Human Footpaths in The Outer Suburbs Of Ottawa: Distribution, Network Connectivity, and Walkability

Karine Saboui

A thesis submitted to the Faculty of Graduate and Postdoctoral Studies in partial
fulfillment of the requirements
for the degree of Masters of Arts in Geography

Department of Geography

Faculty of Arts
University of Ottawa

Abstract

This research has three objectives; 1) describe the distribution of footpaths in the outer suburbs of Ottawa, 2) quantify the impact footpaths have on network connectivity in the outer western suburban neighborhoods of Ottawa, 3) quantify the impact of footpaths on destination-based walkability measures in the outer western suburban neighborhoods of Ottawa. The distribution of footpaths is assessed using a principal component analysis on 86 observations (footpaths) and 11 variables (land usage, transit connection, income, population density). Network connectivity is measured using the link-node ratio, the gamma index, and the alpha index, as well a node betweenness centrality. Walkability is measured in ArcGIS through an origin-destination cost matrix. The results show that the distribution of footpaths cannot be explained by the selected variables. Footpaths slightly decrease overall network connectivity and re-work node betweenness centrality. Footpaths have no impact on destination-based walkability. And so, footpaths may serve as better pedestrian routes but not necessarily as faster routes through the outer western suburbs of Ottawa.

Résumé

Cette recherche comporte trois objectifs : 1) décrire la distribution des sentiers piétonniers dans les banlieues d'Ottawa, 2) quantifier l'impact qu'ont les sentiers sur la connectivité du réseau piétonnier dans les banlieues externes ouest d'Ottawa, 3) quantifier l'impact qu'ont les sentiers sur la marchabilité, telle que calculée par un indice de moindre distance entre destinations et origines. La distribution des sentiers est évaluée par une analyse des composantes principales sur 86 observations (sentiers) et 11 variables (utilisation des terres, connexion au système de transit, le revenu, la densité de population). La connectivité du réseau est mesurée en utilisant le ratio lien- nœuds, l'indice gamma, et l'indice alpha, ainsi que centralité « betweenness » des nœuds. La marchabilité est calculée par le logiciel ArcGIS selon une matrice de coût, en terme de distance, entre origines et destinations. Les résultats montrent que la distribution de sentiers ne peut pas être expliquée par les variables sélectionnées. Les sentiers diminuent légèrement la connectivité globale du réseau et changent la centralité « betweenness » des nœuds. Les sentiers n'ont aucun impact sur marchabilité, telle que mesurée par une matrice des chemins les plus cours entre origines et destinations. Et donc, les sentiers ne servent pas de chemins plus courts entre origines et destinations mais plutôt de chemins plus agréables pour le piéton.

Acknowledgments

This research project would not have been possible without the support and guidance of my thesis supervisor Dr. Michael Sawada. I am sincerely thankful for your time, your patience, your ideas, and your invaluable insights.

I wish to thank my committee members, Dr. Jackie Dawson and Dr. Elizabeth Kristjansson, for their time, knowledge, and feedback.

I would like to express my gratitude to Sara Simpkin and René Duplain from the University of Ottawa Geographic, Statistical and Government Information Centre. I thank you for your time and for helping me locate and access datasets.

I would like to thank the administration staff and my fellow graduate students from the Department of Geography. I warmly thank my friends and colleagues in the HEAL and LAGGISS labs for the stimulating discussions and the times we shared.

I thank my family and friends for their support throughout this journey.

Table of Contents

List of Tables

List of Figures

List of Appendices

Chapter 1
Introduction

1

Walking is one of the most accessible forms of physical activity. The benefits of walking on health are numerous and so there is a growing interest in understanding the elements of walkability to foster an environment that provides opportunities for walking in urban environments. There is now a growing body of literature on urban walkability, most of which focuses on features of the built environment such as, aesthetics, pedestrian network connectivity, or presence of walkable destinations. Most of the existing literature on walkability has for object the formally planned elements of a given urban area. In contrast, very few research are concerned with informal pedestrian elements, such as urban footpaths created by the footfall of pedestrians. Footpaths are almost ubiquitous to any environment –urban or not- where pedestrians are free to diverge from the formally planned walk paths. They form by the action of footfall of pedestrians over a deformable ground. Through repeated use, footpaths are maintained and evolve through time. As common as footpaths may be, almost no data exists in regards to their distributions in urban areas. Furthermore, there are very few methodologies pertaining to empirical research at larger urban scales (i.e. neighborhoods, or an urban region). As a result, no previous literature exists on the relation between urban footpaths and urban walkability.

This research, attempts to bridge the gap in the literature between footpaths and walkability as measured through connectivity analysis and destination-based indices of walkability. Furthermore, the research provides a methodology for identifying footpaths from air photos, which may be used to study footpaths at the scale of a whole urban region. Specifically, taking the outer suburbs of Ottawa as a study site, this research aimed to describe, using a principal component analysis (PCA), the distribution of footpaths in relation to land usage variables, density of public transit stops, population density and average income. This research also investigated the impact of footpaths on the network connectivity in the western suburban neighborhoods of the city Ottawa. Connectivity was assessed using metrics developed for the field of geography; these were the link-node ratio, the alpha index, and the gamma index. As well betweenness centrality-a measure developed in the field of graph theory- was used to measure the impact of footpaths on connectivity at the scale of the node. And lastly, this research measured the contribution of footpaths on walkability as measured using destination-based indices of walkability.

The thesis is organized in a classic monograph into four chapters. The first chapter, serving as an introduction, presents the survey of the literature on walkability and more specifically footpaths. The research questions and objectives are also contained within this first chapter. The second chapter describes the methods for data collection and analysis. The results of the analyses are presented in the third chapter. Lastly, the fourth chapter is a discussion of the main findings, recommendations for future research, research contributions.

1.1 Literature Review

1.1.1 Defining Walkability

Walkability is a general term that describes the extent to which a given space offers opportunities for walking (Southworth, 2005). Elements such as an origin/destination, an area, and a route are essential to walkability (Moudon et al., 2006). Not all spaces offer the same opportunities for walking. The quality of the pedestrian environment, according to Southworth (2005), depends on six criteria: 1) connectivity; 2) linkages to other modes; 3) fine-grained and varied land use patterns; 4) safety; 5) quality of path; and 6) path context (visual interest, landscaping, spatial definition). Walkability is therefore assessed though dimensionless indicators created from some combination of to the six criteria listed or used in perception-based assessments.

Indices of walkability are calculated based on the criteria of a walkable built environment such as weighted indicators of walkable destinations (Moudon et al., 2006) (Frank et al., 2010) , intersection density (Frank et al 2005) (Li et al. 2005), land use entropy (Frank, et al. 2005) (Frank et al., 2009) (Frank, et al 2007) (Frank et al., 2006), and residential density (Frank et al 2005) (Lee & Moudon, 2006) (Li, et al, 2005) (Moudon et al., 2006) (Frank et al., 2009).

While walkability indices are effective in measuring the extent to which the built environment supports opportunity for walking, they do not however inform on the human perceptions of the space. Perception of walking opportunities may vary across demographics depending on the way the act of walking is envisioned. For example, within a same neighborhood, perception of walkability opportunities may indeed be linked with the amount of mobility choices within a household (Larsen, El-Geneidy, & Yasmin, 2010) . Walkability is not only dependent on the presence of appropriate built infrastructures but also on whether the environment is a match with the walking needs and motivations of the local population. According to Alfonzo (2005) an individual’s decision to walk will depend on satisfying some needs. Drawing from Maslow's theory of human motivation and the hierarchy of needs, Alfonzo (2005) states that the most basic walking needs one will seek to satisfy is 'Feasibility' that is the individual’s capacity for mobility, and 'Accessibility’ of the environment (the presence of sidewalks, paths, trails, or features that provide perceived paths on which to walk). Once these two basic needs are met, other needs such as safety, comfort, pleasurability can be met.

Though indices of walkability in the urban environment may not fully address perception of the built environment in relation to walking needs, they are however, an effective measure of one of the fundamental walking needs, accessibility. This is because indices are calculated using indicators that solely relate to formally planned features of the pedestrian network (city sidewalks, walkways, etc.). The absence of formal planned pedestrian infrastructure should not equate the absence of walking opportunities. Pedestrians are not restricted to the planned transport networks like automobiles; rather they are free to cut across spaces if the opportunity presents itself. In so doing, pedestrians develop self-organized network of connections that work in parallel to fill in the missing links of the planned formal network.

1.1.2 Human Footpaths

The literature on self-organized pedestrian network connections in the urban environment is small. The bulk of the literature remains concentrated in the area of nature conservation. Beyond the realm of nature conservation, self-organized pedestrian structures are defined by different terminologies. This section gives an overview of the different terms encountered in the literature and summarized into Table 1. Included are both nature conservation and urban environment.

Table 1. Summary of terminology found in the literature on pedestrian made paths

Terminology	Ground type	Environment	Authors
Desire lines	Any	Any	(Kansky, 1963) (Norman, 2008) (Norman, 2011) (Myhill, 2004)
Informal Trail	Deformable	Protected 'nature' areas [Non-urban]	(Marion & Leung, 2011) (Walden-Schreiner & Leung, 2013) (Newsome & Davies ,2009).
Social Paths	Deformable	Urban	(Gallagher, Marshall, & Atkinson-Palombo, 2013)
Trail	Deformable	Urban	(Helbing, Molnár, & Schweitzer, 1998) (Helbing, Molnár, Farkas, & Bolay, 2001) (Goldstone & Roberts, 2006)

1.1.2.1 Desire lines

The term "desire line" in the context of transport route is used to mean the best connection between an origin and a destination. The term is not expensively used in the literature and may be narrowed down to three researchers: Kanksy (1963), Myhill (2004), and Norman (2008) (2011).

Kansky (1963) employs a graph theory framework to model the structure of connectivity of transport network. It is in this context that Kansky (1963) defines a desire line as "the shortest imaginary between two given vertices (air distance)" (p.31). While Kansky (1963)'s definition of desire line is purely mathematical- it is no more than Euclidian distance- Myhill (2004) and Norman (2008) (2011) add to it a social dimension.

Myhill and Norman are design researchers and therefore use the observation of desire line in the pedestrian environment as a manifestation of discrepancy between user and planners/designers.

Their approach is a descriptive one that aims to find an example of where designer does not meet the users desire/needs beyond the pedestrian experience. Myhill (2004) begins by defining desire lines as being "the ultimate unbiased expression of natural human purpose and refer to tracks worn across grassy spaces, where people naturally walk – regardless of formal pathways."(p. 293) . Norman (2008) (2011) supports this definition and further argues that desire lines are a type of "social signifier" as they are act as signals that are to be interpreted as a way to understand the social usage of a space. Norman (2008) (2011) much like Myhill (2004) proposes that desire lines are ubiquitous to the living experience and therefore are present in not only real work but also virtual world interactions. And so desire lines are formed on any type of surfaces not solely grassy patches.

1.1.2.2 Informal Trail

The term 'informal trail' appears almost exclusively in the literature on walking in 'protected natural' environment, such as provincial, state, or national parks. The use of the word 'informal' is used to mark a difference between the trails defined by a land management plan (the formal network of trails) (Walden-Schreiner & Leung, 2013) (Wimpey & Marion, 2011) (Newsome & Davies 2009).

Wimpey & Marion (2011) used a hot spot statistical analysis to better understand, amongst other things their, distribution in a Great Falls Park (Virginia) found that informal trails develop where users feel that the formal trail system does not reach their desired destination. Furthermore Wimpey & Marion (2011) support that trails acts as 'releasor cues' that draw more visitors off the formal. A similar approach used by Newsome & Davis (2009) highlights that users are motivated to reach further destination for personal reason such as question for solitude or challenge. In this sense much like desire lines, informal trails can be used as indicator in park management plans to mitigate user impact. This is the approach used by Walden-Schreiner & Leung (2013) used an empirical approach to develop a methodology for assessing park usage impact based on data collected on trail usage in Yosemite National Park (California).

1.1.2.3 (Human) Trail and Social Paths

Specific to the urban environment are the terminologies 'social path' (Gallagher et al., 2013) and (human) 'trail' (Helbing et al., 2001) (Helbing et al., 1998) .The former is not as extensively present in the literature as the latter. The specificity of the terms social path and trail are that not only are they particular to the urban but also that they are associated solely with deformable grounds such as green spaces, parks, grassy patches.

Helbing (1998) (2001)'s research seeks to understand the question of how are pedestrian trail systems built and how do they evolve over time. Hebling's uses concepts of fluid dynamics and real-world observation of trail patterns to model the pedestrian movement through space. In such, Helbing (1998) (2001) places a great emphasis on the opportunity for trail development such as the presence of open space and soft ground. Goldstone and Robert (2006) place the similar emphasis spatial opportunity as a determinant for trail development in their modeling research on trail systems.

The term social path is used by Gallagher, Marshall, Atkinson-Palombo (2013) to describe the 'informal routes that emerge in grassy areas because of footfall'. Gallagher, Marshall, Atkinson-Palombo (2013) measured pedestrian connectivity measures around light rail station in Dallas (Texas) and Denver (Colorado) including social path to argue their importance in pedestrian accessibility measures for multimodal urban transport.

1.1.2.4 Beyond terminology: Human footpaths as indicator of walking choices

An individual's behavior may be agent-based to a certain extent in the sense that direct routes with the least amount of directional changes are favored (Hill 1982) (Helbing et al. 1997) and that walking gains should offset costs (Hoogendoorn & Bovy, 2004) .However, the strategies for choosing walking routes are derived from a largely subconscious process on the part of the walker (Hill 1982). This is because humans are social beings and subconsciously will take social cues from their environment to make choices. A self-organized human trail is one such indicator.

The trail forms a mark on the landscape. This mark is a way by which walkers communicate indirectly with each other via the markings - footprints- they make on the land that indicates the way to walk. Reading walking cues allows the pedestrian to save on the cost of clearing a new trail (Helbing et al. 1997) (Goldstone & Roberts 2006) even if it leads to divergence from the most direct route (Helbing et al. 1997) (Goldstone & Roberts 2006). Taking an existing trail also reduces cognitive costs as the users figure that the path will lead to a desired destination because others have walked it thus placing faith in societal choices (Goldstone & Roberts 2006). In such, the trail by its material properties works as a greater social attractant for pedestrian than an organized municipal walkway.

1.1.3 Summary

The literature on footpaths is divided into four distinct categories based on terminology. This terminology seems to be associated with the environment where the network forms, for example, urban (trails, social paths) or non-urban (informal trails), or universal (desire lines). The method used for this literature review has been to query database of peer-reviewed articles/book chapters using specific keywords. Keywords used were “trails”, “shortcuts”, “social trails”, “desire lines”, “desire paths” “informal paths” “informal trails”. Each of these terms were then combined with the terms “walkability measure” OR “walkability index” OR Walkability score” OR “Walk Score”. No relevant results were returned when the terms were combined. Research on measures of point-based walkability and self organized pedestrian walking routes. This is an important lacuna, as this type of walking infrastructure is common, contributes to the network and is developed by users to reach a given destination more efficiently. The ability to efficiently reach a destination is a core component of all calculation of walkability indices and therefore should take into consideration all route options available.

1.2 Study Site and Research Objectives

1.3 Study Site

For this research, the selected study sites are the outer suburbs of the city of Ottawa, which are located beyond the greenbelt (Figure1). Within the outer suburbs are ‘natural neighbourhoods’ that are defined based on socio-economic status, demographics and physical barriers (Parenteau et al., 2008) . For the scope of this research, the neighbourhoods forming the outer suburbs are grouped into three distinct regions a western region, a southern region, and an eastern region based on the physical separation created by the greenbelt.

The rational for selecting the outer suburbs for the study site is that their street pattern is similar (curvilinear), and their primary land usage is residential. As well, most of these neighbourhoods are classified solely as auto suburb (Gordon & Janzen, 2013).

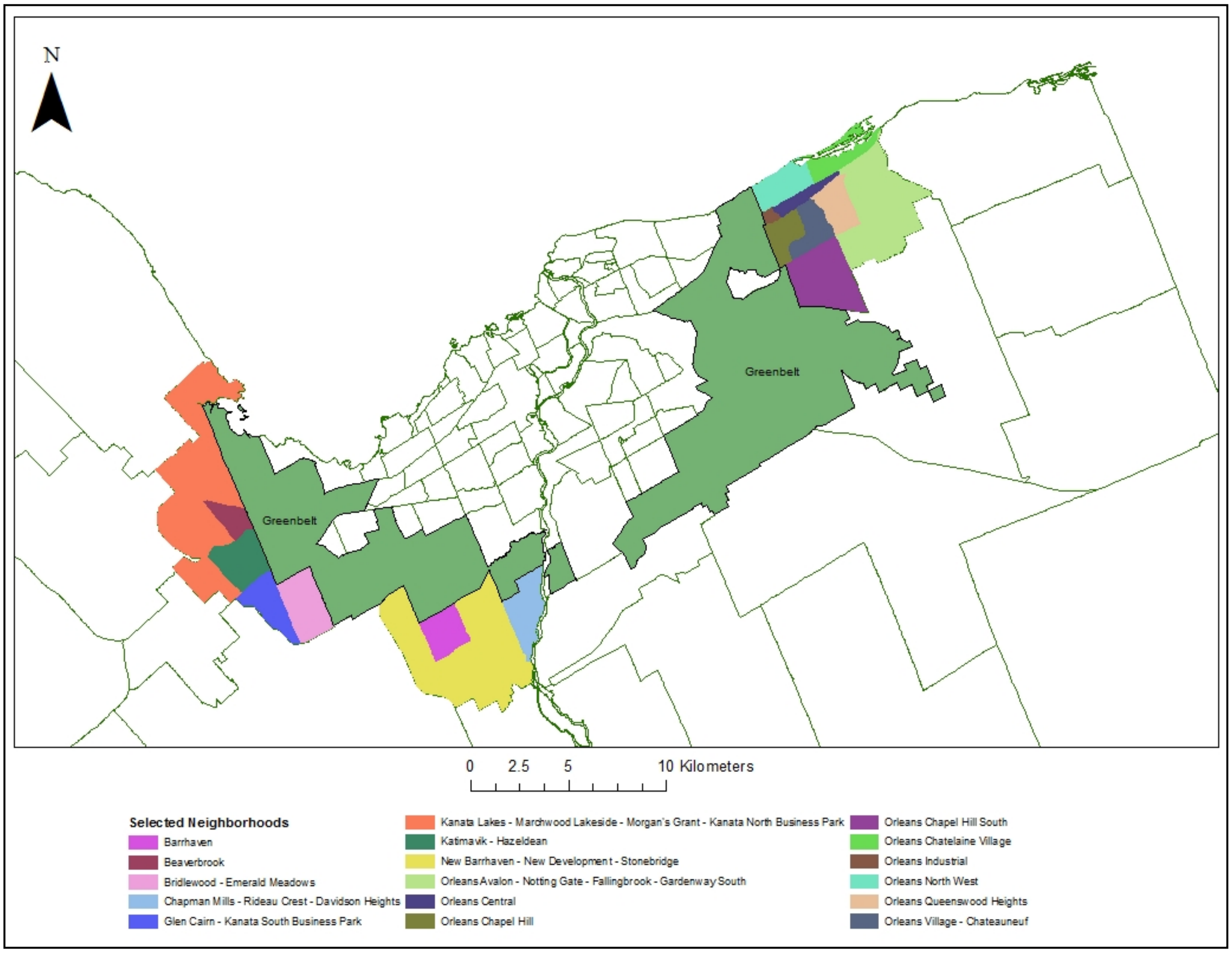

Figure 1. Maps of the outer suburban neighborhoods of Ottawa (data source: Ottawa Neighborhood Study)

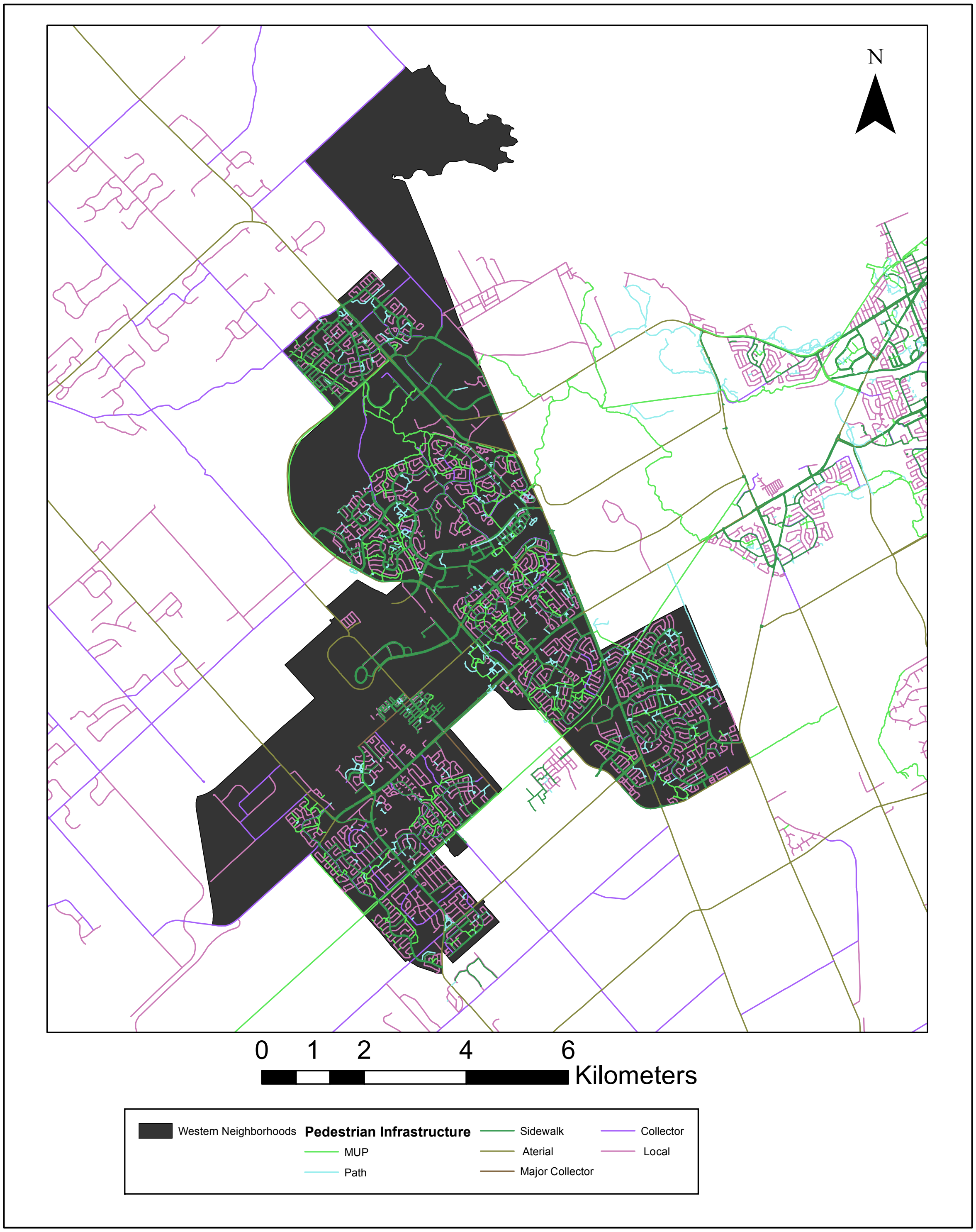

Figure 2. Map of the western suburban neighborhoods of Ottawa showing pedestrian infrastructure (data source: City of Ottawa Pedestrian Network, City of Ottawa Road Network)

1.4 Research Objectives

The role of human footpaths in the context of suburban walkability has not been extensively researched. Most of the existing research focuses on the role of human footpaths as linkages between destinations. None however, relate footpaths to point-based walkability indices. This research seeks to understand the spatial emplacement of footpaths within the city of Ottawa in order to quantify their impact on the existing pedestrian network not only in terms of network connectivity but also on walkability indices.

Based on the existing literature on human footpaths, this research had the following objective: to quantify the impact of human footpaths on walkability especially in relation to network connectivity and point-based measures of walkability. As well, this research aimed to describe the association between footpaths distribution and certain independent variables of the environment.

Specifically this research sought to:

1. Investigate whether a relationship exists between footpaths and allocated land usage, as well as income and population density variables. Deformable grounds- green space- has been identified throughout the literature (Barros , Gonnet , Pickering 2013) (Walden-Schreiner & Leung 2013) (Wimpe & Marion. 2011) (Leung et al. 2011) (Newsome & Davies, 2009) (Gallagher, Marshall, Atkinson-Palombo 2013) (Helbing et al. 2001) (Helbing et al. 1997) providing the proper context for trail development.

2. Quantify how footpaths change the connectivity of the network and node centrality within the western suburban neighbourhoods (see Figure 2).

3. Quantify how measures of point-based walkability change with the addition of footpaths within the western suburban neighbourhoods (see Figure 2).

Chapter 2
Methodology

2

This chapter presents the methods for data collection and data analysis. The chapter is divided into two main sections: 1) Methodology and 2) Data Analysis. In the first section, the data sources and data collection process is explained. The variables are presented and their relation to walkability is justified. This section also explains how the pedestrian network was created from the various data sources. The second section, explains the data analysis process. This second section presents a detailed explanation of the principal component analysis (PCA) and the mathematical basis for the connectivity measures and walkability index.

2.1 Methodology

2.1.1 Data Collection

The data used in this research was obtained from the National Capital Commission (via the University of Ottawa), the City of Ottawa (via the Open Data Portal and the University of Ottawa), Statistics Canada, the Ottawa Neighborhood Study, and DMTI Spatial (via the University of Ottawa). A summary of the data sources and relevance to the research objectives is presented in Table 2.

Table 2. Summary table of the data type and data sources used in this research

<table>
<tr><td colspan="3">Footpaths identification
Air photos (source: NCC and City of Ottawa 2011)</td></tr>
<tr><td>PCA</td><td>Connectivity</td><td>Walkability Measure</td></tr>
<tr><td>Land Zoning (City of Ottawa 2011)
OC Transpo (City of Ottawa 2011)
Census (Statistics Canada 2011)</td><td colspan="2">Pedestrian Network (City of Ottawa 2014)
Road Network (City of Ottawa 2014)
EPOI (DMTI 2011)</td></tr>
</table>

2.1.2 Footpath Identification

Footpaths were identified from the air photo and stored in an ArcGIS shapefile (feature dataset). The specific steps involved heads-up digitalizing of the footpaths from the air photo. This method has been used by Gallagher, Marshall, and Atkinson-Palombo (2013) because of its cost and time efficiency. For quality control purposes, a few rounds of digitalizing were performed. This was necessary because identification of footpath is largely dependent the ability to properly distinguish features on the images; and this ability is learned with practice.

2.1.3 Footpaths and variables

The variables were grouped in the following categories: those pertaining to land usage, those pertaining the socio-economic, and those pertaining to connection to public transit. These variables were identified in the literature as having an influence on walking (Gallagher et al., 2013) (Larsen et al., 2010). A 400-meter circular buffer was created around each footpath. The variables were sampled from within the buffer areas around each footpath.

Land usage variables

The land usage variables were obtained from the City of Ottawa. The City of Ottawa defined 8 classes of zoning for land usage. These land usage are: residential, institutional, open and leisure spaces, environmental, commercial, industrial, rural, and development reserve. These variables were selected because footpaths form most often in area with available deformable ground and this type of ground is most often found in open and leisure spaces. However, other land usage cannot be excluded as they also may provide the right context for footpath formation. The data was imported as a shapefile into ArcMap. The land use dataset was intersected to the buffer dataset. For each buffer, the area covered by each land usage was summed and then expressed as a percent of the total buffer area.

Socio-Economic Variables

The socio-economic variables were extracted from the Canadian Census (2011) data at the scale of the dissemination area. The variables that were selected were the average income household at the scale of the dissemination area and population density. Average income was selected because in household with more income, automobile mode of transport is more accessible (Larsen et al., 2010) while lower income household have fewer choices and may rely on public transit or walking as a mode of transportation. Population density was selected as a variable because footpaths develop through repeated usage, with denser population there may be more pedestrian activity and more use of footpaths. The socio-economic dataset was intersected with the buffer shapefile. The population was expressed as density per area. The income was averaged for each buffer.

Transit Stops

OC Transpo stops were selected as a variable because it has been shown that in the suburb informal paths sometimes develop next to public transit stations (Gallagher et al., 2013). This is because, individuals often walk to access other modes of transports such as public transit. The transit stop data was obtained from the City of Ottawa. The transit dataset was intersected with the buffer shapefile. The transit stop were expressed as density per area.

2.1.4 Creating the Pedestrian Network

The pedestrian dataset provided by the City of Ottawa is composed of three elements: sidewalks, multi-use pathways (MUP), and paths. The City pedestrian dataset represents only those elements that are built specifically for pedestrians. As a result sidewalks are discontinuous polylines and local streets, which do require sidewalks, but are still available to pedestrian, are omitted. Using the City of Ottawa pedestrian network would have been time costly, as sidewalks would have to be repaired into continuous lines for analysis. To overcome the issue of discontinuity and missing local streets, the City of Ottawa roadway guidelines for sidewalks was used to identify which street types were available to pedestrians, either through sidewalks or low-flow traffic. As a result, all street subclass except “Highway”, “Ramp” , “Transitway”, “BusOnly” were extracted. MUPs and paths were extracted from the City pedestrian dataset. The extracted elements from both the road and City pedestrian datasets were merged to create a new pedestrian dataset. This new pedestrian dataset was compared to the air photos. This step was necessary as the air photos were from 2011 and the pedestrian/road datasets from 2014. The portions of the dataset that were not present on the air photos were deleted.

2.1.5 Building the Pedestrian Networks

For each region, two networks were built using ArcGIS. One network was built from formally planned elements (local streets, streets with sidewalks, multiuse pathways, paths). The second network was built from both formally planned elements and the identified footpaths.

2.1.5.1 Topology Rules

A set of topology rules was defined for both networks. Topology rules were necessary as the elements of each network were derived from at least two different datasets and therefore introduced dangles, pseudo-nodes, overlaps, and multipart features. The topology rules enforced are presented in Table 3.

Table 3. Topology Rules Applied to the Network

Rule
Must not overlap
Must not self-overlap
Must not have pseudo-nodes
Must not have dangles[1]
Must be single part

2.1.6 Nodes, Origins and Destinations

The alpha index, the gamma index, and the link-node ratios were calculated using the number of network links and the number of network nodes. The nodes represented the intersection of the links with one another or cul-de-sac end-nodes. Betweenness centrality was calculated for nodes representing walkable destinations (EPOI and Transit Stops).

2.1.6.1 Destination-Based Walkability Analysis

The destination-based walkability analysis was calculated from and origin-destination matrix. The origins nodes represented the network nodes and the destinations nodes represented walkable destinations (EPOI and Transit Stops) along the network.

---

[1] Dangles were permitted for end nodes such as cul-de-sacs.

2.2 Data Analysis

2.2.1 Principal Component Analysis

Principal Component Analysis (PCA) is a descriptive data analysis method. PCA allows to identify groups of inter-correlated variables, reduce the number of variables being studied, and to rewrite the dataset in an alternative form (Johnston, 1980) . In this research PCA is used to identify groups of inter-correlated variables that could describe the distribution of footpaths in the suburban neighborhoods of Ottawa. The selected variables (average income, population density, transit stop density, land usage allocation) are sampled in each of the footpath buffer. The data matrix is stored as a comma separated value file and imported into R. The analysis is performed using the *prncomp* function in R.

The dataset was scaled using the SCALE function. The Principal Components were obtained using the function prcomp. The loadings were obtained from the function DATA$loadings. The sampling adequacy was obtained using the KMO function of the psych package.

2.2.2 Network Connectivity and Node Centrality

2.2.2.1 Network Connectivity Measures

Network connectivity was analyzed using three measures: the link-node ratio, the gamma index, and the alpha index. The alpha and gamma indices have been developed for geography and assume that most connected network are loops or circuit type networks (Haggett, 1969). Circuit type network are a type of planar network. An optimal circuit network is one where there are few to no unconnected end nodes. For this research network connectivity was measured around certain footpaths. Buffers of twice the length of each footpath were created in ArcMap. Overlapping buffers were dissolved to form a continuous area. The networks plus and minus footpaths were intersected to the buffers. Standalone features, such as unattached nodes or edges, were deleted manually using the editing function. As well, buffers that returned indices values above 1 or below 0 were removed. Such values were returned for very small buffer areas where the network was formed by just 1 or 2 links and 1 node.

2.2.2.1.1 Link-Node Ratio

The link-node ratio is a very simple measure of network connectivity. It is the ratio of links to nodes present in a network. The higher the ratio the more connected the network. A ratio of 2.5 represents a fully connected grid network (Tal & Handy, 2012). Ratios of 1.4 to 1.6 are the recommended standard found in most urban planning guidelines (Tal & Handy, 2012).

2.2.2.1.2 Alpha index

The alpha index is a measure of the number of circuits, or cycles, within a graph in comparison with the maximum possible number of circuits, cycles, possible for that graph. The alpha index values range from 0 to 1, where 1 represents a fully connected graph.
$\alpha = E - N + 1 / 2N - 5$

Where E represents the edges and N the nodes.

2.2.2.1.3 Gamma Index

“The gamma index is a simple measure of the density of the network” (Barthélemy, 2011, pg.8) . Gamma index measures the number of observed links in comparison to the maximum possible links for a network. The values range from 0 to 1.

the gamma index of a planar graph, as is the case in this research is calculated using the following equation:

$\gamma_p = E/3N-6$

Where E represents the edges and N the nodes.

2.2.2.2 Node Centrality (Betweenness Centrality)

Betweenness centrality (BC) of a node is defined as the fraction shortest paths between pairs of other nodes in the network that pass by that node (Freeman, 1977).

And thus BC for a node (v) is defined as

$$BC\ (v) = \sum_{i} \sum_{j} \frac{g\ ij(v)}{gij}$$

Where g *ij* (v) represent the shortest paths going through *v* from *i* and *j*

g *ij* represent the shortest paths from *i* and *j*

For this research BC for each node (representing walkable destinations) is calculated for the networks with, and without footpaths using the UNA toolbox[2]. The BC values are normalized so that the range of possible values is between 0 and 1. The change in betweenness centrality value at each node when footpaths are added to the network was calculated by subtracting the node betweenness values of the minus footpath network from the network with footpaths.

The UNA toolbox calculates BC at a specified search radius around each node of on a network. The network was formed of 693 nodes representing EPOI and OC Transpo transit stops, and edges defined as the walkable pedestrian elements (sidewalks, multi-use pathways, paths). BC was calculated at four search radii for both networks with and without footpaths. In order to determine an appropriate starting search radius, the distance between each node was calculated using the ArcMap proximity tools. The maximum distance between nodes was found to be 610.40 meters. A search radius of 615 meters was used as the starting radius. Setting the search radius at the maximum distance between nodes avoid null BC values resulting from a distance greater than the search radius. BC was also calculated at radii of 1000 ,1500, and 2000 meters to assess the effect of scale on BC variation. The search radii were specified as Euclidean distance

---

[2] The UNA (Urban Network Analysis) toolbox is an ArcGIS plugin that has been developed by City Form Lab

on the UNA toolbox because the ArcMap proximity tools calculate distances in terms of Euclidean distance.

2.2.3 Destination-Based Measure of Walkability

Walkability was measured using a similar method as the one created by Walk Score. The Walk Score algorithm searches for pre-defined amenities type from an origin point. Scores are generated from a polynomial decay function where scores are allocated as follows:

> "…full score or near full score for amenities that are within .25 miles of the origin. After this, scores decrease with distance smoothly. At a distance of one mile, amenities receive only about 12%of the score they would have had if they were right next to the origin. After one mile, scores decrease less quickly with greater distance, until they reach 1.5 miles, after which they do not count towards the final score" Walk Score (2015)

A fifth degree polynomial function (see equation 1), which was derived from the Walk Score algorithm, is applied to the OD-Cost matrix.

Equation 1. Fifth Degree Polynomial

$$F(x)= 1.1040\ x^5 + 8.886\ x^4 + -123.731\ x^3 + 86.314\ x^2 + -16.449\ x + 101.984$$

The destination or amenities were defined as EPOI and OC Transpo Transit stop. The origins were network nodes of the pedestrian network without footpaths.

Chapter 3
Results

3

This third chapter presents the results of the data analysis. This chapter is divided into three main sections. The first section presents the results of the principal component analysis (PCA). The second section presents the results of the connectivity analysis. This section is further divided into two subsections. The first subsection presents the connectivity results obtained by calculating the link-node ratio, the alpha index, and the gamma index. The second subsection presents the results of for the node betweenness centrality analysis. Lastly, the third section presents the results of for the walkability analysis. For all sections, the results are presented in the form of tables and figures.

3.1 Principal Component Analysis

There were 86 footpaths identified from the air photos. A PCA was performed on a data matrix of 11 variables on 86 observations. The results of the PCA indicate no clear grouping of inter-correlated variables. The 70% of the variance within the dataset is accounted for by the first 5 components (Table 4). And the first 3 components explain only 49% of the variance within the dataset (Table 4). Furthermore, the proportion of variance of the components is similar with no one component explaining the most of the variance (Table 5, Figure 3). A test for data sampling adequacy, using the Kaiser-Meyer-Olkins method, revealed that the dataset could not be explained by a few factors (Table 6). As well, the correlation between variables is very low and therefore suggests that the dataset does not lend itself well to a PCA.

3.1.1 Extracting the Components

There are two methods for extracting the components. The first is the Kaiser Criterion method. Using the Kaiser Criterion, only those components with an eigenvalues above 1 (Table 4) are retained. The second uses a scree plot; the eigenvalues are plotted against the variance of the components (Figure 3). The slope of the graph is interpreted and components to the left of the sharp change in the slope (called the “elbow”) are retained. Both methods were used in this research (Table 4 and Figure 3).

Using the Kaiser Criterion, the first 6 components should be retained as their eigenvalue is above 1 (Table 4). The scree plot (Figure 3) shows an elbow at the third component.

Table 4. Standard Deviation, Proportion of Variance, Cumulative Proportion, and Eigenvalues for the 11 Components. The principal components with an eigenvalue above 1 are represented in a black while those below are represented in grey.

	Component										
	1	2	3	4	5	6	7	8	9	10	11
Standard deviation	1.44	1.40	1.15	1.12	1.05	1.01	0.91	0.81	0.60	0.55	0.37
Proportion of Variance	0.19	0.18	0.12	0.11	0.10	0.09	0.08	0.06	0.03	0.03	0.01
Cumulative Proportion	0.19	0.37	0.49	0.60	0.70	0.79	0.87	0.93	0.96	0.99	1.00
Eigenvalues	2.09	1.95	1.32	1.26	1.10	1.01	0.83	0.65	0.36	0.31	0.14

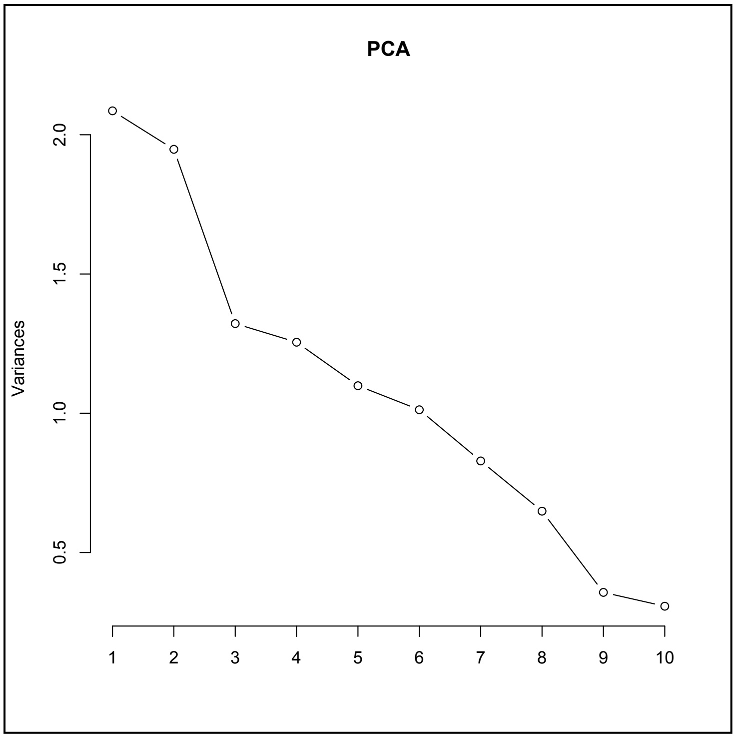

Figure 3. Scree plot of the variance against the components

3.1.2 Component Loading

3.1.2.1 Principal Component 1

The first component, which accounts for 19% of the variance, is defined by high (above 0.5) loading on the average income variable and average loading (<0.4 and >0.5) for rural land usage and connection to transit variable (negative loading) (Table 5).

3.1.2.2 Principal Component 2

The second component, which accounts for 18% of the variance, is defined by high (above 0.5) loading on of the residential land usage (negative) variable and average loading (<0.4 and >0.5) for commercial land usage variable (Table 5).

3.1.2.3 Principal Component 3

The third component, which accounts for 12% of the variance, is defined by high loading on environmental land usage and average loading (<0.4 and >0.5) for open space land usage (negative), environmental land usage, and industrial land usage (Table 5) variables.

Table 5. Variable loading (high loadings are shown in red and average loadings are shown in green) on the retained principal components .

	Principal Components		
	1	2	3
Socio-Economic			
Population Density	-0.20	0.01	-0.16
Average Income	0.52	-0.16	0.05
Land Usage			
Open Space	0.27	0.32	-0.44
Institutional	0.23	-0.33	-0.37
Industrial	-0.05	0.33	0.48
Environmental	-0.07	0.01	0.50
Development Reserve	0.29	-0.11	0.08
Commercial	-0.29	0.40	-0.37
Rural	0.40	0.34	0.03
Residential	-0.13	-0.61	-0.01
Connection to Transit			
Transit Stop Density	-0.46	-0.05	-0.14
Eigenvalues	2.09	1.95	1.32

3.1.3 Correlations between variables

Overall, few variables of the data set have significant correlations between one another. The highest correlations are between population density and open space (0.57) and Residential and open space (-0.58) (Figure 5). Low correlations between variables explain why the PCA does not show any patterns or grouping of variables in areas around footpaths.

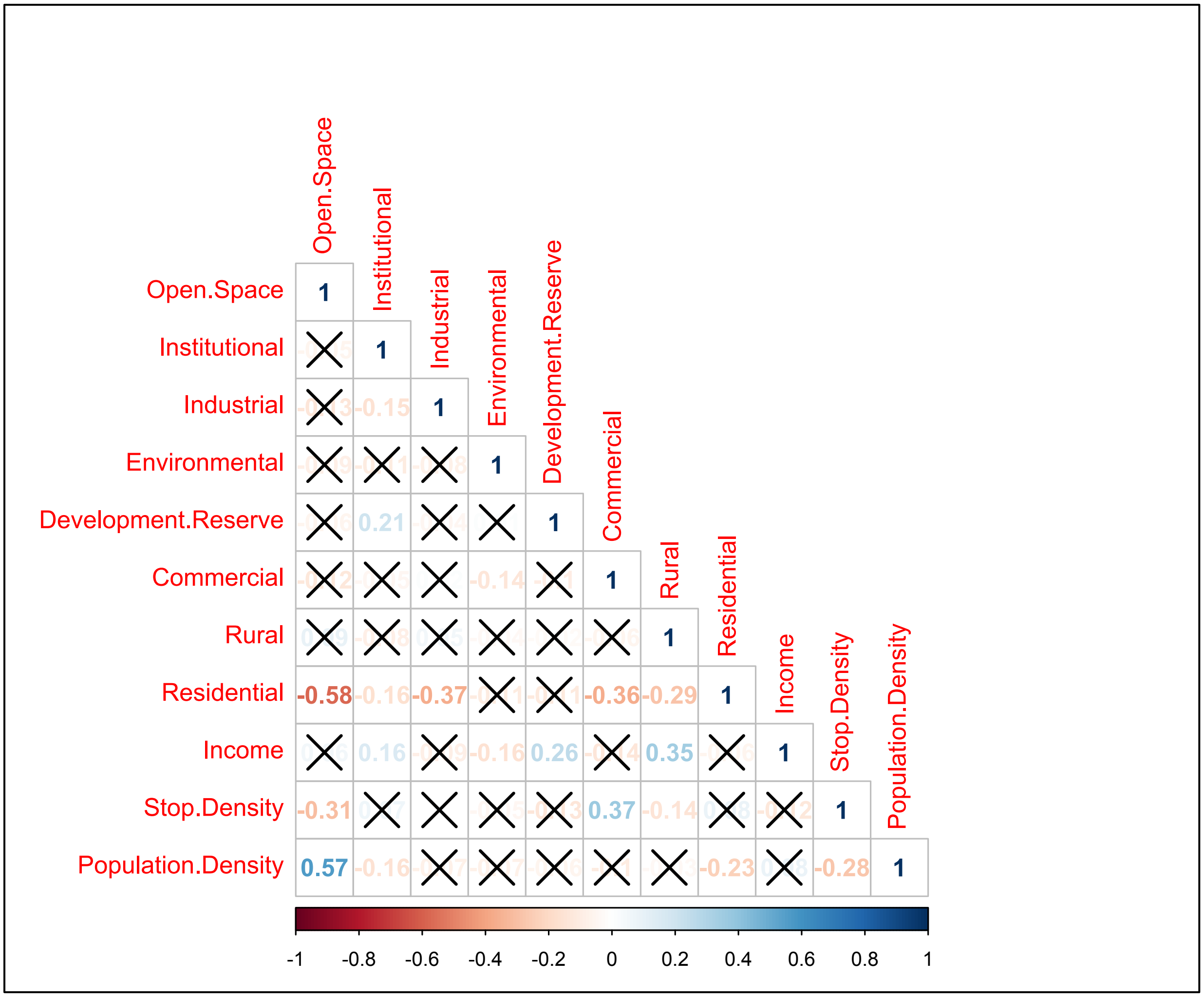

Figure 4. Correlation between variables with test for significance (p < = 0.05)

3.1.4 Testing the Validity of the Data

The Kaiser-Meyers-Olkins (KMO) test verifies the sampling adequacy for a dataset for PCA. Overall and variable values above 0.5 are acceptable. The overall MSA value is 0.33 well below the acceptable 0.5. Only variables relating to connection to public transit and average income have MSA scores above 0.5. This indicates that the lack of conclusive results in the PCA may be due to the structure of the dataset and the distribution of the variables.

Table 6. MSA Values for each of the variables and for the whole dataset

	MSA
Open Space	0.35
Institutional	0.37
Industrial	0.19
Environmental	0.14
Development Reserve	0.26
Commercial	0.32
Rural	0.37
Residential	0.29
Transit Stop Density	0.55
Income	0.52
Population Density	0.34
Overall	0.33

3.2 Connectivity Analysis

3.2.1 Indices and Ratio

3.2.1.1 Link-node ratio

The link-node ratio for the buffered area around each footpath increases from 1.133 to 1.172 with the addition of footpaths to the network. However, for the western region (combining all neighborhoods into one region) of the study site the addition of footpaths decreases the link-node ratio from 1.442 to 1.400 (Table 7).

3.2.1.2 Gamma Index

The gamma index is a ratio of the number of edges present in the network to the maximum possible number of edges possible for that same network. The possible range of value is between 0 and 1. Where 1 represents a fully connected network while 0 an unconnected one. The Gamma index for the whole western area (combining all neighborhoods into one larger zone) of the study site is very high; for the network with footpaths, the index is 0. 934 and without footpaths it is 0.961 (Table 7). For the buffered area around the footpaths, the gamma index increases from 0.756 to 0.782 with the addition of footpaths (Table 7).

3.2.1.3 Alpha Index

The gamma index is a ratio of the number circuits present in the network to the maximum possible number of circuits possible for that same network. The possible range of values is between 0 and 1. Where 1 represents a fully connected and 0 an unconnected network. The alpha index, much like the gamma index, is decreases slightly from 0.943 to 0.901 for the western suburban area (combining all neighborhoods into one region) of the study site. For the buffered area, the index is increases from 0.638 to 0. 677 when footpaths are present.

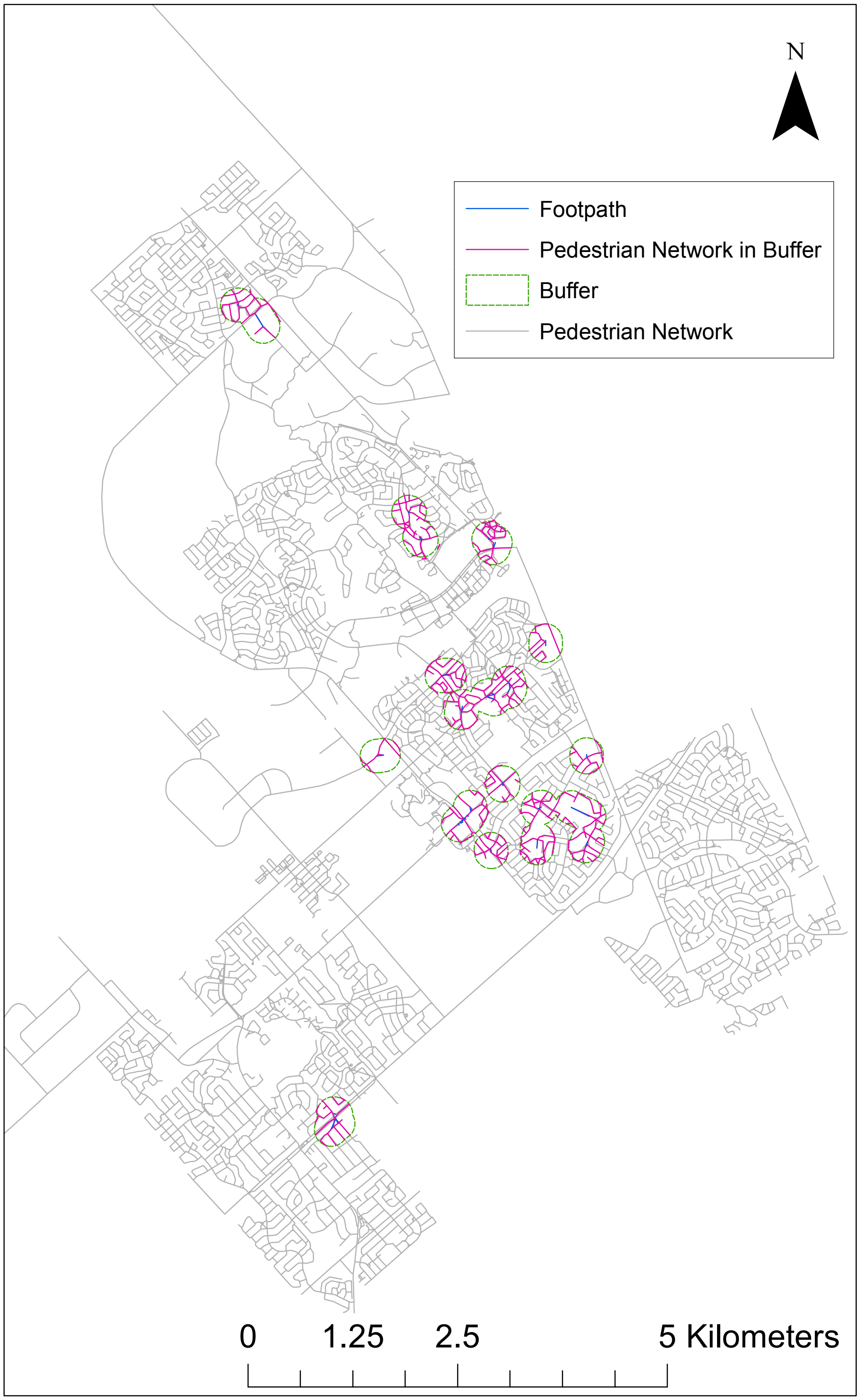

Figure 5. Delineation of Buffers around footpaths used for the connectivity analysis as measured using LNR, Gamma Index, Alpha Index.

Table 7. Comparison of connectivity measure between the networks with footpaths (plus) and the network without footpaths (minus) for the whole western region of the study area and for buffered area around the footpaths

	Connectivity Measure					
	Link Node		Gamma		Alpha	
	Plus	Minus	Plus	Minus	Plus	Minus
Buffer around footpath	1.172	1.133	0.782	0.756	0.677	0.638
Study Site	1.400	1.442	0.934	0.961	0.901	0.943

Table 8. Summary of difference between the networks with footpaths and the network without footpaths

	Connectivity Measure		
	Link-Node	Gamma	Alpha
Buffer around footpath	0.039	0.026	0.039
All	-0.042	-0.028	-0,042

3.2.2 Node Betweenness Centrality

Betweenness centrality (BC) was calculated for the nodes representing pedestrian destinations (EPOI and OC Transpo transit stops). BC was calculated using 4 distinct Euclidian search radii (615 meters, 1000 meters, 1500 meters, 2000 meters). The results are presented as figures showing the difference in BC for each node. The BC values of nodes on the network with footpaths are subtracted from the BC value for that some node on the network without footpaths. BC values are normalized so that the BC values range between 0 and 1. As the search radius around nodes increases, more nodes exhibit variations in betweenness centrality values (Table 9).

This is because as the radius is increased the number of node pairs considered in the calculations is also increased. The overall amount of nodes with a change in connectivity represents at most one fifth of the total amount of nodes (Table 9).

Table 9. Number and Percentage of Nodes with a BC value change

	Amount of node with a changed BC value	
	Count	Percentage
Radius 615 meters	18	2,61
Radius 1000 meters	49	7,10
Radius 1500 meters	102	14,78
Radius 2000 meters	138	20,00

The BC may increase or decrease when footpaths are added to the network. Comparing the results for the four radii, it is apparent that nodes whose BC are changed by the addition of footpaths are grouped within the same central area. This area grows larger as the search radius is increased (Figure 6 to Figure 16)

3.2.2.1 Radius 615 meters

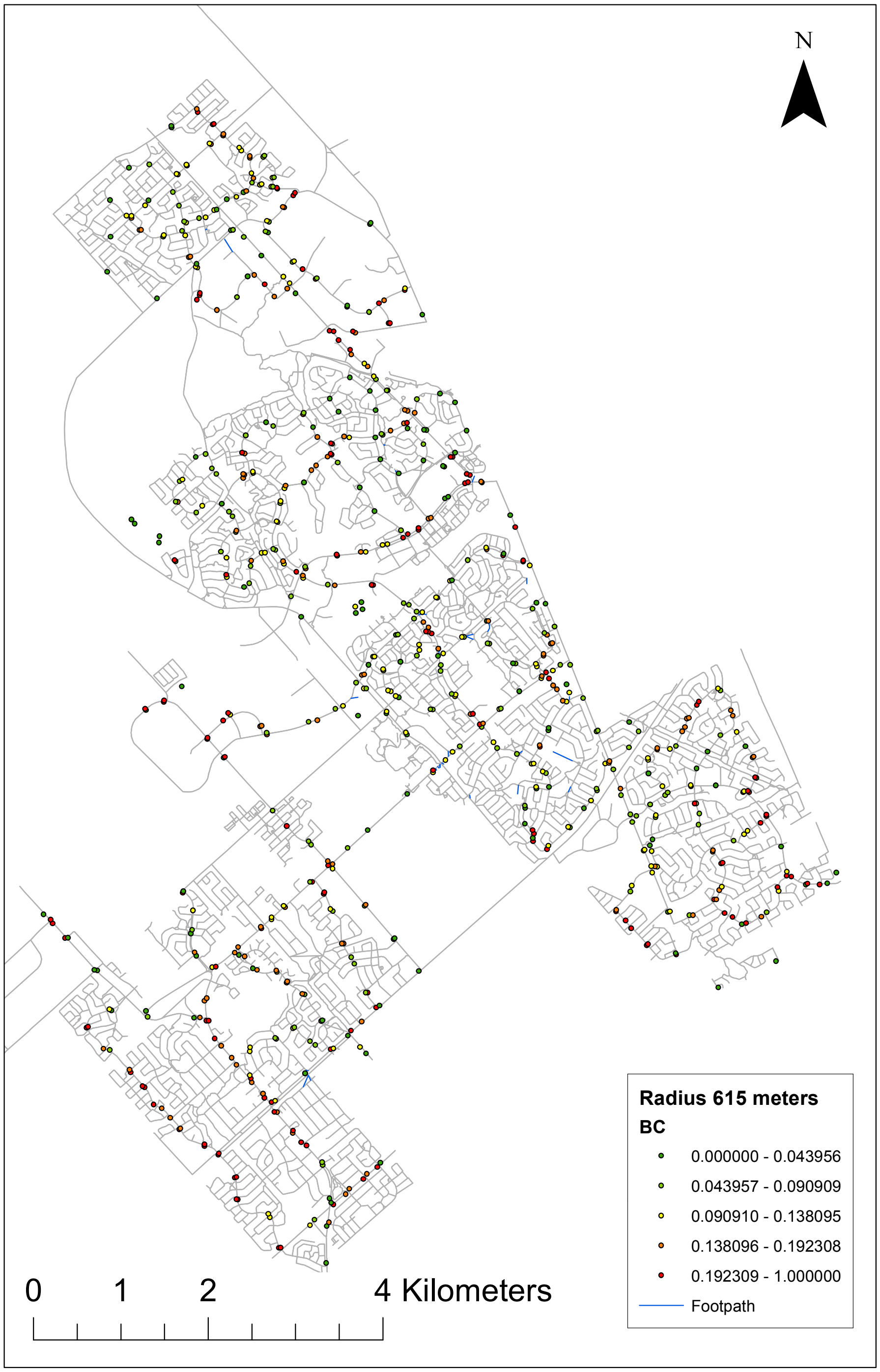

Figure 6. Betweenness Centrality for nodes on the network without footpath (search radius set at 615 meters)

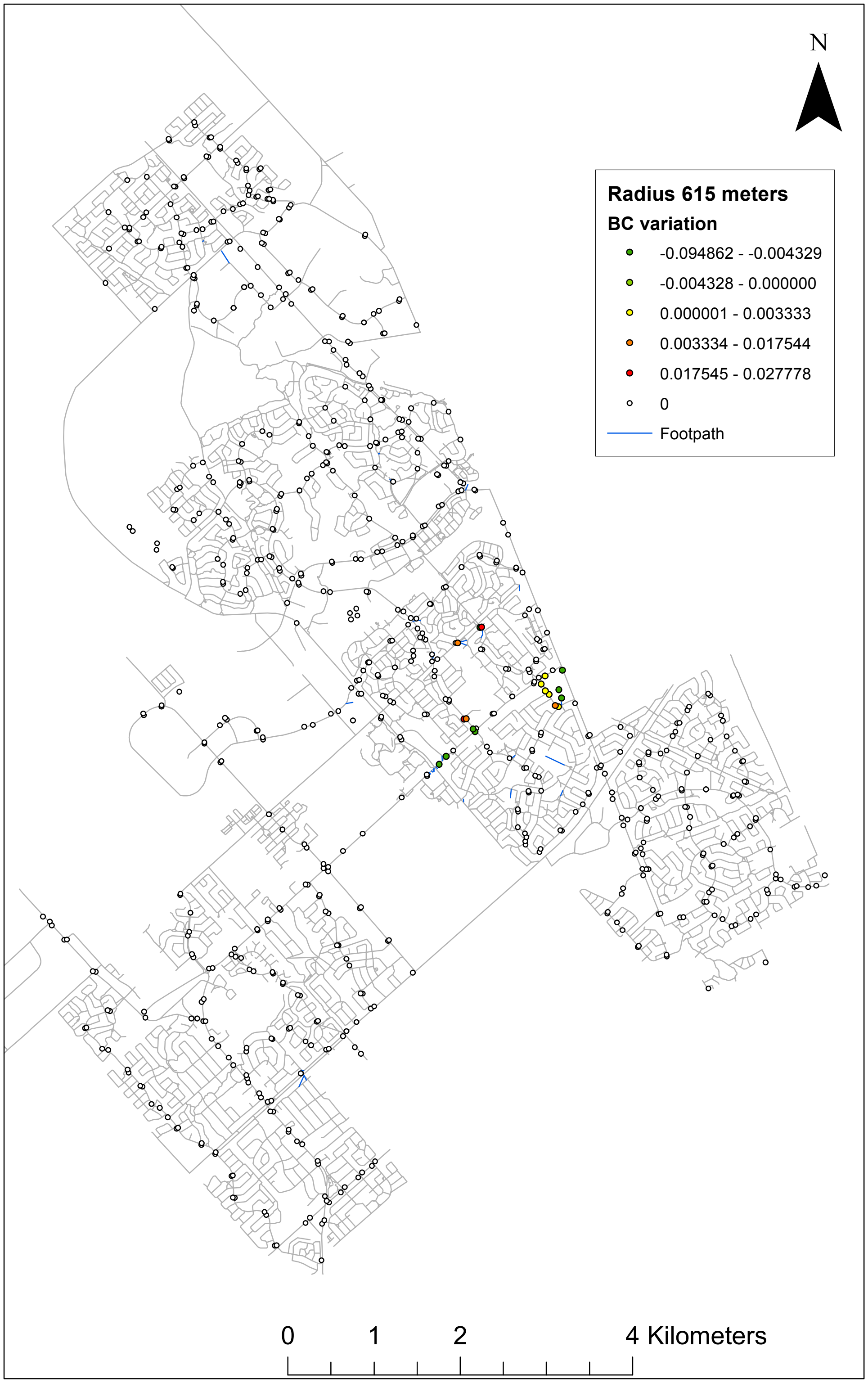

Figure 7. Variation in node BC values comparing the network with footpath to the network without footpaths using a search radius of using a search radius of 615 meters

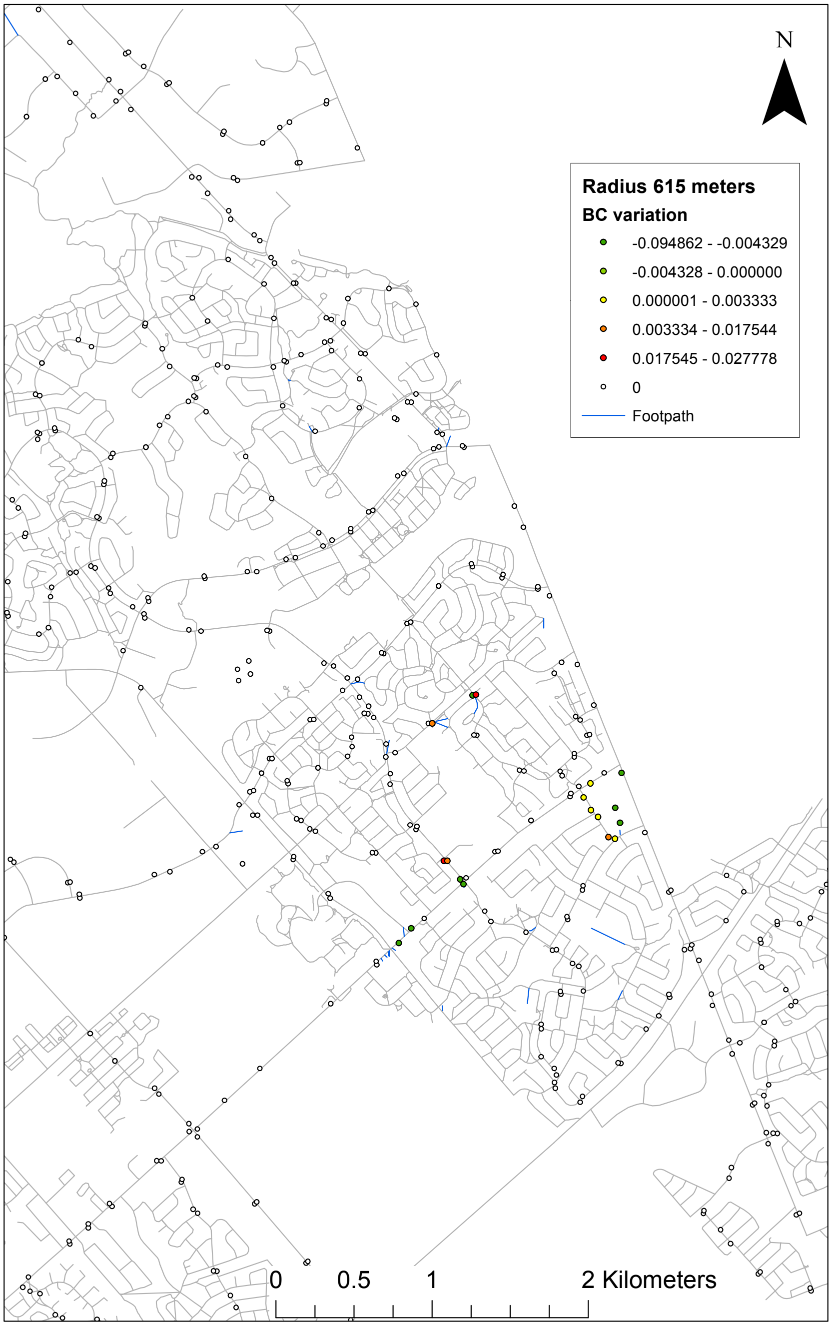

Figure 8. Close-up of area where node BC changes with the addition of footpaths to the network (search radius 615 meters)

3.2.2.2 Radius 1000 meters

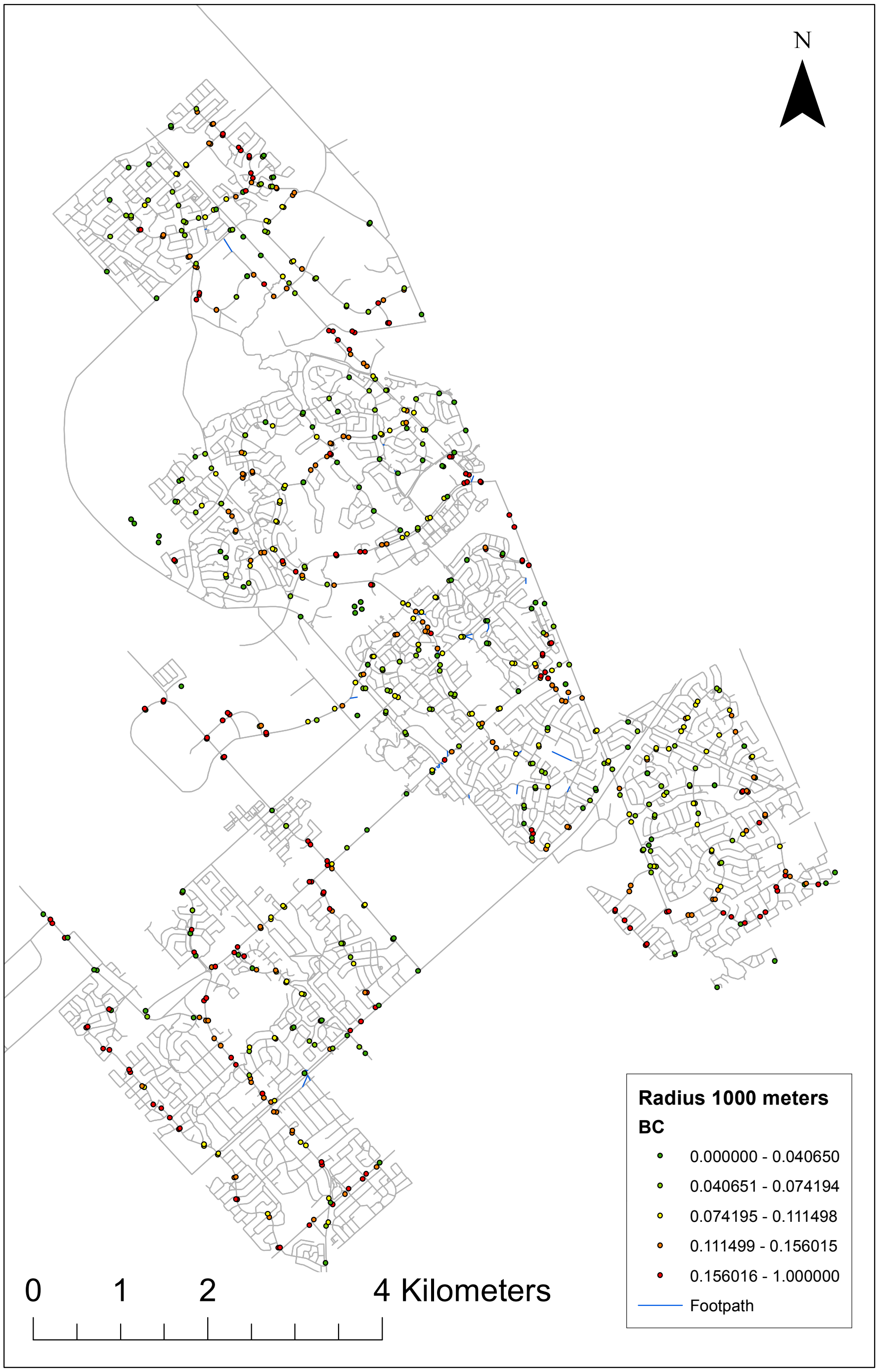

Figure 9. Betweenness Centrality for nodes on the network without footpath (search radius set at 1000 meters)

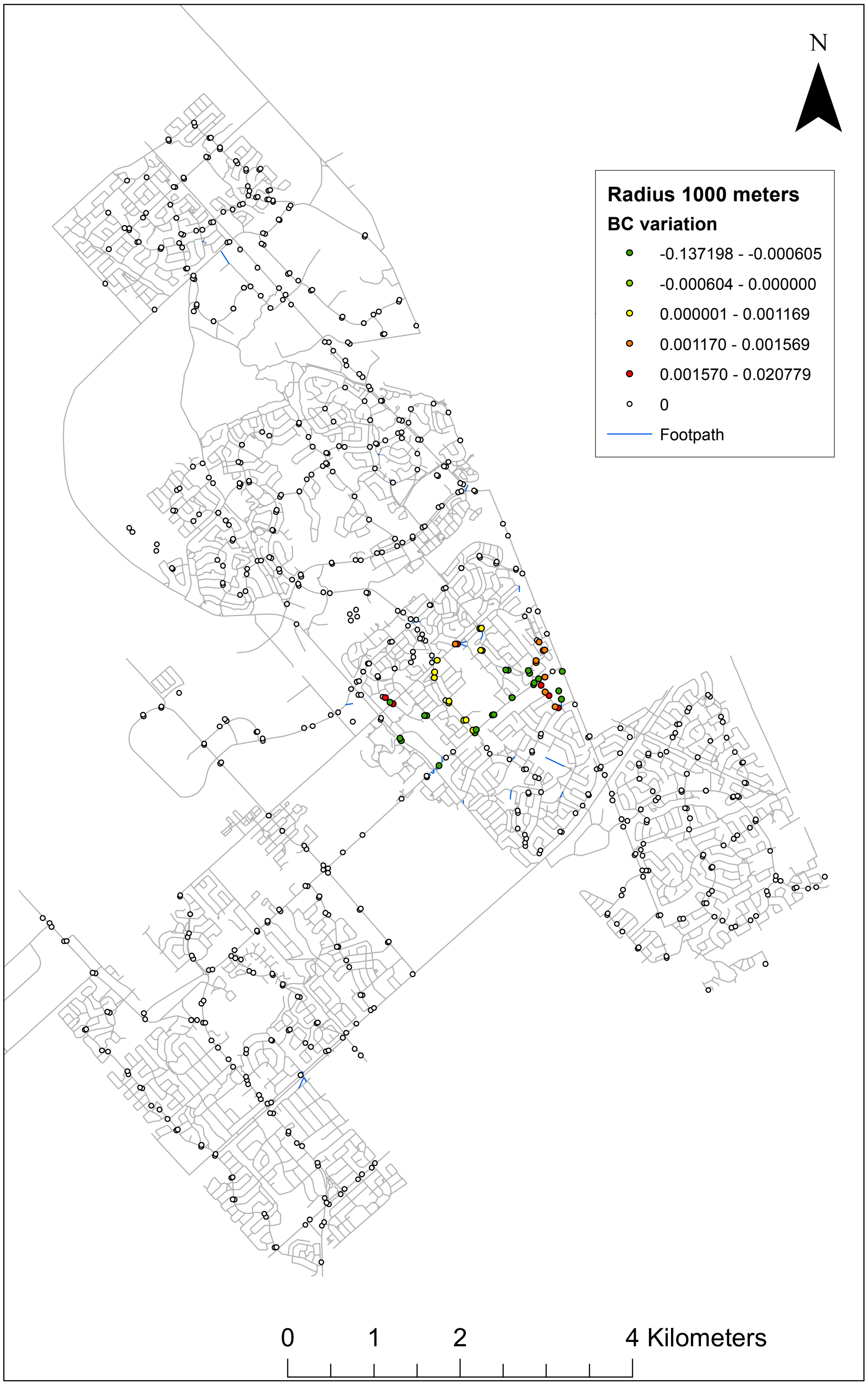

Figure 10. Variation in node BC values comparing the network with footpath to the network without footpaths using a search radius of 1000 meters

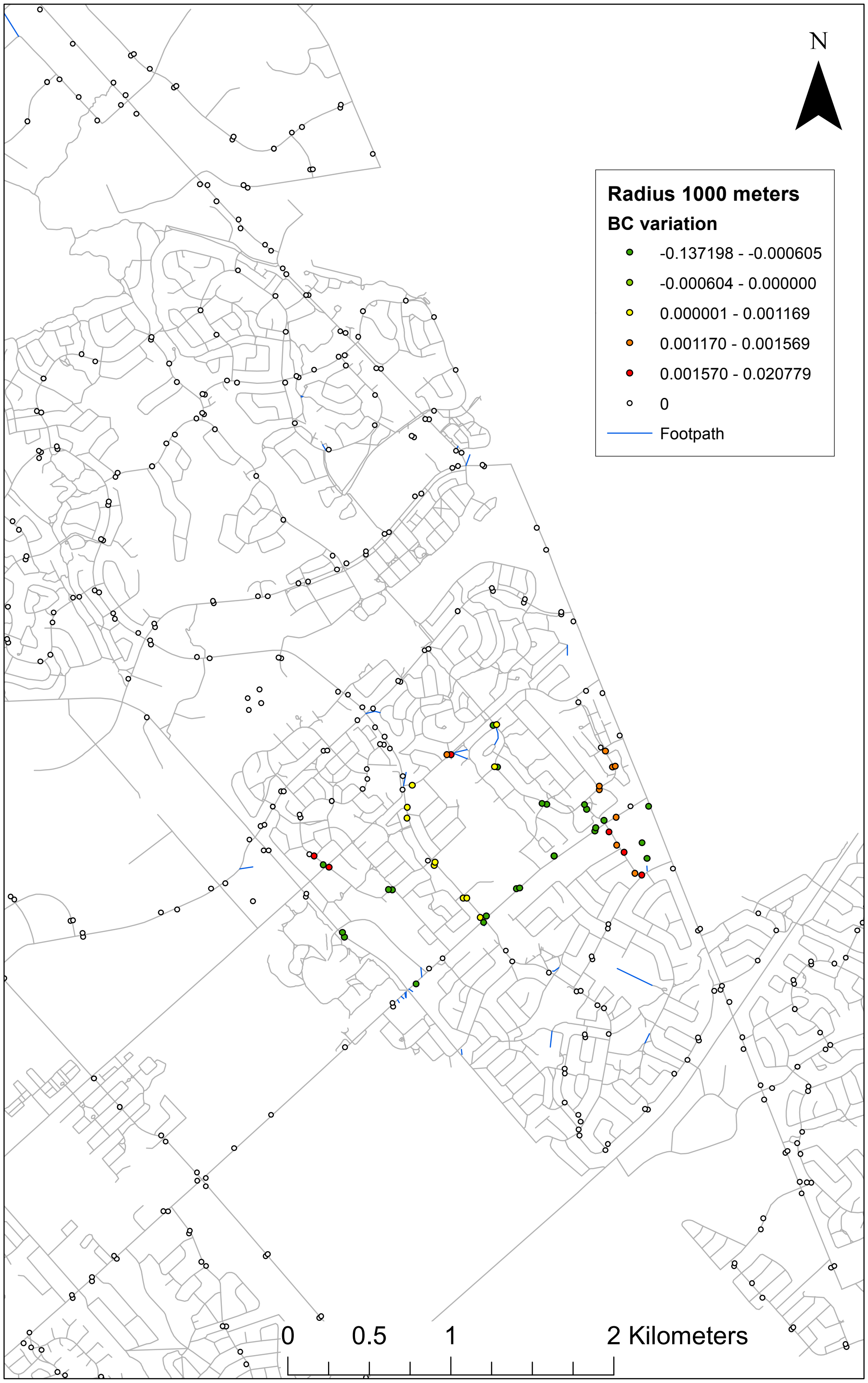

Figure 11. Close-up of area where node BC changes with the addition of footpaths to the network (search radius 1000 meters)

3.2.2.3 Radius 1500 meters

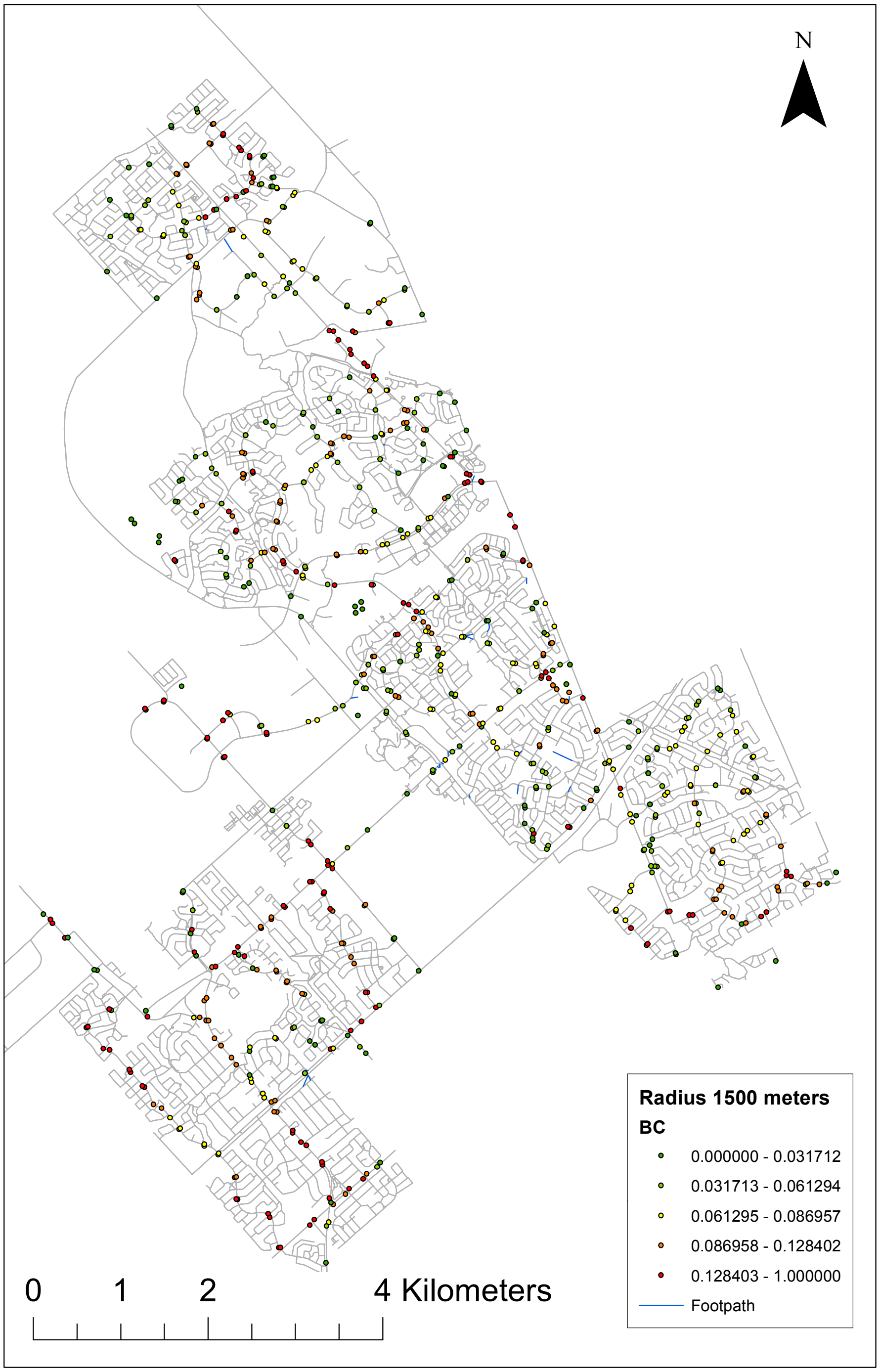

Figure 12. Betweenness Centrality for nodes on the network without footpath (search radius set at 1500 meters)

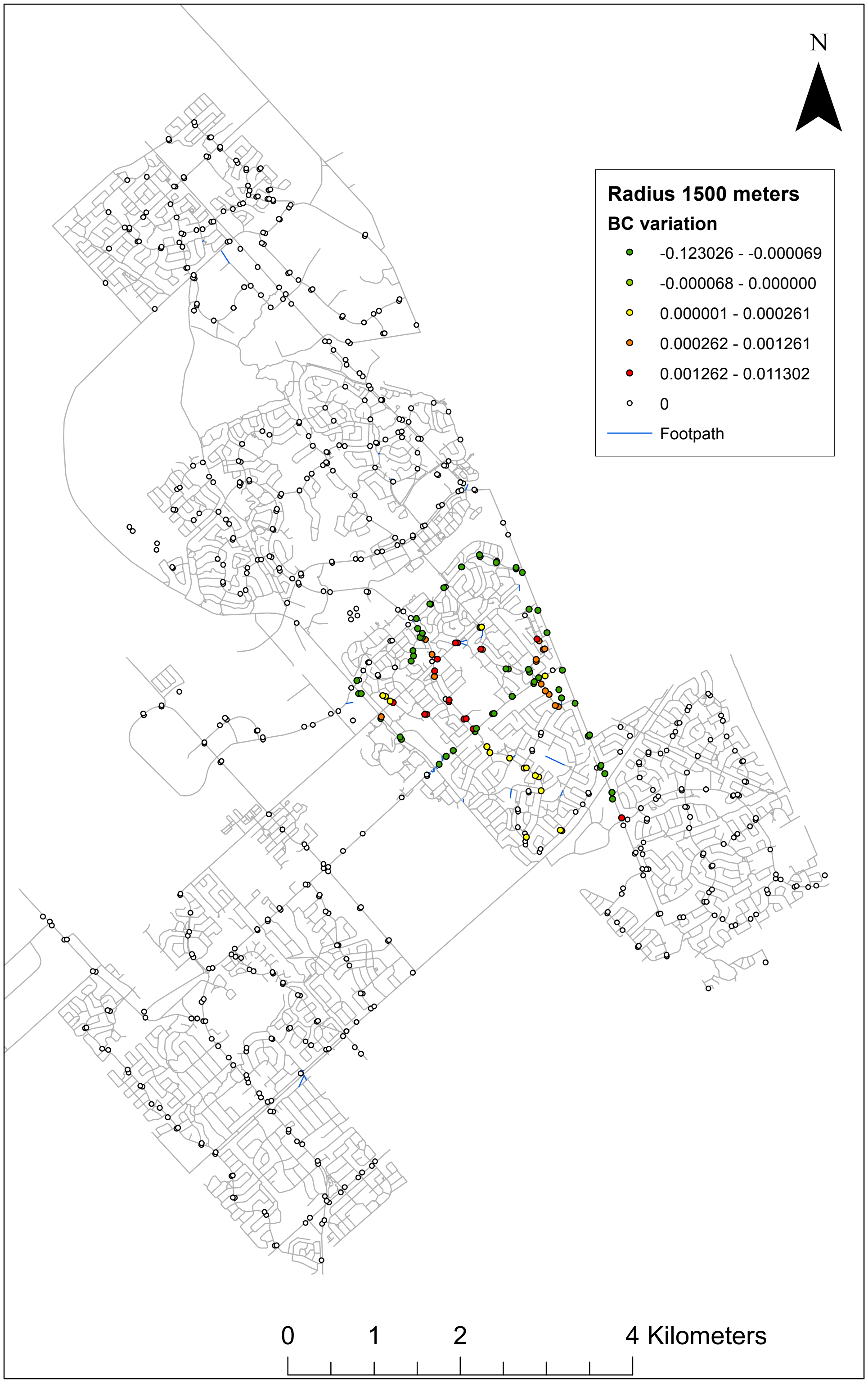

Figure 13. Variation in node BC values comparing the network with footpath to the network without footpaths using a search radius of using a search radius of 1500 meters

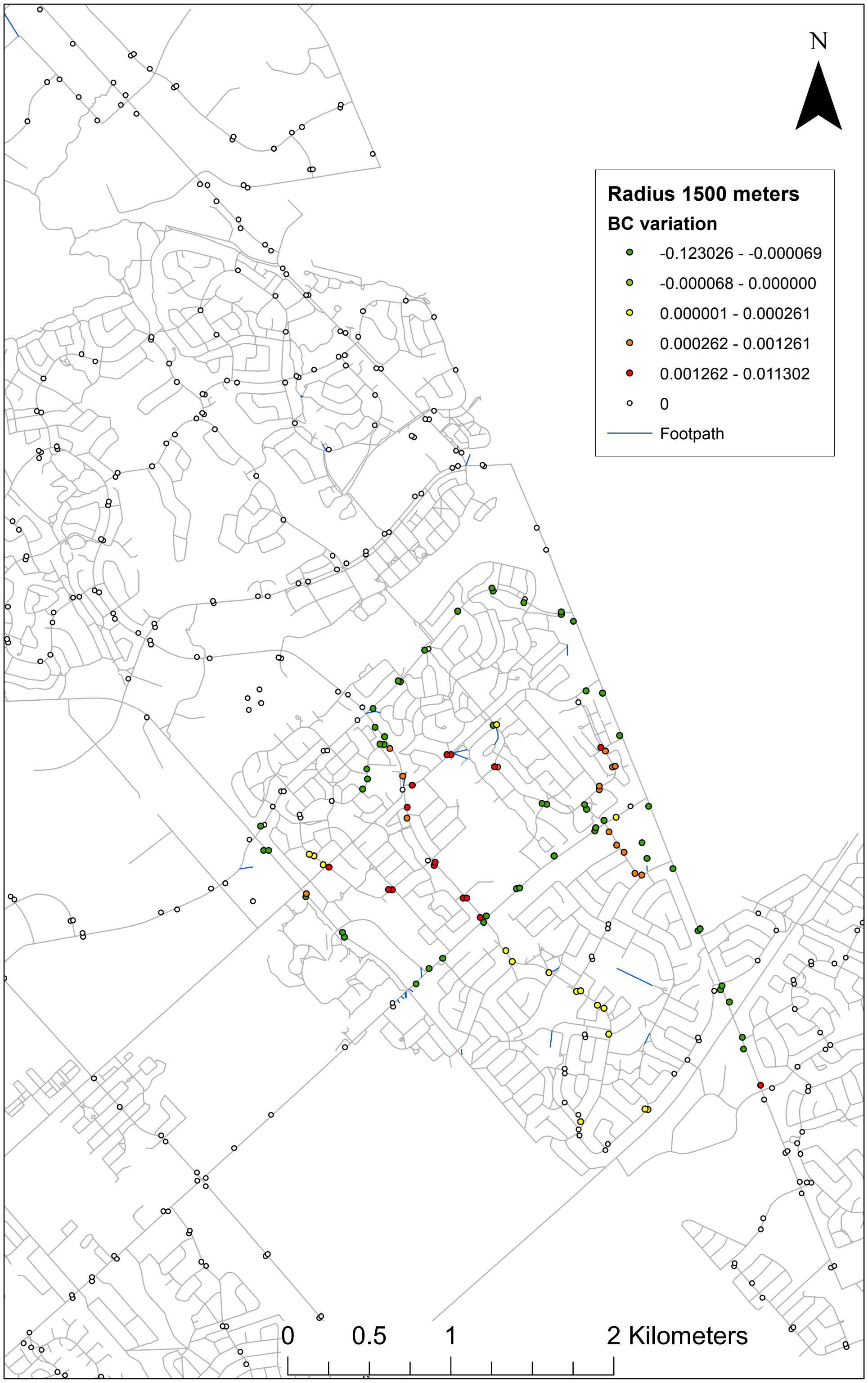

Figure 14. Close-up of area where node BC changes with the addition of footpaths to the network (search radius 1500 meters)

3.2.2.4 Radius 2000 meters

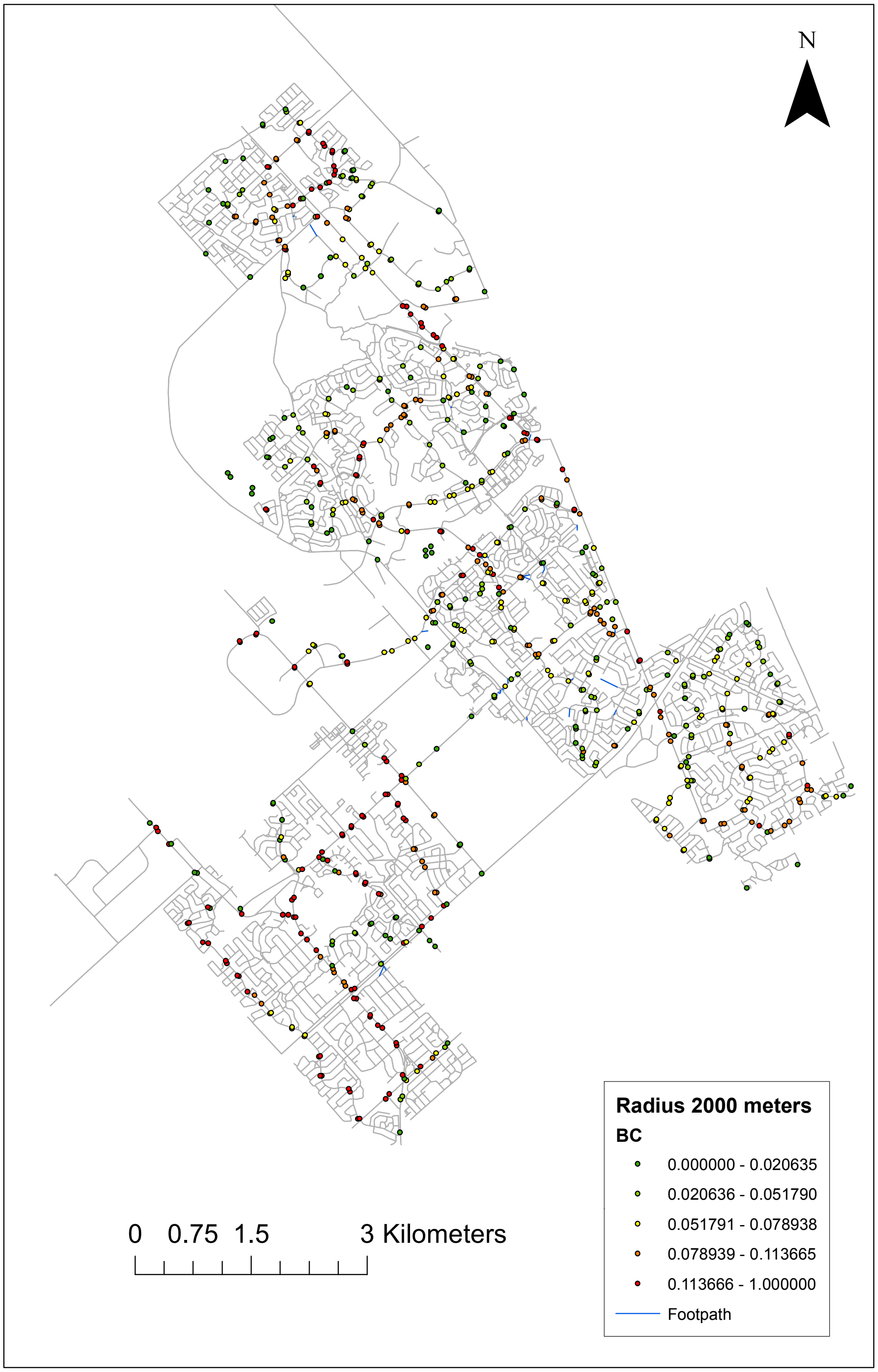

Figure 15. Betweenness Centrality for nodes on the network without footpath (search radius set at 2000 meters)

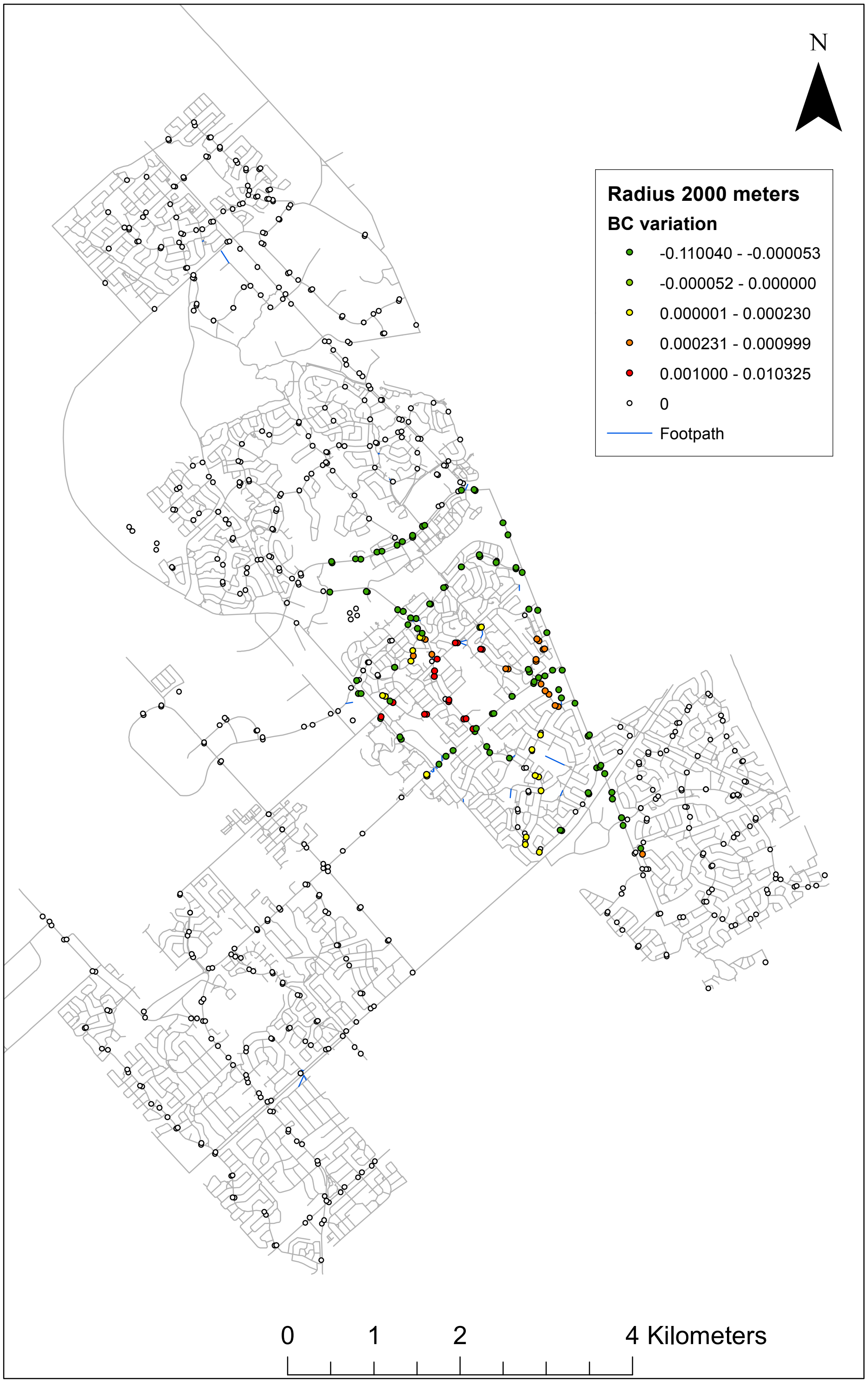

Figure 16. Variation in node BC values comparing the network with footpath to the network without footpaths using a search radius of using a search radius of 2000 meters

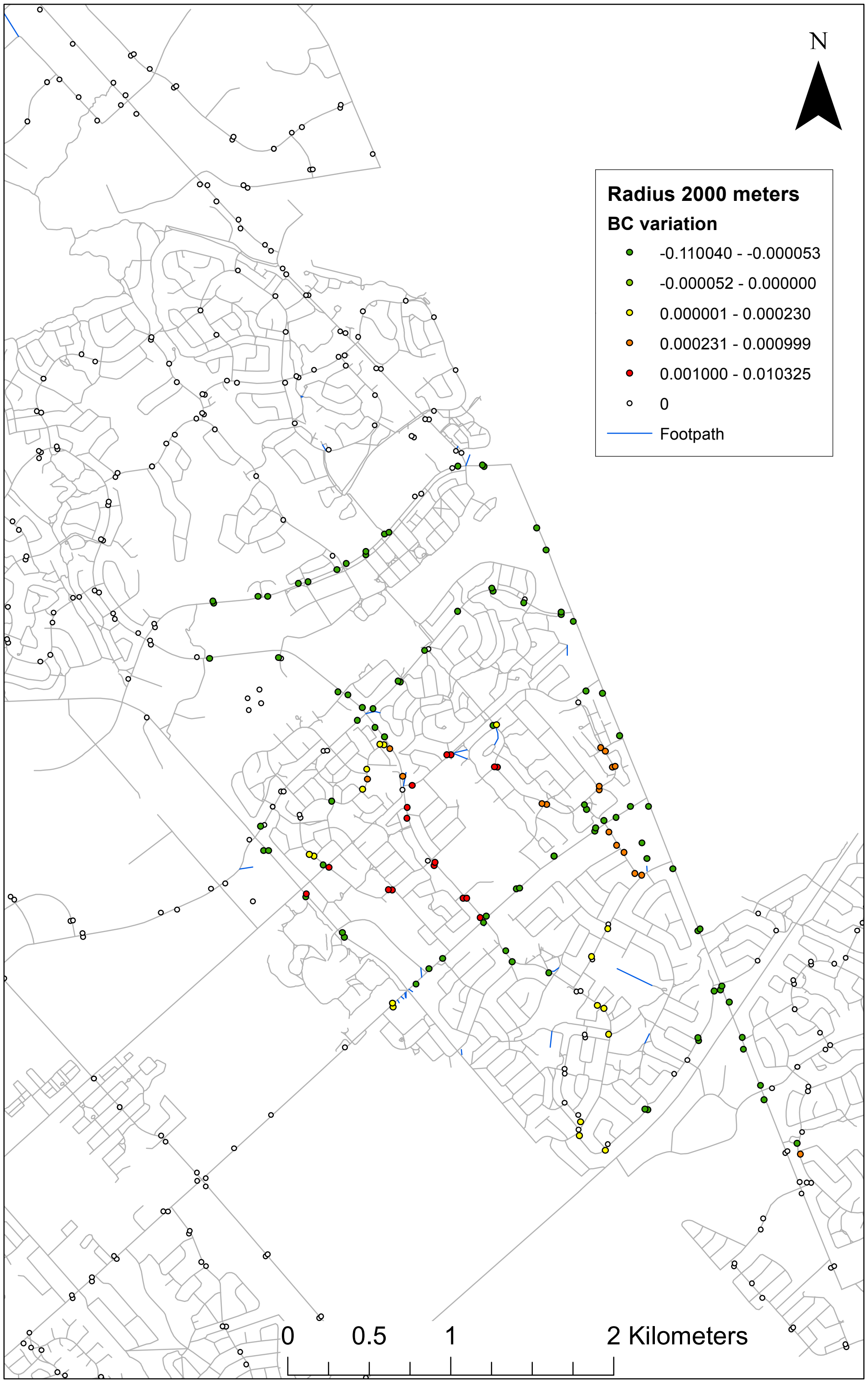

Figure 17. Close-up of area where node BC changes with the addition of footpaths to the network (search radius 2000 meters)

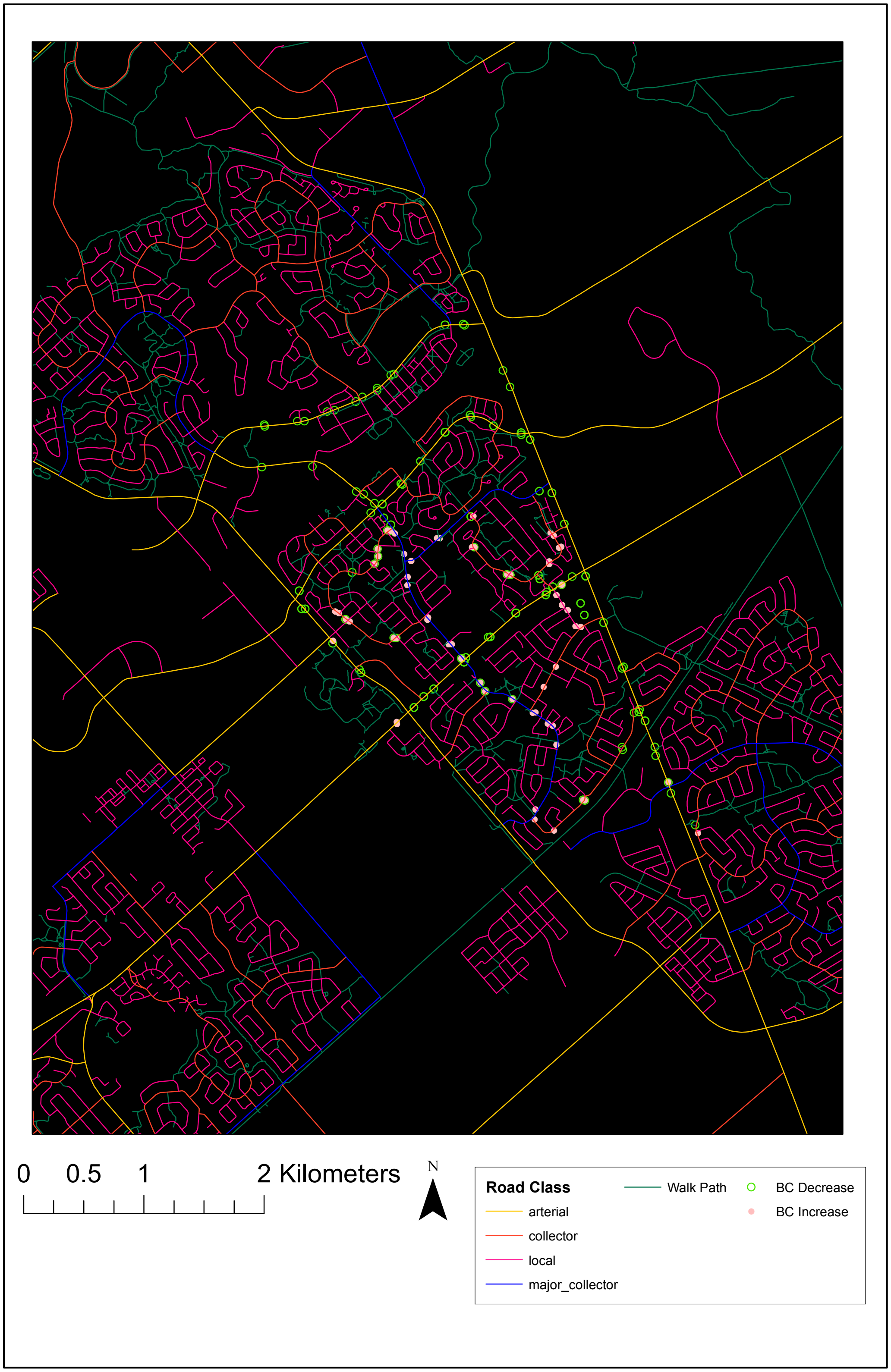

Figure 18. BC variation overlaid unto road type (search radius of 2000 meters) (Data source: City of Ottawa Road Network)

3.3 Destination-based Walkability

Destinations-based measures of walkability were calculated from a set of origins nodes (the network intersections) and destinations (EPOI and Transit stops) within a 2400-meter area. The origin-destination distance cost matrix was generated using the network analysis tool of ArcGIS. A polynomial distance decay function is applied to the average path length to obtain a score out of 100. From the OD-cost matrix analysis, the results show that the addition of footpaths to the existing pedestrian network had no impact on the average path length (Table 10). And therefore, the walkability score is also unchanged by the footpaths (Table 11).

Table 10. Average path length in meters between origins and destinations for the networks with and without footpaths

	Average Path Length in Meters	
	With Footpath	Without Footpath
Glen Cairn	1804.38	1804.38
Kanata Lakes	3693.67	3693.67
Stitsville	2762.18	2762.18
Katimavik	1652.20	1652.20
Beaverbrook*	--	1066.94
Bridlewood*	--	1640.56

*Neighbourhood with no footpaths

Table 11. Walkability score (out of 100) for the networks with and without footpaths

	Score	
	With Footpath	Without Footpath
Glen Cairn	6.44	6.44
Kanata Lakes	0	0
Stitsville	0	0
Katimavik	10.59	10.59
Beaverbrook*	--	52.04
Bridlewood*	--	11.05

*Neighbourhood with no footpaths

Chapter 4
Discussion

4

This chapter discusses the results presented in chapter 3 of the thesis. This chapter is organized into seven sections. The first section focuses on footpaths and network connectivity. The second section addresses the distribution of footpaths within the outer suburban neighborhoods of Ottawa. The third section discusses walkability measures for suburban areas. The fourth section presents and discusses the research limitations. Following this section is the fifth section that proposes recommendations for future researches. The sixth section discusses the contributions of the research. And lastly, the seventh section concluded both the thesis and the chapter by summarizing the main finding of the research.

4.1 Footpaths and Networks Connectivity

One of the objectives of this research was to quantify the impact of footpaths on network connectivity. Network connectivity was measured using geographic metrics. These were the Link-Node Ratio (LNR), the Gamma Index, and the Alpha Index. As well, network node connectivity was also measured using a graph theory metric. The specific metric was Betweenness Centrality (BC). In the first part of this section, the results relating to the geographic metrics are discussed. The second part of this section discusses the result of the BC analysis in relation to pedestrian activity.

4.1.1 Link-Node Ratio, Gamma Index, Alpha Index

Network connectivity was measured using Link-Node Ratio, Gamma Index, and Alpha Index. Network connectivity was calculated for all of the elements of the whole western region (the continuous western neighborhoods beyond the greenbelt). Network Connectivity was also calculated for delineated areas surrounding each of the footpaths. The radius of these buffers was defined as twice the length of the footpath it encircled. In instances where buffers of different footpaths overlapped with one another, the buffers were dissolved to form a continuous area (see Figure 6, Chapter 3).

The results of the network analysis showed two different end results depending on how the network was delineated. When buffers are used to delineate the network, the results of the connectivity analysis show an increase in network connectivity. However, when the network is restricted to the boundaries of the western region, a decrease in connectivity is noted. The results of the connectivity analysis when buffers are used to delineate the network are most likely an artifact of way the pedestrian dataset was clipped to the buffered and the scale at which the network is connected.

The pedestrian network in the western outer suburbs is made-up largely of low-flow local streets where walking is possible on the street, and of sidewalks on higher-flow streets where walking on the street is not possible. These elements-local streets, and higher-flow street- are all created for automobile transport and therefore are at a scale that is greater than that of the pedestrian realm. The speed at which automobile travel is much greater than that of a pedestrian, and so the distances covered by an automobile is generally much greater than what a pedestrian would set out to travel. In the autosuburb, the planning of the streets is oriented towards automobile transport and the walkable elements are then depending on that pre-established scale. These elements-local streets and street-dependent sidewalks- are meant to connect at the distance scale of the automobile and not of the pedestrian. And so when the buffers were created around the footpaths-and defined as twice the length of the footpath- and the network was clipped to that area, many dangle end nodes were created. The reason why connectivity is increased is that footpaths, which are at the scale of the buffer, actually connect at least via one node, sometimes two nodes, to other elements of the network while the other walkable elements are truncated. And so network connectivity appears to be increased with the addition of footpaths.

When connectivity is assessed for the network covering the whole western suburban region, the results show a decrease in the gamma index, the alpha index, and the LNR. This result is explained by how footpaths connect to the network. There are three possible scenarios for how a footpath may connect to the network (see Table 12). Footpaths in many instances connect to the network either to an existing network node or by creating a new node. The gamma index, and the alpha index are metrics based on the assumption that the ideal network should be a circuit network, where the most closed looped there are the more connected the network is. In this research, footpaths were found to, in many instances, terminate without making a connection to the network. This results in a dangle end-node that is not connected to the network. If a footpath introduces 1 link (the footpath itself) and 2 nodes-one to connect to the network, and another dangle end- then this reduces the number of closed loops within the network and is reflected in a decreased in the LNR, the gamma index, and the alpha index.

Although changes in connectivity occur with the additions of footpaths to the network, these changes are rather small. This is because the metrics (LNR, gamma index, alpha index) used to calculate connectivity sums, or ratios, of the total number of link and nodes in a network. In a large network, adding 57 links will likely not result in a drastic change in the connectivity index or ratio. In this research, footpaths accounted for only 0.8% of the total number of links in the network.

Table 12. Summary of how new footpath connection to the network impact connectivity measures (link-node ratio, gamma index, alpha index)

	Link-Node Ratio	Alpha Index	Gamma Index
Footpath connects on both ends at existing network nodes 1 link added 0 node added	Increased	Increased	Increased
Footpath connects on one end at existing network nodes 1 link added 1 node added	No change	Decreased	Decreased
Footpath connect to network by making 2 new connections 1 link added 2 nodes added	Decreased	Decreased	Decreased

4.1.2 Betweenness Centrality

Betweenness centrality is a metric that is calculated for each node within a network and expresses the fraction of shortest paths from surrounding pairs of nodes that pass through that node. As opposed to connectivity metrics such as the LNR, or the gamma and alpha indices, which sum up all the elements of the network, BC provides a global value for a specific network element (Boccaletti, Latora, Moreno, Chavez, & Hwang, 2006) . In this sense, BC comparison analysis for a network with footpaths versus one without footpaths may capture flow-or routing-changes at the scale of the network elements, which is something that is not achievable with the geographic metrics of connectivity used in this research

In this research, BC was calculated for a set of nodes representing the walkable destinations-– EPOI and transit stops- of the network. BC was calculated for those elements at four Euclidian distances radii; 615 meters, 1000 meters, 1500 meters, and 2000 meters. The BC values of the nodes for the network without footpaths were subtracted from the BC values of the same nodes on the network with footpaths. The results show that the BC of some nodes in the center of the network changes when footpath are added to the network. In this area 17 of the 32 of footpaths are found. Other areas with similar node density but with fewer numbers of footpaths do not show BC changes when footpaths are added.

4.1.3 Footpaths in the Suburbs: possible role

The results of the BC analysis may serve to shed some light on the role of footpaths within the existing pedestrian network. As presented in the literature review found in chapter 1, footpaths are conceptualized as elements that serve as shorter routes (desire lines) or better paths (informal trails/paths) towards desired destinations. The results of the connectivity analysis as measured using the LNR, the gamma index, and the alpha index, contributed little towards improving network connectivity. In this sense, they are not short cuts. The results of the BC analysis show areas where footpaths increase and decrease the connectivity of specific nodes. Compared to other regions of the western neighborhoods, the nodes showing a change in BC are located in an area with more than one footpath. This may perhaps indicate that single footpath does little to change connectivity of nodes, but a combination of footpaths though not necessary in vicinity of one another work collectively to provide better paths across space.

Betweenness centrality is connectivity measure that is different from other measure such as the LNR or the gamma and alpha indices. This is because BC is concerned with the spatial structure of the network that is how the different elements of space are connected to one another. When the nodes exhibiting a change in BC values are superposed to the street network, one may notice that the nodes located along arterial streets most often show a decrease in BC. The nodes located on lower flow traffic streets, such as collectors, show both increase and decrease in BC (see chapter 3 Figure 19). A decrease in node BC indicates that fewer shortest paths are going through that one node. While an increase indicates that a higher number of shortest paths now pass through that one node. It is common that pedestrians do not make use of the pedestrian space in the way envisioned by the planners; they instead develop other routes that offer a more comfortable experience (Tomko, Winter, & Claramunt, 2008). Pedestrians then develop another hierarchy of routes that is different from the ones envisioned by the planners and this can be reflected in the BC of nodes (Tomko et al., 2008). The suburban street pattern in Ottawa is formally planned for the automobile where a hierarchy is developed according to traffic flow speed. The decrease in BC values of nodes along high traffic street could indicate that footpaths reorganize pedestrian flows towards streets with a lower flow of automobile traffic, which may be more pleasant walking routes.

Some of the footpaths in the area where most node BC variations are observed attach to the network via only one node, thus terminating in a dangle end-node. This may indicate that 1) pedestrian backtrack along the route to return to their point of origin, or 2) the unconnected end is more dynamic in the sense that it may be reconnected to the network at different position at different moments in time. Helbing et al. (1997), and Goldstone & Roberts (2006) describe the process of footpath formation as a dynamic one. Footpaths develop from the footfall of pedestrian across a deformable surface. It is the repeated use of a path that maintains it. According to Helbing et al. (1997), and Goldstone & Roberts (2006) different users contribute to different portion of footpaths and the segment that is most frequently used persists through time. Those segments that are less travel upon, and therefore less visible, and eventually dissolve over time.

In this research, the footpaths were identified from air photos that were taken in the spring of 2011. This means that following winter, only the portion of footpaths that are the most used are visible. And thereby a footpaths may be reconnected anywhere along the network depending on the needs of the pedestrians. That more than one footpath has a dangle end-node in the same area could indicate that the pattern of movement across this space is not constant over time. In this sense, it could be redefined over time. And so footpaths could serve as temporary paths towards destinations via more pedestrian friendly walking routes.

4.2 Footpath distributions in the outer suburbs of Ottawa

This research attempted to describe the distribution of footpaths within the outer suburbs of Ottawa. The variables were selected in light of existing literature on walkability or footpaths. These variables were related to land usage, public transit connections, income, and population density. A principal component analysis (PCA) was performed on the variables that were sampled in buffered areas surrounding each footpath. The results if the PCA showed no strong grouping of variables in the vicinity of the footpaths. And so it is concluded that the distribution of footpaths in the outer suburbs of Ottawa could not be explained by the selected variables. The results of a PCA analysis are dependent upon the structure of the dataset. The dataset was assessed using the Kaiser-Meyers-Olkins (KMO) test for sampling adequacy. The results of the KMO test indicate that the structure of the data sample at the observation is not suitable for a good PCA. That is the original variables are either too correlated of not at all to one another. In this research, the latter seems to be the case. The correlations between variables are very low and are mostly insignificant (Figure 5, chapter 3). The results of the KMO test and the PCA may be explained by the homogeneity of land usage, demographics, and the evenly spaced transit stops in the outer suburbs.

The North American suburbs are defined by having a predominance of residential land usage with a clear separation between land usages (Gordon & Janzen, 2013). As well, suburbs in Ottawa are mostly auto-suburbs (Gordon & Janzen, 2013) meaning that the scale at which the amenities are placed is one that is meant for automobile transport and not for walking. This implies that most of the sampling sites –the 400 meter buffers around the footpaths for the PCA analysis- will be predominantly residential with other land usage, such as commercial land usage under-represented. In addition, the population density will be more closely associated with certain areas defined by residential land usages and not commercial ones. This makes for a dataset structure that does not lend itself well to data analysis such as PCA.

4.3 Walkability: Indices and ways of walking

Walkability expresses the extent to which a given area is conductive to walking. The walkability of an area, a neighbourhood for example, may be calculated using indices. In this research walkability was calculated using a destination-based index. Destination-based indices, such as the Walk Score, quantify walkability by calculating the distance costs of routes between origins and destinations, which often belong to categories such as retail. The underlying assumption of destination-based walkability indices is that individuals walk from origin to predestined destinations and will chose shortest routes over longer ones. In this sense, walking is always framed as a mode of transport. However, walking may be performed for other reasons, such as for pleasure or health.

In areas where the organisation and planning of the space is geared towards car transportation, such as auto-suburbs, walking may become dissociated from transportation needs. Furthermore, individuals often opt to live in areas of the city knowing in advance what type of transportation is favoured (Jonietz, 2016). Individuals who prefer to walk for transportation will often elect to live in denser areas where retail points are accessible by foot (Jonietz, 2016). In these areas, walking is a mode of transportation and assessing walkability using destination-based indices will be appropriate.

However, in an auto-suburb the walks performed will perhaps be closer to the ones associated with non-urban walking where the emphasis is on the experience of the environment surrounding the pedestrian rather than reaching a retail destination. In this type of walking, aesthetics, terrain, weather, light perceptions are important to the pedestrian to travel through the space (Ingold, 2010). Streets with sidewalks but high flowing car traffic are, in this sense, less attractive than would be a local lesser-traveled street. And so human footpaths may serve to connect, when possible, disconnected attractive route segments. The emplacement of footpaths may also be related to the quality of the environment it traverses thus making for more pleasant path than a paved MUP, trail or sidewalk. As such, the walkability of a space in which walking is performed for non-transportation purposes will not be well expressed by a destination-based walkability index. This is because a destination-based index does not take into consideration any factors related to the aesthetic qualities of the area. In the context of this research, such an index will remain un-altered by the addition of footpaths to the pedestrian network.

4.4 Research Limitations

As is often the case in research, this research has some limitations. The first limitation relates to the footpath data collection from the air photos. The second limitation relates to the method by which walkability is measured.

4.4.1 Data Collection

The footpaths were identified from air photos. This method of data collection is time and cost efficient (Gallagher et al., 2013) . However, it is also dependent on the quality of the air photo. The air photos used in this research had a resolution of 0.2 meters but were collected in the spring of 2011. The time at which the air photos were taken meant that the contrast between footpath and grassy land is not always evident because the grass has not fully grown back after the winter and is therefore pale in color. This makes it difficult to identify lesser-used footpaths, or footpath segments, that are not as strongly etched into the land. This data limitation is perhaps why footpaths appear to end in dangles rather than connect to the pedestrian network. The result is apparent in the connectivity analysis where the addition of footpaths to the network seems to lower connectivity as measured through link-node ratios, as well as gamma and alpha indices.

4.4.2 Calculating Walkability

A second limitation is associated with the methods by which walkability is measured. In this research, as is the case for most assessments of walkability, an index is calculated using the assumption that individuals walk purposefully along shortest routes from an origin to a destination. However as previously discussed the purpose of walking in Ottawa is rarely in the purpose of a commute to and from service points. As well, destination-based walkability indices are calculated using a distance-decay function. The one used in this research assumes that pedestrian will walk a maximum distance of 2400 meters to reach a destination. But the maximum distance a pedestrian is willing to walk is context specific to the city and the activity (Larsen et al., 2010). While developing an Ottawa context specific distance-decay function would have perhaps returned better results, developing such a function would have been too time-costly and beyond the scope of this research. The time-cost associated with creating a context specific function is due to the need to survey the local population regarding their walking habits in order to then create the distance-decay function.

4.5 Future Research Recommendations

Future research is recommended for the distribution of footpaths, and regarding the association of footpaths to non-utilitarian walking. Such future researches should include variables other than the ones used in this research, and should focus on the spatial arrangement of the built environment.

4.5.1.1 Other Variables

The distribution of footpaths could not describe by the variables selected for this research. The land usage variables did not capture small-scale terrain characteristics that make for a more agreeable walking route. For instance in non-urban environments the slope of the terrain plays a role in the formation of trails. In this research, population density was used as a proxy measure for pedestrian flow intensity within the bounded area adjacent to footpaths. While footpaths are a function of deformation group and intensity of pedestrian passage (Goldstone & Roberts, 2006), population density does not necessarily imply heavier pedestrian activity. It is recommended that future research on regarding footpath formation, in suburbs, include variable of the number passages for a given time period (e.g. passages per day)

4.5.1.2 The built environment and footpaths

The morphology of the built environment has been shown to influence pedestrian movement (Hillier & Hanson, 1984)(Westin, 2014). The results of the node betweenness centrality analysis, presented in this research, seem to indicate that footpaths develop as better, but not shorter, routes through the built environment. It is recommended that more research be carried on the role of footpaths in relation to pedestrian movement. The recommended methodology should be to calculate the betweenness centrality of the edges instead of the nodes. Edge betweenness centrality is defined as the numbers of shortest that go through an edge in a network

(Lu & Zhang, 2013). “An edge with a high betweenness centrality represents a bridge like connector between two parts of a network” (Lu & Zhang, 2013). An edge betweenness centrality analysis on the network with footpaths would provide better insights on the role of the footpaths within the pedestrian network in terms of (re)orienting pedestrian movement.

4.6 Contributions

This research contributes to expanding the field of walkability by providing insights on pedestrian made paths, which are elements normally excluded from walkability measures. The results and conclusions of this research should be particularly useful to those involved in creating and planning urban, and suburban, spaces, as well as to researchers.

Understanding walkability from the pedestrian’s perspective provides valuable insights on the real needs of pedestrian within a given urban area. The literature on walkability in urban, including suburban, areas is focused on elements-sidewalks, paved paths- of the formally planned pedestrian network and how it connects to amenities. Human footpaths are normally excluded from most walkability measures because they represent a very small percentage of the pedestrian network and have little impact on how well amenities are connected by the network. The City of Ottawa’s pedestrian plan walkability assessment justifies excluding human footpaths for this very reason (City of Ottawa 2013). In addition, the data as to where human footpaths are located is not readily available. However, although small in numbers, human footpaths may reveal qualities of the environment that make walking more appealing to pedestrians. Such qualities may perhaps contribute less to shortening route distances and more towards a more aesthetically pleasant experience for the pedestrian. These are important variables to take into account when planning and developing better walkable spaces in areas where walking is not transport oriented.

This research also provides a contribution to expending the body of literature on human footpaths in the context of walkability in suburban areas. The novelty of this research is to suggest that walkability in suburban areas should be assessed less in terms of transportation but more in terms of the experience of the environment. Some recommendations for future research are presented in this section. These recommendations suggest using of variables that relate to the aesthetics of the environment, and alternative method for network analysis.

4.7 Conclusions

The objectives of this research were to investigate the distribution of footpaths and their role within the formally planned pedestrian network. While the results of the distribution analysis did not show any patterns of variables around footpaths, the network analysis provided some insight unto the possible role of human footpaths and walkability in a suburban environment.

Footpaths were found to rework network connectivity by increasing the betweenness centrality of some nodes and decreasing the connectivity of other nodes. Nodes whose connectivity was increased are most often located along lower flow traffic roads. In contrast nodes whose connectivity were decreased are most often located on higher flow traffic roads. The role of footpaths in light of these results could be that they are used and created by pedestrian as better but not shorter routes to navigate the pedestrian space. Some footpaths had one well-traced segment that connects to the formally planned pedestrian network but often have an unconnected end node. This unconnected end-node may be reconnected at specific locations of the network that are perhaps changed over time. And so footpaths may serve as dynamic elements of the pedestrian network that evolve over time in response to the needs of pedestrians.

The results of this research provide some valuable insight into walkability metrics, such as indices, used to describe the walkability of an area. Using walkability indices that frame walking as transport-driven may not fully represent the real walkability of suburban areas where walking is performed to meet other needs beyond those related to transportation.

References

Barthélemy, M. (2011). Spatial networks. *Physics Reports*, *499*(1-3), 1–101. http://doi.org/10.1016/j.physrep.2010.11.002

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., & Hwang, D. (2006). Complex networks: Structure and dynamics. *Physics Reports*, *424*(4-5), 175–308. http://doi.org/10.1016/j.physrep.2005.10.009

Frank, L. D., Sallis, J. F., Saelens, B. E., Leary, L., Cain, K., Conway, T. L., & Hess, P. M. (2010). The development of a walkability index: application to the Neighborhood Quality of Life Study. *British Journal of Sports Medicine*, *44*(13), 924–933. http://doi.org/10.1136/bjsm.2009.058701

Freeman, L. (1977). A set of measures of centrality based on betweenness. *Sociometry*, *40*(1), 35–41.

Gallagher, P., Marshall, W., & Atkinson-Palombo, C. (2013). Missing Links: How Social Paths Can Improve Pedestrian Accessibility to Light Rail. *Transportation Research Record: Journal of the Transportation Research Board*, *2393*, 59–65. http://doi.org/10.3141/2393-07

Goldstone, R. L., & Roberts, M. E. (2006). Self-organized trail systems in groups of humans. *Complexity*, *11*(6), 43–50. http://doi.org/10.1002/cplx.20135

Gordon, D. L., & Janzen, M. (2013). Suburban nation? Estimating the size of Canada's suburban population. *Journal of Architectural and Planning Research*, *30*(3), 197.

Haggett, P. (1969). *Network analysis in geography*. London: Edward Arnold.

Helbing, D., Molnár, P., Farkas, I. J., & Bolay, K. (2001). Self-Organizing Pedestrian Movement. *Environment and Planning B: Planning and Design*, *28*(3), 361–383. http://doi.org/10.1068/b2697

Helbing, D., Molnár, P., & Schweitzer, F. (1998). Computer simulations of pedestrian dynamics and trail formation. *arXiv Preprint Cond-mat/9805074*. Retrieved from http://arxiv.org/abs/cond-mat/9805074

Hillier, B., & Hanson, J. (1984). *The social logic of space*. Cambridge: Cambridge University Press. Retrieved from http://ebooks.cambridge.org/ref/id/CBO9780511597237

HILL, M. R. (1982). SPATIAL STRUCTURE AND DECISION-MAKING ASPECTS OF PEDESTRIAN ROUTE SELECTION THROUGH AN URBAN ENVIRONMENT. *ETD Collection for University of Nebraska - Lincoln*, 1–276.

Hoogendoorn, S. P., & Bovy, P. H. L. (2004). Pedestrian route-choice and activity scheduling theory and models. *Transportation Research Part B: Methodological*, *38*(2), 169–190. http://doi.org/10.1016/S0191-2615(03)00007-9

Ingold, T. (2010). Footprints through the weather-world: walking, breathing, knowing. *Journal of the Royal Anthropological Institute*, *16*(s1), S121–S139. http://doi.org/10.1111/j.1467-9655.2010.01613.x

Johnston, R. J. (1980). *Multivariate statistical analysis in geography : a primer on the general linear model*. London ; New York: Longman.

Jonietz, D. (2016). Personalizing Walkability: A Concept for Pedestrian Needs Profiling Based on Movement Trajectories, 279–295. http://doi.org/10.1007/978-3-319-33783-8_16

Kansky, K. J. (1963). *Structure of transportation networks : relationships between network geometry and regional characteristics.* --. Chicago.

Larsen, J., El-Geneidy, A., & Yasmin, F. (2010). Beyond the quarter mile: Re-examining travel distances by active transportation. *Canadian Journal of Urban Research: Canadian Planning and Policy (supplement)*, *19*(1), 70–88.

Lu, D. L., & Zhang, D. M. (2013). Edge Betweenness Centrality. In W. Dubitzky, O. Wolkenhauer, K.-H. Cho, & H. Yokota (Eds.), *Encyclopedia of Systems Biology* (pp. 647–648). Springer New York. Retrieved from http://link.springer.com/referenceworkentry/10.1007/978-1-4419-9863-7_874

Marion, J. L., & Leung, Y.-F. (2011). Indicators and protocols for monitoring impacts of formal and informal trails in protected areas. *Journal of Tourism and Leisure Studies*, *17*(2), 215–236.

Moudon, A. V., Lee, C., Cheadle, A. D., Garvin, C., Johnson, D., Schmid, T. L., … Lin, L. (2006). Operational definitions of walkable neighborhood: theoretical and empirical insights. *Journal of Physical Activity & Health*, *3*, S99.

Myhill, C. (2004). Commercial Success by Looking for Desire Lines. In M. Masoodian, S. Jones, & B. Rogers (Eds.), *Computer Human Interaction* (pp. 293–304). Springer Berlin Heidelberg. Retrieved from http://link.springer.com/chapter/10.1007/978-3-540-27795-8_30

Newsome, D., & Davies, C. (2009). A case study in estimating the area of informal trail development and associated impacts caused by mountain bike activity in John Forrest National Park, Western Australia. *Journal of Ecotourism*, *8*(3), 237–253. http://doi.org/10.1080/14724040802538308

Norman, D. A. (2008). THE WAY I SEE IT Signifiers, not affordances. *Interactions*, *15*(6), 18–19.

Norman, D. A. (2011). *Living with complexity*. Cambridge, Mass: MIT Press.

Parenteau, M.-P., Sawada, M., Kristjansson, E. A., Calhoun, M., Leclair, S., Labonté, R., … Herold, S. (2008). Development of neighborhoods to measure spatial indicators of health. *URISA Journal*, *20*(2), 43–55.

Southworth, M. (2005). Designing the walkable city.(Author Abstract). *Journal of Urban Planning and Development*, *131*(4), 246.

Tal, G., & Handy, S. (2012). Measuring Nonmotorized Accessibility and Connectivity in a Robust Pedestrian Network. *Transportation Research Record: Journal of the Transportation Research Board*, *2299*, 48–56. http://doi.org/10.3141/2299-06

Tomko, M., Winter, S., & Claramunt, C. (2008). Experiential hierarchies of streets. *Computers, Environment and Urban Systems*, *32*(1), 41–52. http://doi.org/10.1016/j.compenvurbsys.2007.03.003

Walden-Schreiner, C., & Leung, Y.-F. (2013). Spatially Characterizing Visitor Use and Its Association with Informal Trails in Yosemite Valley Meadows. *Environmental Management*, *52*(1), 163–178. http://doi.org/10.1007/s00267-013-0066-0

Westin, S. author. (2014). *The paradoxes of planning : a pyscho-analytical perspective*. Farnham, England: Ashgate Publishing Limited.

Appendices

Appendix 1. North American Industry Classification System (NAICS) codes and their description

Code	Industry
11	Agricultural, Forestry, Fishing and Hunting
21	Mining
22	Utilities
23	Construction
31-33	Manufacturing
42	Wholesale Trade
44-45	Retail Trade
48-49	Transportation and Warehousing
51	Information
52	Finance and Insurance
53	Real Estate Rental and Leasing
54	Professional, Scientific, and Technical Services
55	Management of Companies and Enterprises
56	Administrative and Support and Waster Management and Remediation Services
61	Educational Services
62	Health Care and Social Assistance
71	Arts, Entertaining and Recreation

72	Accommodation and Food Services
81	Other Services (Except Public Administration)
92	Public Administration

Appendix 2. R Script for Principal Component Analysis

```
#Importing the data into R
Data = read.csv (“/../…/../FOR_PCA_DATA.csv”)
#Scaling the data
Data=scale (Data)
#Importing psych package for kaiser-meyers-olkins test
Library(psych)
#KMO test on the data
KMO(Data)
# Principal Component Analysis on dataset
PCA=princomp(Data)

#Plot component as histogram
plot(PCA)
#Scree plot of components
plot(PCA, type= “line”)
#export loadings
write.csv(PCA$loadings, “/../../loadings.csv”)
```

Appendix 3. City of Ottawa Land Zoning Classification

Zone Name	Zone Code
Residential Zones	
(1) Residential First Density Zone	R1
(2) Residential Second Density Zone	R2
(3) Residential Third Density Zone	R3
(4) Residential Fourth Density Zone	R4
(5) Residential Fifth Density Zone	R5
(6) Mobile Home Park Zone	RM
Instititutional Zones	
(7) Minor Institutional Zone	I1
(8) Major Institutional Zone	I2
Open Space and Leisure Zones	
(9) Parks and Open Space Zone	O1
(10) Community Leisure Facility Zone	L1
(11) Major Leisure Facility Zone	L2
(12) Central Experimental Farm Zone	L3
Environmental Zones	
(13) Environmental Protection Zone	EP
Commercial/mixed use Zones	

(14) Local Commercial Zone	LC
(15) General Mixed Use Zone	GM
(16) Traditional Mainstreet Zone	TM
(17) Arterial Mainstreet Zone	AM
(18) Mixed Use Centre Zone	MC
(19) Mixed Use Downtown Zone	MD
Industrial Zones	
(20) Business Park Industrial Zone	IP
(21) Light Industrial Zone	IL
(22) General Industrial Zone	IG
(23) Heavy Industrial Zone	IH
Transportation Zones	
(24) Air Transportation Facility Zone	T1
(25) Ground Transportation Facility Zone	T2
Rural Zones	
(26) Agricultural Zone	AG
(27) Mineral Extraction Zone	ME
(28) Mineral Aggregate Reserve Zone	MR
(29) Rural Commercial Zone	RC
(30) Rural General Industrial Zone	RG
(31) Rural Heavy Industrial Zone	RH

(32) Rural Institutional Zone	RI
(33) Rural Residential Zone	RR
(34) Rural Countryside Zone	RU
(35) Village Mixed Use Zone	VM
(36) Village Residential First Density Zone	V1
(37) Village Residential Second Density Zone	V2
(38) Village Residential Third Density Zone	V3
Other Zones	
(39) Development Reserve Zone	DR

Appendix 4. Values used to derive distance decay polynomial function

0	102,0066
0,030801	102,0066
0,064682	102,0066
0,101643	102,0066
0,123203	101,7763
0,147844	101,3158
0,169405	100,8553
0,194045	100,1645
0,218686	99,24342
0,249487	98,09211
0,264887	96,94079
0,283368	96,01974
0,308008	94,40789
0,332649	92,33553
0,35729	90,26316
0,37885	88,42105
0,406571	85,65789
0,421971	83,81579
0,440452	82,20395
0,449692	80,82237

0,462012	79,44079
0,483573	76,90789
0,498973	74,83553
0,514374	72,99342
0,539014	69,76974
0,563655	66,31579
0,585216	63,32237
0,603696	60,55921
0,619097	58,25658
0,634497	56,18421
0,649897	53,88158
0,665298	51,57895
0,680698	49,27632
0,696099	47,20395
0,717659	44,21053
0,729979	42,13816
0,74846	39,83553
0,770021	37,07237
0,788501	34,76974
0,800821	32,92763
0,819302	30,625
0,840862	27,86184

0,856263	26,01974
0,874743	23,94737
0,899384	21,18421
0,920945	19,11184
0,939425	17,26974
0,960986	15,42763
0,982546	13,81579
1,001027	12,20395
1,028747	10,36184
1,062628	8,519737
1,090349	7,368421
1,12423	6,217105
1,161191	5,296053
1,198152	4,605263
1,225873	4,375
1,256674	4,375
1,284394	4,375
1,312115	4,375
1,330595	4,375
1,358316	4,144737
1,389117	3,684211
1,416838	3,223684

1,438398	2,302632
1,463039	1,151316
1,478439	0

Appendix 5. R Script for assigning walkability scores

```
#importing the values used to derived a polynomial distance decay function

> x=read.csv("Values used to derive distance decay polynomial function ")

y=x[,2]

x=x[,1]

#convert the distances to kilometers

x=x*1.60934

plot(x,y)

#Fit the data to a polynomial function

xy.fit=lm(y~poly(x,5,raw=TRUE))

lines(x,predict(xy.fit,data.frame(x=x)),col='red')

#Import the OD-Cost matrix

 r2r=read.csv("FILE")

xl=r2r[,2]

xl=xl/1000

#fit the data to the polynomial function to assign a score

xl.wt=predict(xy.fit,data.frame(x=xl))

xl.wt=xl.wt/100

r2r=cbind(r2r,DISWT=xl.wt,dwtXwt=r2r[,2]*xl.wt)

# export the results to a comma separate value file

write.csv(r2r,"FILE ")
```

Glossary

Terminology and use in the context of this research

Circuits: closed paths loops within the network

Connectivity: A measure of how the elements of a network connect to one another.

Dangle End-Nodes: Dangle end-nodes are formed when the end of an edge, or link, is not attached to another link/edge. For example, cul-de-sac streets are represented as dangle end-nodes in a network.

Links/Edges: The terms 'links' and 'edges' are use interchangeable to describe defined paths upon which pedestrians walk. E.g. Streets, human footpaths, multi-use pathways, trails, sidewalks, local streets

Network: The spatial arrangement of links/edges and nodes.

Nodes: Nodes are formed by the intersection of links, or edges.

Origin-Destination (OD) Cost Matrix: A matrix that represents the cost (in distance or time) of paths between origins and destinations for a given area.

Pedestrian Space: All of the land area available upon which a pedestrian may travel. The pedestrian space included defined paths (streets, human footpaths, multi-use pathways, trails, sidewalks, local streets) and also any unobstructed open area that do not have any pre-traced paths but are still walkable (e.g. Green spaces, fields...) .

Copyright Acknowledgements

City of Ottawa

Pedestrian Network- 2014

Accessed from the City of Ottawa Open Data Portal

http://data.ottawa.ca/dataset/pedestrian-network/resource/f688ed01-f8d8-4358-a582-9af3316094c3

Road-Road Segment – 2014

Accessed from the City of Ottawa Open Data Portal

http://data.ottawa.ca/dataset/roads/resource/f72339eb-769f-48ec-a18a-9965a879527e