Evaluating Walkability in the Age of Open Data: OpenStreetMap and Community-level Transportation Analysis

Austin Dunn[1], Bailey Hanson[1], Christopher J. Seeger[2]

[1]Iowa State University, Iowa/USA

[2]Iowa State University, Iowa/USA · cjseeger@iastate.edu

Abstract: In recent literature, "walkability" is evaluated and measured in terms of both infrastructure connectivity and human-scale streetscape features. Such analyses rely on geospatial data about walking and bicycling facilities, which, in many rural communities, is either non-existent or inaccessible. Situated in the larger "open data" movement, OpenStreetMap (OSM) is a crowdsourced, web-based map and repository for geospatial data, including infrastructure features such as roads, trails, sidewalks, building footprints, parks, and more. We assert that OSM is a powerful source of information that can be used to understand and to model the built environment, especially in rural areas where official spatial data is lacking. In addition to infrastructure geometry, the OSM platform supports the storage of attribute data through the use of tags. We operationalize the idea that OSM can be used as a viable data source for evaluating walkability at the community scale. Using Perry, IA, as a case study, we have employed multiple measures of walkability using both OSM pedestrian data and Iowa Department of Transportation street data for comparison. We found that the inclusion of OSM pedestrian data in walkability analyses greatly improved the results for some measures. Further research will focus on ways to incorporate OSM tags and OSM-sourced community assets and barriers into walkability analysis.

Keywords: Walkability, open data, transportation, pedestrian, design

1 Introduction

As "walkability" emerges as a key concern for designers and planners of the built environment, there is a need to understand and measure it (FRANK et al. 2006). Walkability refers to how friendly an area is to pedestrians. On a macro-scale, this concerns the connectivity of pedestrian infrastructure such as sidewalks and trails. On a micro-scale, features of the pedestrian streetscape, such as crosswalks, sidewalk condition, and perceptions of safety, play an important role in walkable places. Accordingly, walkability has been evaluated and understood both in terms of connectivity and the quality of the pedestrian streetscape (SCHOLOSSBERG & BROWN 2004, MANAUGH & EL-GENEIDY 2011, LESLIE et al. 2007). For landscape architects and planners, the ability to document and measure these parameters in the existing environment and in future scenarios, can facilitate the delivery of safer and more vibrant pedestrian spaces.

The open data movement calls for data that can be freely used and redistributed to facilitate the exchange of information and ideas. Landscape architects can and should tap into growing sources of open data for the built environment. OpenStreetMap (OSM) is a repository for geospatial data for the entire planet and is a powerful source of information that can be used to understand and to model the built environment (HAKLAY & WEBER 2008, HAKLAY 2010). The OSM platform is built and contributed to by a community of mappers. Contributors add and edit information related to all types of infrastructure features such as roads, trails, sidewalks, building foot prints, and more. Local contributors confirm edits using multiple sources

Journal of Digital Landscape Architecture, 3-2018, pp. 119-129. © Wichmann Verlag, VDE VERLAG GMBH · Berlin · Offenbach. ISBN 978-3-87907-642-0, ISSN 2367-4253, e-ISSN 2511-624X, doi:10.14627/537642013.

such as aerial imagery, GPS devices, and low-tech field maps along with their local knowledge to ensure accurate data.

Pedestrian connectivity can be evaluated using information about the built environment from OSM, such as sidewalks, trails, and alleyways, as well as pedestrian infrastructure attributes such as pavement condition, slope, and curb cuts. This information can help designers consider walkability more holistically and in more detail. This research addresses how OSM's open data can contribute to understanding walkability, how both macro- and micro-scale features of the pedestrian environment can be evaluated, and which methods are most appropriate for analyzing OSM-sourced pedestrian data.

2 Methods for Evaluating Walkability

The authors operationalized various methods for evaluating walkability in the city of Perry, Iowa, in the United States Midwest. Geospatial data for the street network was downloaded from the Iowa Department of Transportation. Complete pedestrian network (or) infrastructure data was digitized in OSM, and then downloaded to ArcGIS via the ArcGIS Editor for OpenStreetMap tool. We defined pedestrian infrastructure to include sidewalks, trails, alleyways, and lower-traffic roads. Lower-traffic roads are roads not classified as "motorway," "trunk," or "primary" by OSM. A network dataset was created for the street network and for the pedestrian network. Additional data was sourced from the 2014 I-WALK program in Perry (SEEGER 2014). I-WALK is a participatory mapping program that gathers volunteered geographic information (VGI) about community walkability. The procedures and data collected are consistent with that of the Microscale Audit of Pedestrian Streetscapes (MAPS) framework (CAIN et al. 2014).

In the city of Perry, we used a 100-meter grid to measure and represent connectivity and streetscape factors across the entire urban area. The centroid of each grid cell was used to create a half-mile buffer, a distance that has been established as "walkable" by Weinstein AGRAWAL et al. (2008). The half-mile buffer for each point became a study area, with the results of the connectivity measures attributed to the corresponding grid cell.

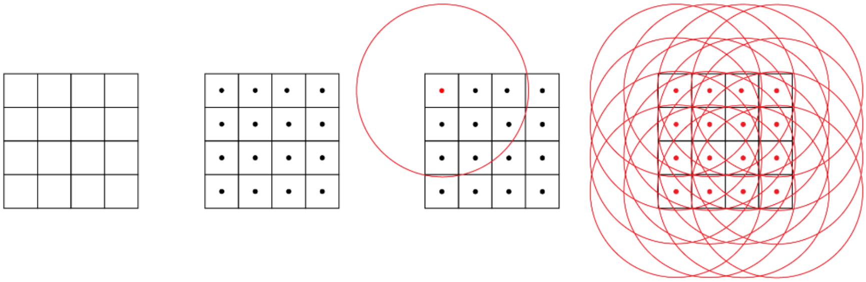

Fig. 1: A 100 m grid was created for the entire community. For the centroid of each grid cell, a half-mile buffer was created. Each half-mile buffer area was a study area for the three connectivity measures outlined below. The connectivity scores for each buffer area were related back to the grid cells for representation.

We evaluated three connectivity measures: intersection density, link-node ratio, and pedestrian shed (ped shed). Further, we evaluated the distribution of five pedestrian streetscape features in Perry. These included sidewalk completeness, sidewalk condition, pleasantness of walk, presence of intersection curb cuts, and presence of intersection signage or painted crosswalks.

2.1 Connectivity Measures

1. Intersection density is a measure of the number of intersections per square mile. For this study, we counted the number of intersections within a half mile of each grid cell centroid and expressed it as number of intersections/square mile.
2. Link-node ratio is the ratio of links (road segments) to nodes (intersections), with a high number indicating better connectivity. Stangl notes that this measure is actually measuring the presence of cul-de-sacs and dead-ends, which detract from connectivity (2012, 229). For this study, links and nodes were counted if they were completely contained within the half-mile buffer of the grid cell centroid.
3. Ped shed measures the area accessible via the street or pedestrian network as a percentage of the area defined by the "as the crow flies" distance. That is, a network buffer as a percentage of a Euclidean buffer is the ped shed (LARCO AND PARKER 2013). A higher percentage means greater accessibility. For this study, we used half mile as a walkable distance.

Each connectivity measure was calculated using both the street network and the pedestrian network.

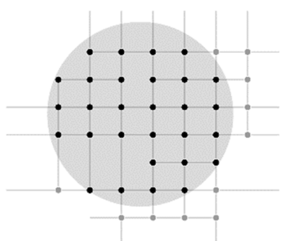

Fig. 2:
Intersection density takes a count of intersections within a half mile and expresses density as intersections per square mile

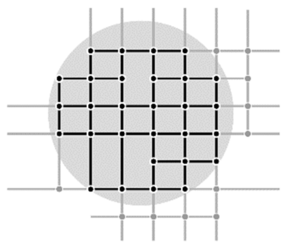

Fig. 3:
Link-node ratio counts the number of road segments (links) and intersections (nodes) within a half mile. Link-node ratio is expressed as links/nodes.

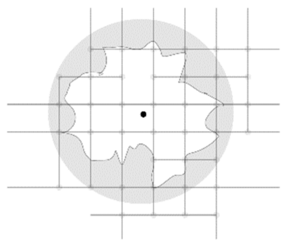

Fig. 4:
Ped shed utilizes a network analysis to calculate the area reachable by half mile via the network. Ped shed is expressed as area reachable via network/area of half-mile Euclidean buffer.

2.2 Pedestrian Streetscape Features

Features of the pedestrian streetscape in Perry were collected in 2014 via I-WALK, a participatory mapping program facilitated by Iowa State University Extension and Outreach Community and Economic Development. The program equips community members with smart phones to collect spatial data about pedestrian infrastructure (SEEGER 2014).

3 Results and Analysis

3.1 The Contribution of OSM Pedestrian Data

For each of the three methods, we evaluated the difference in connectivity scores between the street network and pedestrian network. We found substantial differences in patterns of connectivity between the street network and OSM pedestrian network with the ped shed method. We found that the addition of pedestrian data had a small effect on patterns of connectivity in the community for both the intersection density and link-node ratio methods.

When using the intersection density method to evaluate the street network in the city of Perry, we found that the central core of town has the highest density of intersections. The central core of Perry is older and has a grid-like street network typical of traditional neighborhoods. Lower intersection density scores around the edge of the town could be attributed to the fact that we measured the number of intersections within a half mile of each grid cell; the grid cells near agricultural areas outside the city have substantially fewer intersections within a half mile. This represents a limitation of using the intersection density method at the community level, especially in rural communities. An "edge effect" will give lower scores to outer areas, which might not reflect the spatial arrangement of streets as the measure is intended to do.

Intersection density was also used to evaluate the pedestrian network. The addition of sidewalks, trails, and alleyways made for many more intersections in the network and resulted in a very high number of intersections per square mile. An intersection of low-traffic roads would account for one intersection using the street network, but when counting sidewalks on both sides of the streets, this one intersection became 9 intersections. Accordingly, there is no way to compare the numerical results between the street and pedestrian networks "as is." When the results of the intersection density for the pedestrian network were normalized, we saw comparable patterns. Although sidewalks and pedestrian infrastructure add density, they generally correspond with the street network, so the pattern of intersection density is mostly unchanged.

Like intersection density, the results of the link-node ratio method for the street network are affected by the rural context of the city of Perry. We see an edge effect with high scores around the outside of town where there are longer links and fewer nodes. To address the edge effect, we counted only the road segments that were completely contained within the half-mile buffer, instead of any segment that intersected the buffer. This resolved most of the edge effect and from these results we found higher link-node ratios in the core of town (with some high scores around the edge, attributed to the aforementioned reason), and lower scores around the edge.

When we used the link-node ratio method to analyze the pedestrian network, we saw a similar pattern to that of the street network; however, the extent of higher scores in the core was larger. The overall average link-node ratio for the entire community was reduced with the inclusion of pedestrian infrastructure. This is consistent with the findings of TAL & HANDY (2012), who found that the link-node ratio method was not a good indicator for pedestrian networks. As in our study, Tal and Handy found that link-node ratio scores may be reduced with the inclusion of pedestrian infrastructure. For example, when an alleyway is added to the street network, one link is added with two nodes, thereby lowering the link-node ratio, despite the area having greater connectivity. Further, some cities have established standards for link-node ratios with minimum values of 1.2 or 1.4 (DILL 2004). These values are based on street network geometry and are not easily compared to pedestrian network scores.

We found higher ped shed scores in the core compared to the edge of the town when analyzing the street network, which is consistent with the findings of the other methods. There are holes in the coverage of our analysis for this method because we used a snapping distance of 50 meters for the network analysis. Grid cells with a centroid greater than 50 meters from the network did not yield results. We also found that highways were a significant source of connectivity for areas throughout the community and contributed to the overall ped shed score.

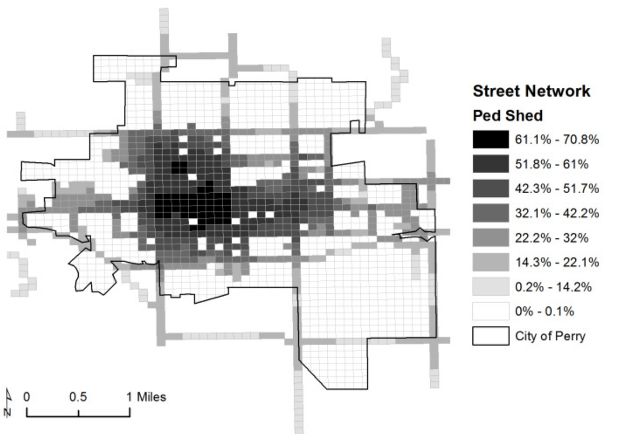

Fig. 5:
Results of the ped shed analysis using the street network

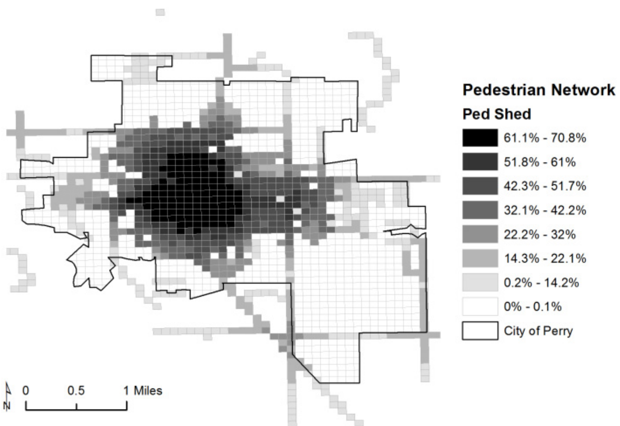

Fig. 6:
Results of the ped shed analysis using the pedestrian network

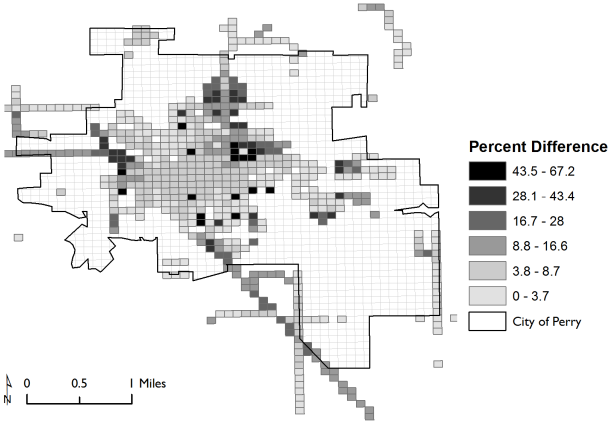

Fig. 7:
Increases to ped shed scores between street and pedestrian network analyses

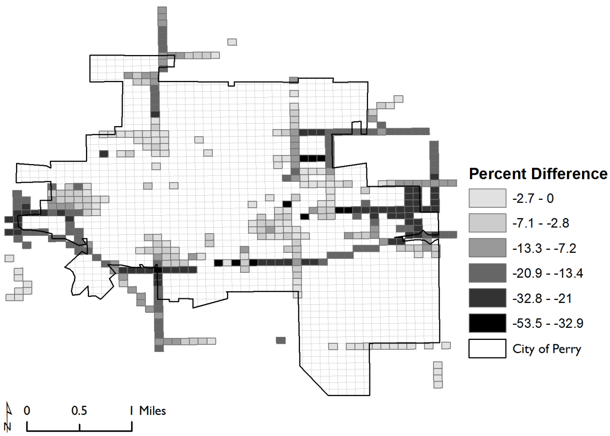

Fig. 8:
Decreases to ped shed scores between street and pedestrian network analyses

Using the ped shed method to analyze the pedestrian network, we found substantial changes to the results. First, the gaps in coverage were filled as more infrastructure was present within 50 meters of the grid cell centroid. Where infrastructure such as trails, alleys, and park paths were added, there was an increase in the ped shed score (see Fig. 6). In the pedestrian network analysis, highways were not included. This resulted in lower scores where sidewalks and low-traffic roads were only connected to highways.

The ped shed method yielded interesting results when we compared the street network to the pedestrian network. We saw an overall average decrease in ped shed scores between the street and pedestrian network analyses, while the core saw an increase. Where infrastructure such as highways was removed, the scores decreased substantially (see Fig. 7 & 8). We believe this to be reflective of the actual pedestrian experience. Greater sidewalk and trail density in the core does positively affect walking, while the presence of highways, especially with no pedestrian infrastructure, has a negative effect. Based on the results, we posit that ped shed scores derived from the pedestrian network more accurately reflect the experience of pedestrians on the ground.

3.2 Integrating Microscale Pedestrian Streetscape Features

Using volunteered geographic information (VGI) from the I-WALK program conducted in the city of Perry in 2014, we were able to analyze pedestrian streetscape features. The data collected is consistent with the Microscale Audit of Pedestrian Streetscapes (MAPS) framework. The addition of streetscape features to our analysis of connectivity yielded important findings.

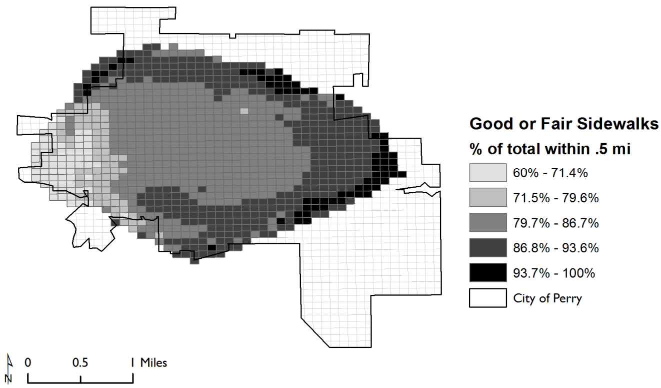

Fig. 9: The percentage of sidewalks within a half mile that were rated as good or fair. When compared with Fig. 6, it is clear that connectivity and quality-of-streetscape features do not necessarily correspond.

Namely, the most connected areas of town are not necessarily the same areas that have the highest quality infrastructure or perceived pleasantness. We measured the percentage of sidewalks rated in good or fair condition within a half mile of each grid cell centroid. We found that areas with a higher percentage of sidewalks in good or fair condition are on the periphery of town and not in the core (Fig. 9). This means it is more likely that you find poor sidewalk conditions in the areas that are more connected and find better quality sidewalk conditions in areas with less connectivity. This comparison reveals places where one could focus efforts to improve connectivity and places where to improve the quality of sidewalks within the highly connected area. Overall, it is important to consider on-the-ground features such as sidewalk condition, crosswalks, and perceptions of safety, because they are not captured in the previously discussed connectivity measures, but are crucial aspects of walkability.

3.3 Strengths and Limitations of Each Measure

Based on our experience using multiple measures of pedestrian connectivity and walkability in the city of Perry, we found that each approach presents strengths and limitations, which are summarized in Table 1. The ped shed method provides the most meaningful results and measurement of connectivity. The use of pedestrian data with the ped shed method showed changes that intuitively reflect the pedestrian experience. As TAL & HANDY (2012) noted, link-node ratio is not well-suited to evaluate pedestrian data. It is a useful measure to understand the patterns of street networks in various neighborhood types. While street patterns are intimately connected to walkability, they do not present the whole story. Likewise, we found that measuring intersection density with the pedestrian network does not provide any meaningful results when compared to the intersection density of the street network. The LEED-

ND standards for intersection density were developed for evaluating street networks and accordingly, are not applicable to pedestrian infrastructure. Further, STANGLE & GUINN (2011) have written about the limitations of intersection density to accurately measure connectivity. Both intersection density and link-node ratio are relatively simple measures that could be done using a paper map. The ped shed method provides the most meaningful results, but it is the most computationally intensive of the three methods discussed here. Calculating ped shed scores require GIS software with Network Analysis capabilities.

Table 1: Strengths and limitations of each measure

Method	Strengths	Limitations
Intersection density	Easy to calculate; standards established by LEED-ND	Unable to compare results between street and pedestrian networks
Link-node ratio	Easy to calculate, measures dead-ends and cul-de-sacs	Not well-suited for pedestrian data (TAL & HANDY 2012)
Ped shed	Most accurately reflects pedestrian connectivity	Computationally intensive; requires network analysis

4 Discussion

Connectivity, the extent to which pedestrian transportation facilities provide uninterrupted service, is an important aspect of walkability. In the literature, measuring pedestrian connectivity has often meant measuring street network connectivity. Such measures make the assumption that pedestrian infrastructure mirrors vehicular infrastructure. While street and pedestrian networks are closely linked, our research is consistent with CHIN et al. (2008) in finding that the addition of pedestrian networks does indeed affect measures of connectivity. A major challenge to assessing pedestrian networks is the dearth of available spatial data for pedestrian infrastructure. While GIS datasets for streets are made available by public agencies, GIS data for pedestrian networks, including sidewalks, trails, and footpaths, are often non-existent or not available. The OSM platform can address this need for publicly-accessible, “open” spatial data for pedestrian facilities. For neighborhoods and communities where pedestrian data is not available, it can be created quite easily using the OSM platform. Facilitators of this process can look to the facilitated Volunteered Geographic Information (f-VGI) framework as a guide (SEEGER 2008, ELWOOD 2008). Once a part of OSM, the data can be freely edited and improved by other users, as well as downloaded and used for a multitude of applications. Further, OSM has recently added functionality that enables users to add attributes to new or existing sidewalk data.

Beyond the geometry of pedestrian features, micro-scale features of pedestrian infrastructure are important to the walkability of an area. OSM can store attribute information about such features via additional tags. For example, users can add tags to sidewalk features to describe surface type, smoothness, slope, wheelchair accessibility, curb cuts, tactile paving, and pedestrian signage. The data structure of OSM is well-suited to store the information collected in a MAPS-type assessment of streetscape features (Fig 10). These features are essential to understanding walkability at the pedestrian scale. When incorporated with analyses of connectivity, information about pedestrian streetscape features can provide a more holistic picture of walkability. Currently, the effectiveness of this approach is limited by inconsistent data completeness and quality. One way to address data deficiencies is for researchers and professionals to facilitate mapping workshops in which residents are trained and instructed to use OSM to map walking and bicycling facilities in their community.

```
sidewalk:both:surface=paving_stones
sidewalk:left:smoothness=*
sidewalk:right:width=* / est_width=*
sidewalk:right:bicycle=yes
sidewalk:both:incline=*
sidewalk:both:kerb=*
sidewalk:both:wheelchair=*
sidewalk:both:tactile_paving=*
sidewalk:right:traffic_sign=*
```

Fig. 10: OpenStreetMap tags that could support attribute information for sidewalk data

Of the three connectivity measures we evaluated, the ped shed method was the most successful in incorporating pedestrian infrastructure data. The ped shed method most accurately reflects how pedestrians are able to move through the landscape. Intersection density and link-node ratio methods are simple, easy-to-calculate proxies for connectivity (STANGL 2012), but they fall short as true measures of connectivity. Further, standards for each score have been developed using street networks, so the values obtained from analysis of pedestrian infrastructure are not easily evaluated using established benchmarks.

5 Conclusion and Outlook

If landscape architects, urban designers, and planners hope to effect walkable places, they must equip their practice with tools for evaluating walkability. A range of these tools are presented above. In addition to the tools, practitioners need pedestrian GIS data. OpenStreetMap (OSM) is an existing platform that can be a source for data for the analysis macro- and micro-scale features of the pedestrian transportation environment. Researchers and practitioners alike should embrace and invest in OSM as a repository of open data about the built environment. As a platform for highly-structured volunteered geographic information, OSM is at once a tool for facilitating data collection and a source for consistent geospatial information. Contributing to OSM can become an opportunity for community outreach. Researchers and practitioners can facilitate workshops whereby community members are actively engaged in creating geospatial walkability data about their community. Indeed, OSM can provide a platform for facilitated Volunteered Geographic Information (f-VGI).

Currently, the authors are replicating the methods used in Perry to evaluate walkability in communities across the State of Iowa. Future work can be done to develop tools that utilize OSM data for “what-if” scenario planning. At present, users must download and edit OSM data in order to evaluate potential changes. A user-friendly and web-based interface for assessing walkability in current and potential pedestrian environments would make these tools more accessible to designers and decision makers. Future work should also focus on ways to integrate measures of connectivity with streetscape features as well as amenities. The web-based WalkScore service assesses connectivity and destinations but fails to incorporate important features such as sidewalk condition and crosswalks. These and other streetscape features are especially important for the mobility of vulnerable groups such as youth, persons with disabilities, and aging populations.

References

CAIN, K. L., MILLSTEIN, R. A., SALLIS, J. F., CONWAY, T. L., GAVAND, K. A., FRANK, L. D. & GLANZ, K. (2014), Contribution of streetscape audits to explanation of physical activity in four age groups based on the Microscale Audit of Pedestrian Streetscapes (MAPS). Social Science & Medicine, 116, 82-92.

CHIN, G. K., VAN NIEL, K. P., GILES-CORTI, B. & KNUIMAN, M. (2008), Accessibility and connectivity in physical activity studies: the impact of missing pedestrian data. Preventive medicine, 46 (1), 41-45.

DILL, J. (2004, January), Measuring network connectivity for bicycling and walking. In 83rd Annual Meeting of the Transportation Research Board, Washington, DC (pp. 11-15).

ELWOOD, S. (2008), Volunteered geographic information: future research directions motivated by critical, participatory, and feminist GIS. GeoJournal, 72 (3-4), 173-183.

FRANK, L. D., SALLIS, J. F., CONWAY, T. L., CHAPMAN, J. E., SAELENS, B. E. & BACHMAN, W. (2006), Many pathways from land use to health: associations between neighborhood walkability and active transportation, body mass index, and air quality. Journal of the American planning Association, 72 (1), 75-87.

HAKLAY, M. (2010), How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets. Environment and planning B: Planning and design, 37 (4), 682-703.

HAKLAY, M. & WEBER, P. (2008), Openstreetmap: User-generated street maps. IEEE Pervasive Computing, 7 (4), 12-18.

LARCO, N. & PARKER, R. (2013), Overlooked Destinations: Suburban Nodes, Centers, and Trips to Strips. OTREC-RR-12-12. Portland, OR: Transportation Research and Education Center (TREC). https://dx.doi.org/10.15760/trec.43.

LESLIE, E., COFFEE, N., FRANK, L., OWEN, N., BAUMAN, A. & HUGO, G. (2007), Walkability of local communities: using geographic information systems to objectively assess relevant environmental attributes. Health & place, 13 (1), 111-122.

MANAUGH, K. & EL-GENEIDY, A. (2011), Validating walkability indices: How do different households respond to the walkability of their neighborhood? Transportation research part D: transport and environment, 16 (4), 309-315.

SEEGER, C. J. (2008), The role of facilitated volunteered geographic information in the landscape planning and site design process. GeoJournal, 72 (3-4), 199-213.

SEEGER, C., LILLEHOJ, C., WILSON, S. & JENSEN, A. (2014), Facilitated-VGI, smartphones and geodesign: Building a coalition while mapping community infrastructure. Digital landscape architecture, 300-308.

STANGL, P. & GUINN, J. M. (2011), Neighborhood design, connectivity assessment and obstruction. Urban Design International, 16 (4), 285-296.

WALK SCORE (2011), Walk Score Methodology. http://www2.walkscore.com/pdf/WalkScoreMethodology.pdf.

WEINSTEIN AGRAWAL, A., SCHLOSSBERG, M. & IRVIN, K. (2008), How far, by which route and why? A spatial analysis of pedestrian preference. J. of urban design, 13 (1), 81-98.