared use paths in Chittenden County do not have open access continuously along the trail, therefore, the non-motorized users might be unaffected by the immediate surrounding land use and are not interacting with it versus pedestrians in a downtown may be shopping in adjacent retail stores. For example, a path used mainly for recreation and commuting purposes might pass very close to large agricultural lands, public institutions, major highways, or industry, however, the types of biking or walking traffic it carries may be unrelated to any of those specific proximate land use types. Second, non-motorized traffic demand levels are typically generated by activities and humans at or between land uses. For planning purposes, land uses adjacent but with no access to trails may not be a factor in non-motorized traffic volumes or the daily patterns in those volumes. On the other hand, the lack of access or proximate green space may be an attraction for some users. To account for these factors, at each non-motorized count location, we identified the land use at the nearest access points within a 1.5 km linear distance along the path from the count location. Sequentially, we investigated the probable connection between the hourly volume distributions and land use characteristics.

While the land use immediately surrounding a count location can be calculated with a simply GIS overlay, the land use at access points requires more manual consideration. The access points nearest the existing count locations and within 1.5 km, as shown in Figure 1, were used. For each access point a 0.5 km by 0.5 km square buffer area was generated. The area in the buffer was overlaid with the land use data to calculate area by land use in the buffers. Table 2 shows for each count station the total number of access points within its approximate range and the percentages of total area by land use type. Stations 7 and 8 have higher accessibility than the rest of stations. Five (1, 2, 3, 7 and 8) of the count stations have relatively high surrounding residential land use. Two (4 and 9) have only approximately 10 percent residential. Stations 1, 2 and 3 have 0 percent commercial land use, while the downtown Burlington station 9 has commercial land use for 34 percent of the buffer area. This table suggests this set of 9 count locations on shared use paths has a broad range of different mixes of land use types (note that our analysis of land use patterns county-wide presented later in the paper reveals these sites are not as diverse or representative as the table suggests at first consideration).

Table 2 Count Locations and their surrounding land use patterns

	Island Line Trail					UVM Trail	Kennedy Driver Trail		Downtown Burlington
Station #	1	2	3	4	5	6	7	8	9
# of access points	4	4	4	3	4	5	12	8	2
Residential	40	40	41	9	15	19	37	38	11
Commercial	0	0	0	12	8	19	7	9	34
Recreational	12	12	20	17	23	5	6	1	7
Public institutional	0	0	19	2	2	5	12	5	15
Agricultural	13	13	4	2	3	20	19	30	1
Transportation	8	8	10	28	16	20	5	6	25
Other	27	27	6	30	33	12	18	11	7
Total	100	100	100	100	100	100	100	100	100

Hourly distribution patterns at count locations

By aggregating the counts from every each location, we developed average hourly distribution of non-motorized traffic for weekdays, Saturdays and Sundays as illustrated in Figure 2A through 2I. For any of the day-of-week types at a particular count station, the hourly percentages were computed based on average hourly volumes for an average weekday, Saturday, or Sunday normalized over the duration of counts, as shown by Equation (1)

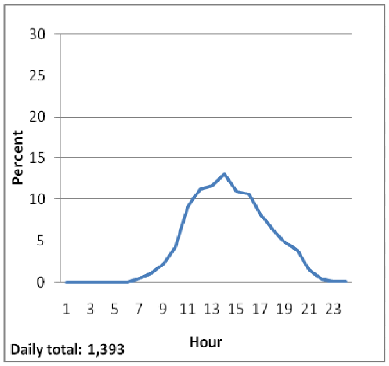
$$p_{ik} = \frac{(\sum_{n=1}^{N} h_{ikn}) / N}{(\sum_{n=i}^{N} D_{in}) / N} \quad \text{Equation (1)}$$

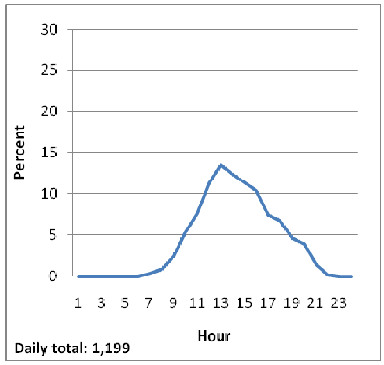
where p_{ik} (used for illustration in Figure 2) stands for the average hourly percentage at hour *k* for day-of-week type *i*; h_{ikn} stands for hourly volume of hour *k* on day *n* for day-of-week type *i*; D_{ikn} stands for daily volume on day *n* for day-of-week type *i*; and *N* stands for the count duration (as given in Table 1); While the nine sets of graphs show some differences in hourly non-motorized traffic patterns on weekdays and on Saturdays & Sundays, the distributions are not as different as we expected they might be based on initial evaluation of land use differences. Note the volume level indicated in Figure 2 varies widely from location to location. Note also the trails with higher volume have smoother distributions but that even locations with a large number of days counted have jagged patterns so we conclude the jagged patterns are not due to smaller number of days being counted.

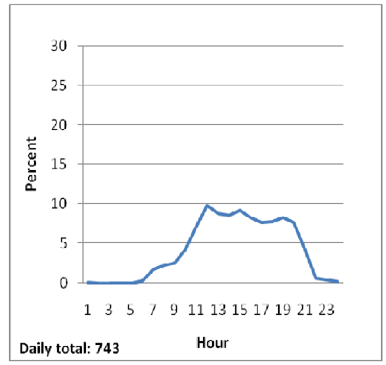
Figure 5 Hourly distribution of non-motorized traffic on an average Saturday, Sunday and Weekday (x-axis: hour of the day; y-axis: normalized hourly proportions of daily volume)

A Island Line Trail station 1 (Colchester)

Saturday (7 days of counts) Sunday (7 days of counts) Weekday (36 days of counts)

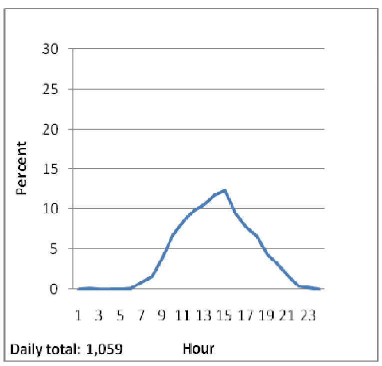


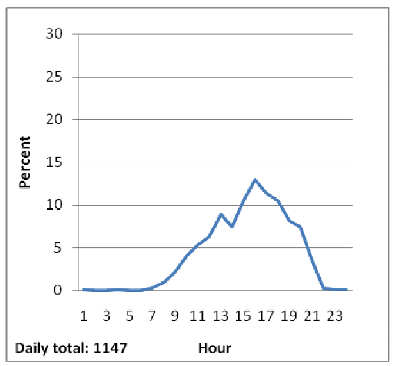


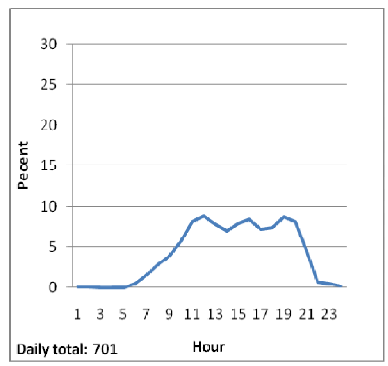


B Island Line Trail station 2 (Burlington)

Saturday (4 days of counts) Sunday (4 days of counts) Weekday (20 days of counts)

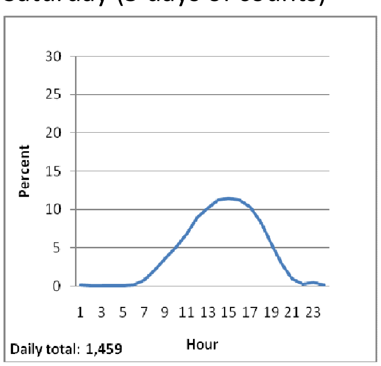


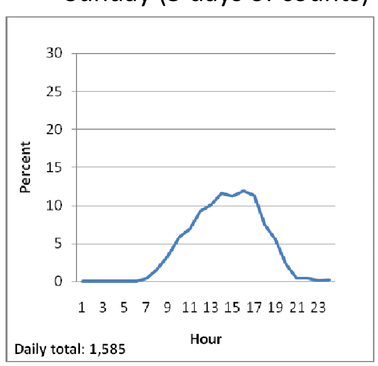


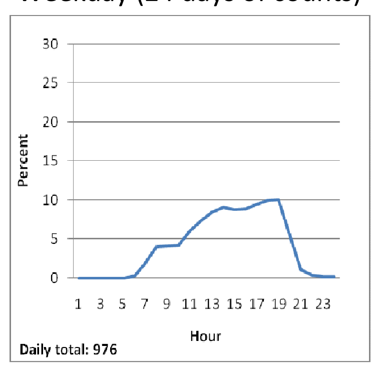


C Island Line Trail Station 3 (Burlington)

Saturday (5 days of counts) Sunday (5 days of counts) Weekday (24 days of counts)

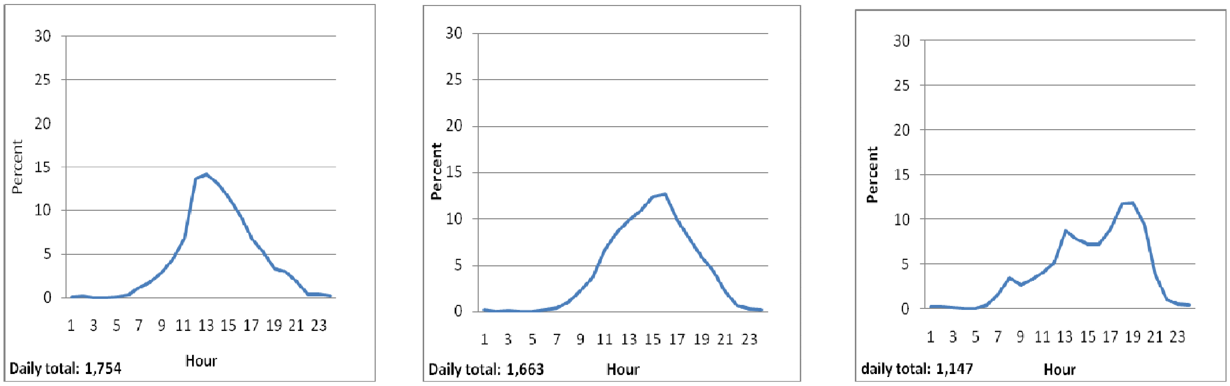






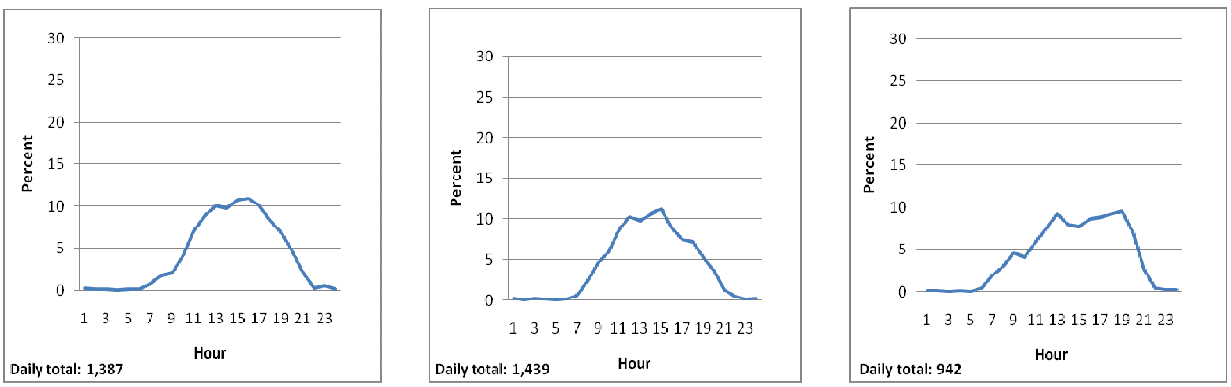
D Island Line Trail station 4 (Burlington)

Saturday (3 days of counts) Sunday (3 days of counts) Weekday (12 days of counts)



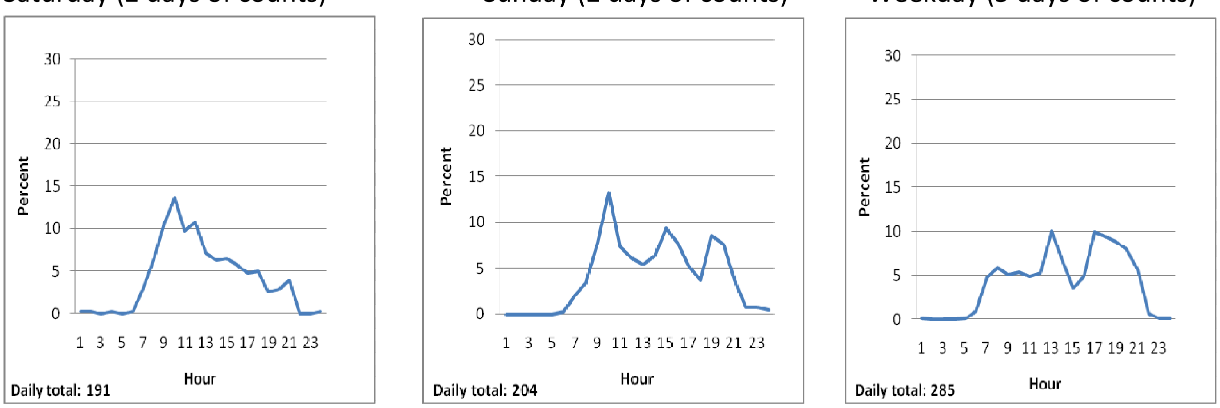
E Island Line Trail station 5 (Burlington)

Saturday (4 days of counts) Sunday (4 days of counts) Weekday (19 days of counts)



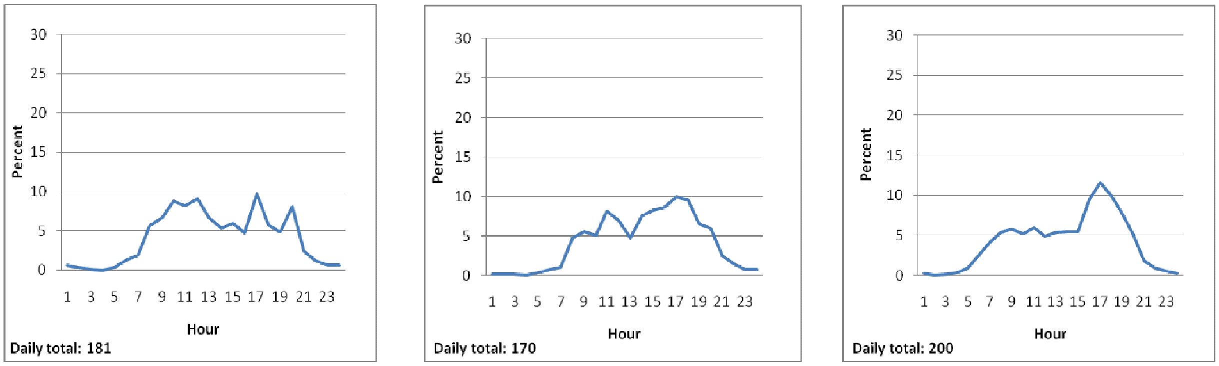
F UVM Trail station 6 (South Burlington)

Saturday (2 days of counts) Sunday (2 days of counts) Weekday (5 days of counts)



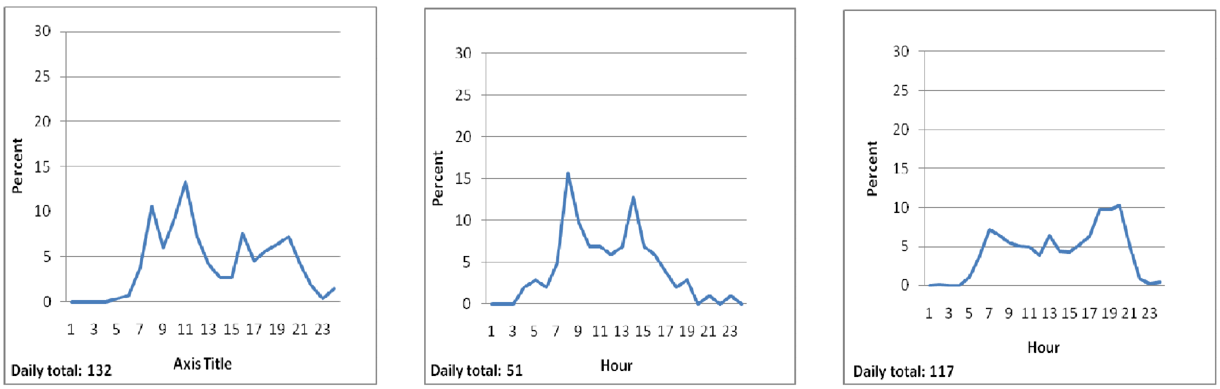
G Kennedy Drive Trail station 7 (South Burlington)

Saturday (8 days of counts) Sunday (8 days of counts) Weekday (38 days of counts)



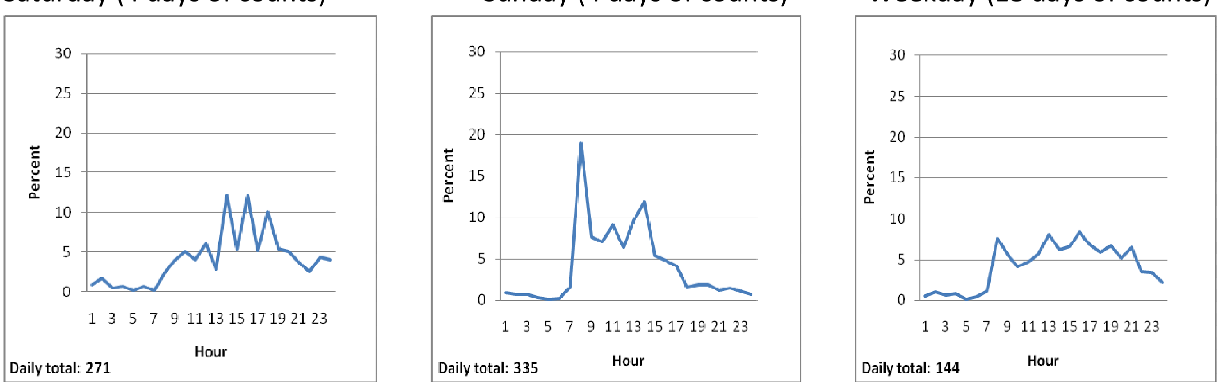
H Kennedy Driver Trail station 8 (South Burlington)

Saturday (2 days of counts) Sunday (2 days of counts) Weekday (9 days of counts)



I Downtown Burlington Station 9 (Burlington)

Saturday (4 days of counts) Sunday (4 days of counts) Weekday (18 days of counts)



While some limited differences between weekday and weekend and accessibility versus limited access shared use path volume distributions are evident from Figure 2, for the most part the impact of surrounding land use is not evident. The limited access trail has univariate distributions which are flattened on weekdays and more peaked on weekends. The shared use paths with more access points have much more variable volume distributions throughout the day. No obvious patterns were found based on the surrounding land use tabulated in Table 1. Locations 1, 2, 3, 7 and 8 might be considered residential, 3, 6 and 9 higher commercial. Locations 4, 5 6 and 9 have proximate significant roadways.

In order to sample representative land use areas, a stronger and more robust measure of land use pattern is needed. Count locations 7 and 8 demonstrated an AM and PM peak on weekdays and Saturdays. Location 9, in the very commercial downtown area, has some volume over all 24 hours which is different from all other count locations.

Sampling Based on Land Use Patterns

One main goals of this research was to develop a method and procedure to select a set of locations for non-motorized traffic counts that is representative of the entire county. The previous section suggested a limited connection between seemingly different land use areas and the hourly distribution pattern of volume. Bearing this in mind, we developed a framework which allowed us to consider the overall land use patterns for the whole County in relationship to this set of shared use path count locations.

A grid system was created dividing the entire County into a large number of sub-sections. The purpose of this grid system was to categorize the land use patterns in entire study area at a refined detailed level. Continuously distributed 0.5 kilometer by 0.5 kilometer square polygons were automatically generated for the entire county dividing the Vermont street system into short links contained within nearly six thousand of regular-shaped cells. The total length of street links was then computed for every each cell. Since the cells with no street links captured inside are normally located in areas extremely remote or hardly being accessed, we examined only those ones, in total 3,450, with non-zero total length of streets (non-zero cells). This represented 60% of the area of the county. K-means clustering method was then adopted to categorize those non-zero cells into a number of relatively more homogenous groups based on the surrounding land use types.

Clustering Analysis

K-means clustering method has been a popular data mining tool in various fields after it was first introduced in 1967 by James MacQueen (9). In our case, the data set consists of all the non-zero cells each attributed by 6 variables which represent the percentages of total cell area for the different land use. For example, a cell that is mostly occupied by residential buildings might be coded as having over 50 percent residential land use and small percentages of other types such as commercial and recreational. K-means cluster analysis was conducted to group all cells using the percentages by land use type as reference variables based on the Euclidean “distance”[1] calculated using land use percentages. In the clustering procedure, the non-zero cells were partitioned into five groups such that a pair of elements in the same cluster tends to be more similar than a pair of elements belonging to different clusters. The aim of the K-means cluster algorithm is to classify a set of data points into K categories through the K clusters defined a priori (9). The main idea is to define K initial cluster centers $C^o = \{c_1^o, c_2^o, \ldots, c_k^o\}$ one for each cluster, and then to relate the rest of the data points to these centers depending on the closest Euclidean distance,

$$d(x_i, x_j) = \left[\sum_{v=1}^{V} (x_{i,v} - x_{j,v})^2 \right]^{1/2} \qquad (2)$$

[1] Note this is distance in vector space not distance within the transportation network.

where $d(x_i, x_j)$ is the Euclidean distance between point *i* and *j*, and *V* is the dimension of the reference variable vector.

The following procedure then applies two steps iteratively. First, it assigns each data point to the nearest cluster center according to the Euclidean distance using equation (3).

$$C_j^{i+1} = \{c_1^{i+1}, c_2^{i+1}, \ldots, c_k^{i+1}\},$$
$$c_j^{i+1} = \{x \mid d(x, \mu_j^i) \le d(x, \mu_{j'}^i), 1 \le j, j' \le k \quad and \quad j \ne j'\} \qquad (3)$$

Then when all points have been assigned to a cluster, the cluster centers are recalculated based on their assigned data points using equation (4),

$$\mu_j^i = \frac{1}{|c_j^i|} \sum_{x \in c_j^i} x \qquad (4)$$

where *j* refers to the clusters, *i* is the iteration counter, $|c_j^i|$ is the number of points of cluster *j* after the *i*th iteration, and μ_j^i is the center of cluster *j* after the *i*th iteration. The procedure using equations (3) and (4) is repeated until the cluster centers do not change any more (that is, they converge), and then the K clusters each with a subset of data points become the final categories.

Following the above procedure, we categorized those non-zero cells into five clusters based on their distribution of the six land use percentages. As Table 2 shows, the five typical land use groups defined by the clustering algorithm have been named by our group as mixed use, public institution, residential, recreation, and agriculture with each having a unique blend of land use types. Figure 3 illustrates the distribution of these cluster types in the county. Mixed use clusters possess a relatively high concentration in the mostly urban area of Burlington (encompassing the City of Burlington and its surrounding towns including South Burlington, Essex Junction, and Williston) which is more developed and economically active than the rest of the county. Characterized by a great mixture of all land use types, the mixed use cluster has the highest commercial, transportation land use density, and also includes some level of residential land use. The public institution cluster is represented by a high level of public institution land use density, mainly counting schools and government buildings, colleges or other facilities. Similarly, residential, recreation and agriculture clusters, respectively, have their highest representative land use type as residential, recreational, and agricultural. Recalling that grids without any roads or trails have been excluded, one might expect there are many agricultural and residential grids in the county with fewer institutional and recreational cluster. There are also a reasonable number of mixed use areas.

Table 3 Cluster categories and their land use patterns

Land use type	Mixed use	Public institutional	Residential	Recreational	Agricultural
	Cluster centers (with land use percentages)				
Residential	23	9	75	17	17
Commercial	8	1	1	1	0
Recreation	4	2	2	63	1
Public institution	3	80	1	0	0
Transportation	16	3	5	5	4
Agriculture	16	3	16	9	75
Others	30	2	2	6	2
# of cells in County	518(15%)	101(3%)	1451(42%)	140(4%)	1242(36%)
# of Shared Use Path Count Locations	1(11%)	1(11%)	6(67%)	1(11%)	0(0%)
# of all CCMPO Count Locations	2 (12%)	1(6%)	12(71%)	2(12%)	0(0%)

The clustering results in Table 2 (the shaded row) indicate the percentage of cells in the county which fall into each of the five categories. Baring in mind, although this method is applicable elsewhere to generate clusters representing various land use patterns, the results presented here would be a county-specific distribution. The bottom two rows of the table show the type of land use cluster for the 9 shared use path locations analyzed in this paper as well as the full set of 17 locations where the CCMPO counted non-motorized traffic during the last 2 years (a larger number than the multi-day locations used in this study due to inclusion of single-day counts locations). Of the 9 shared use path count locations described above six were in one category. This may explain the similarity in most of the daily distribution patterns. We propose for future counts that random grid cells be selected so that the locations are representative of the land use clusters. The CCMPO locations presented in the table were selected for sound management goals, but need to be diversified for obtaining travel exposure estimation for the whole county.

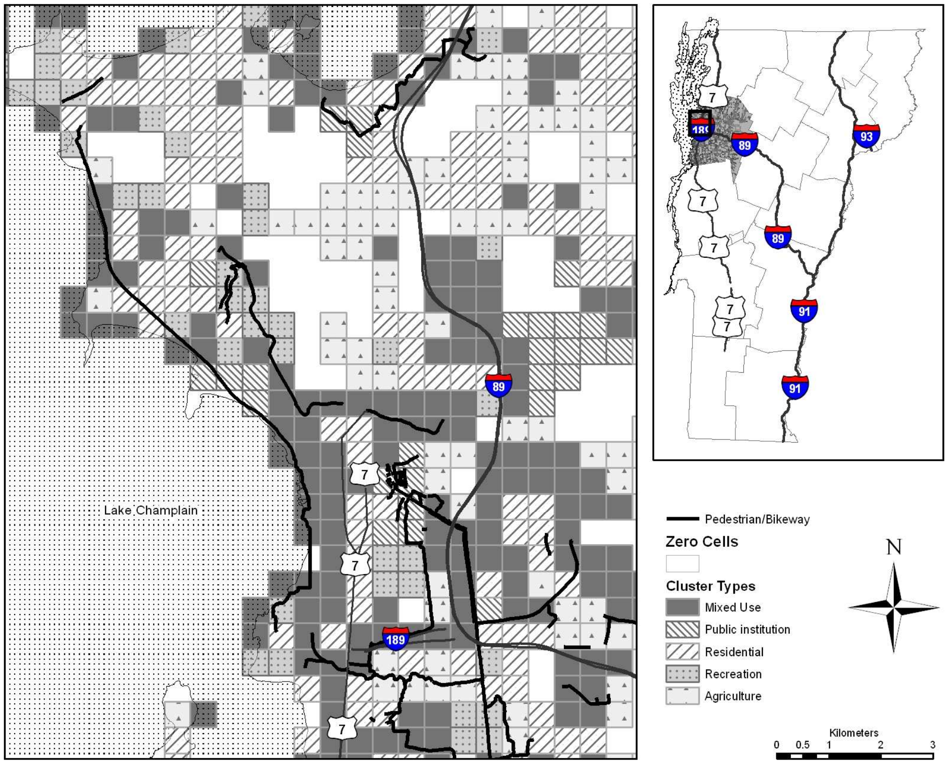


Figure 3 Land use patterns identified by cluster analysis for the grid cells in study area

CONCLUSIONS

Previous literature clearly establishes that non-motorized travel demand is a function of land use patterns. There is recently more wide-spread technology available to obtain automated counts of non-motorized traffic volumes which facilitates larger datasets managed to develop time of day and day of week factors for Highway Capacity Manual style analysis. Development of such factors to predict bicycle and pedestrian traffic volumes and extrapolate from limited counts to region-wide total travel for safety and other applications is important for mainstreaming non-motorized transportation into the overall transportation system. As we proceed to deploy counters and obtain data it is essential to develop sampling frameworks based on spatial land use patterns that allow random counts to be representative of a whole region.

This study utilized a set of shared use path counts for which the locations had been chosen for solid management reasons: traffic management and facility maintenance. Our analysis here demonstrated that these counts did not represent the range of land use mixes found within the county even when only a limited number (five) of land use categories were defined using a K-means cluster analysis. The shared use path daily volume distributions followed similar patterns for weekends versus weekdays. Differences were noted between limited access versus easy access shared use paths. Clearly much more data from different regions of the county will be needed to fully characterize non-motorized traffic volumes, however this study demonstrates that we must carefully consider a more random sampling procedure up-front or the data we collect will be limited in representing the full range of land use areas found in our regions. The consistent daily distribution observed in the results here further support the notion that correction factors that allow extrapolation of counts in pace and time for non-motorized modes are feasible.

The sampling method we propose in this paper for count location selection was conducted with publicly spatial land use data that is widely available. Similarly, the GIS-based grid overlay and the K-means clustering method are straightforward and could be adopted by planning agencies to determine the proportion of their region within a number of land use categories and then to randomly select grid locations for counting. Once grids for counting were selected a random point on the roads or trails within the cell would also require being selected. While the count results would require weighting based on miles of facility within the grid, we would argue this effort makes the data within a counting program more valuable for a wider range of policy questions than simply having counts at the highest volume locations of the region.

ACKNOWLEDGEMENTS:

This research was funded by the U.S. DOT through the University Transportation Center program at the University of Vermont.

REFERENCES

1. Katz, P. The New Urbanism: toward an architecture of community, New York, McGraw-Hill, Inc. 2004

2. Zahran, S., Brody, S. D., Maghelal, P., Prelog, A., Lacy, M., Cycling and walking: explaining the spatial distribution of healthy modes of transportation in the United States, Transportation Research Part D, volume 13, 2008, pp 462-470

3. Rodriguez, D., Joo, J, The relationship between non-motorized choice and the local physical environment, Transportation Research Part D, volume 9, 2004, pp 151-173

4. Pulugurtha, S. S., Pepaka, S. R., Assessment of models to measure pedestrian activity at signalized intersections, Transportation Research Record, volume 2073, 2008, pp 39-48,

5. Guo, J. Y., Bhat, C., Copperman R. B., Effect of the built environment on motorized and nonmotorized trip making, Transportation Research Record, volume 2010, 2007, pp 1-11

6. Bell A., Technology Innovations: Infrared Bicyclist & Pedestrian Counter, *Bike/Ped Professional: Journal of the Association of Pedestrian and Bicycle Professionals,* Association of Pedestrian and Bicycle Professionals, Hamilton Square, NJ, 2006, pp 4-5

7. Aultman-Hall, L., Damon L. and Rebecca L., Assessing the Impact of Weather and Season on Pedestrian Traffic Volumes, forthcoming *Transportation Research Record (*TRB paper number 09-2765), 2009

8. Vermont Center for Geographic Information, http://www.vcgi.org/dataware/?page=./search_tools/search_action.cfm&query=newupd

9. McQueen, J. B. Some methods for classification and analysis of multivariate observations, *5th*

Berkeley symposium on mathematical statistics and probability, Berkeley, University of California Press, 1967, pp 281-297