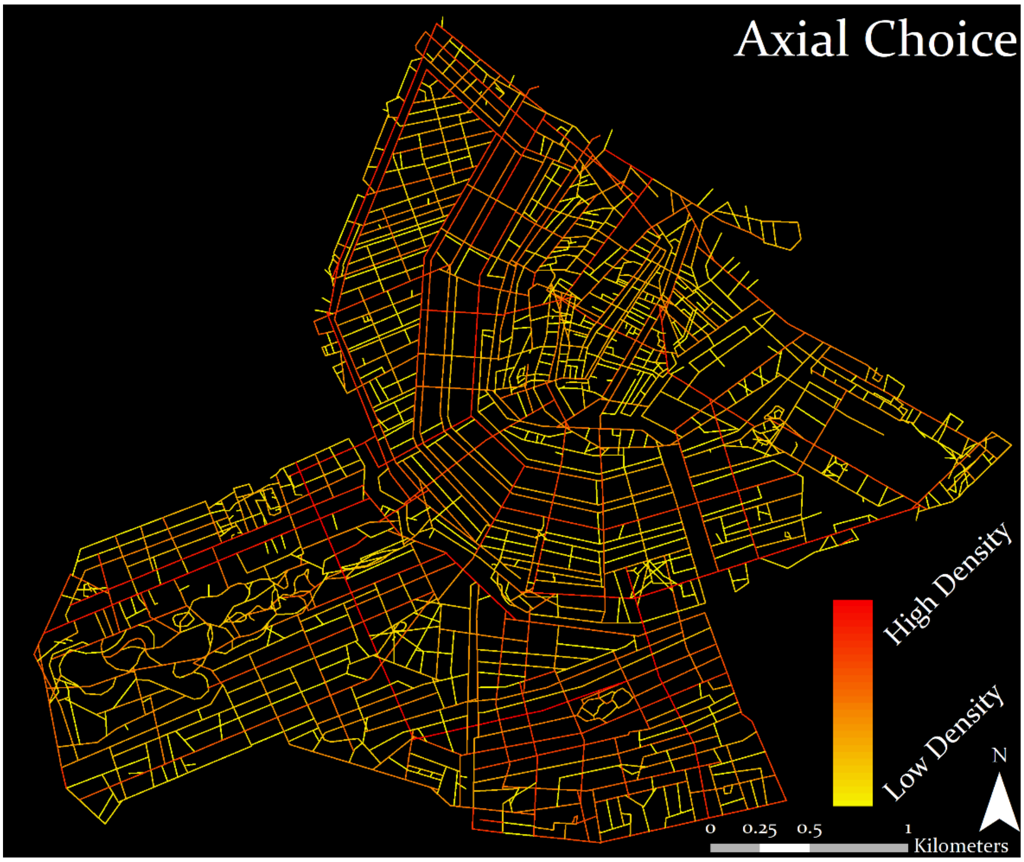
Laboratory for Geo-Information Science and Remote Sensing

Thesis Report GIRS-2019-26

Using agent-based modelling for simulating the effects of space syntax values on pedestrian movement patterns

Simon Veen



08-03-2019

WAGENINGEN
UNIVERSITY & RESEARCH

Using agent-based modelling for simulating the effects of space syntax values on pedestrian movement patterns

Simon Veen

Registration number: 921228856080

Supervisors:

Dr. ir. A (Arend) Ligtenberg

A thesis submitted in partial fulfilment of the degree of Master of Science

at Wageningen University and Research Centre, The Netherlands.

08-03-2019

Wageningen, The Netherlands

Thesis code number: GRS-80436
Thesis Report: GIRS-2019 - 26
Wageningen University and Research Centre
Laboratory of Geo-Information Science and Remote Sensing

Abstract

Pedestrian movement has been studies for more than five decades. Space syntax is a network analysis method that investigates the global and local urban grid using axial lines. It also has techniques for observing movement within a network. While originally created for the analysis of spatial configurations, a correlation between the measured values and pedestrian movement patterns was found. Also included in this thesis are the angular segment analysis, which is an improvement on the space syntax and includes the angles between streets into the measurements, and PageRank measurements, which values are found to better correlate with movement flows than integration values.

These values are calculated for a study area in the city centre of Amsterdam and used in an ABM. In the model, the agents display exploratory behaviour and base their wayfinding on the characteristics of the network. The purpose of the model is to gather movement patterns for the space syntax and PageRank values and compare them to each other.

Three different movement pattern maps can be distinguished, and the choice values give the most logically human movement patterns. The other values have the same result as a non-preference walk. A reason for this could be the size of the study area, and a smaller study area could provide more meaningful results.

Keywords: *Space Syntax, ABM, Human movement patterns, PageRank, Integration, Choice, Axial lines, angular segment analysis*

Acknowledgements

Starting this thesis back in October 2017, I knew this would be a challenging project at the end of my master MGI at Wageningen University. Combining working on this thesis with other courses, an internship at Movares and setbacks in my personal life proved to be difficult at times. I would like to thank my family, friends and roommates for the personal support which has proven to be invaluable. In the end, while it took longer then expected, I'm happy with the final result.

Special thanks go to Arend Ligtenberg, my supervisor during the course of the thesis. His guidance and input on the subject has helped me in the right direction when I strayed on less important things and the sparring sessions together resulted in excellent ideas.

Table of content

Abbreviations

ABM	Agent Based Model
ASA	Angular Segment Analysis
BGT	Basiskaart Grootschalige Topografie
CH	Choice
EVA	Exosomatic Visual Architecture
MD	Mean Depth
MD_α	Angular Mean Depth
NACH	Normalized angular choice
NAIN	Normalized angular integration
TD	Total Depth
RA	Relative Asymmetry
RRA	Real Relative Asymmetry

1 Introduction

1.1 General Background

Pedestrian movement have been studied for more than five decades (Helbing & Johansson, 2013). Understanding the way pedestrians move around their environment is important for predicting congestion, pedestrian traffic and crowd control. Another benefit is the assessment on how the urban layout and different buildings function (Penn & Turner, 2001). Collecting pedestrian data for every urban area and street network is an unachievable task, therefore pedestrian volume models are used for calculations on pedestrian movement across a network (Omer & Kaplan, 2017). Previous studies divide these models according to two main approaches, each with a different scale (Omer & Kaplan, 2017; Raford & Ragland, 2005). Network analysis models are applied on a city-wide scale or neighbourhood levels whereas agent-based models (ABM) focus on a small number of streets or intersections.

1.1.1 Space Syntax

Space Syntax is one of the network analysis models that investigates the global and local urban grid and has techniques for observing movement within a network (Hillier et al., 1993). Space syntax was originally created for the analysis of spatial configurations and to help understand the configured space itself (B Hillier & Hanson, 1984). It consists of two different analytic theories, the convex map and the axial map. The convex map is consisting of the "fattest" convex spaces, and is created by identifying the biggest convex space, and then identifying the second biggest convex space, continuing until the entire area is subdivided (Bafna, 2003). The advantage of the convex space is that it can easily describe the sociologically relations within a network, but the disadvantage is that the resulting picture is static. To capture the possible mobility options within a network, the axial map can be used. The creation of the axial map is also an iterative process. First the longest possible straight line is created, followed by the second longest straight line until there are no more options left. The simplicity of the axial map makes it an analytic method of choice for researchers who focus on movement within spatial settings (Bafna, 2003). Using the axial map, it is possible to calculate different measures about the network, such as the mean depth (shortest path between two nodes) and integration (how well a node is connected to the rest of the network) (B Hillier & Hanson, 1984). While space syntax was created for better understanding of configured space, a correlation between the measured values and pedestrian movement patterns was found (Bafna, 2003; Kim & Piao, 2017; Penn & Turner, 2001). This discovery led to more interest in the topic spatial cognition among space syntax researchers. More analytic theories to represent space were developed as well as other measures to calculate different values of the axial map indices (Kim & Piao, 2017). An example of the better representation of space can be found in the angular segment analysis (ASA), in which the angular sum of the angles is treated as the cost of the journey. This method has found an even better correlation between ASA measures and movement (B Hillier & Iida, 2005).

1.1.2 *ABM*

ABMs are simulation models that describe individual entities, commonly called "agents", and have become a useful tool in different disciplines that deal with complex systems made up of autonomous entities (Grimm et al., 2006). It is a bottom-up approach in which each agent interacts both with each other and their environment and can make decisions based on this interaction. Each agent makes these decisions based on a set of rules (Bonabeau, 2002; Matthews et al., 2007). In contrast to the top-bottom approach of space syntax, ABMs have a more dynamic dimension which makes it a suitable tool for the simulation of movement flows (Omer & Kaplan, 2017). While useful on a small scale, the ABM approach has some limitations. Due to its complexity, it is an approach that is hard to upscale and the accuracy and completeness of the inputs determine whether the output is a correct forecasting or just a qualitative insight (Castle & Crooks, 2006).

1.2 Problem definition

Space syntax literature has proven that there is a relation between the space syntax measures and human movement behaviour (B Hillier & Hanson, 1984; B Hillier et al., 1993; Penn & Turner, 2001; Alasdair Turner, 2007). This relation is lost when an area becomes less intelligible (that is, the correlation between local and global spatial measures decreases)(Penn, 2003). Space syntax can explain up to 80% of the variance in pedestrian flows from location to location using the spatial configurational properties of these locations within the network (B Hillier et al., 1993; Penn & Turner, 2001). This is surprising, since the network analysis method looks from a topological perspective to the network, without considering factors such as distance, land use and population density (Jiang & Jia, 2009). It is a method which excludes the preferences of individuals but is still able to predict the movement behaviour of a population of individuals (Penn, 2003).

In ABMs the set of rules are applied on an individual level, with each agent having its own preferences and goals. Most agents in pedestrian models are using the shortest path method to go from origin to destination. While shortest path is an interesting method, not all types of pedestrians use this method. Local inhabitants might know the shortest path to their destination, a tourist who is unfamiliar in the area does not and will therefore rely more on the environment. Since space syntax is a method for the analysis of the urban environment and can predict pedestrian movement flows, it might prove to be an interesting alternative to the current shortest path method. Some research has already been done on the possibility to incorporate space syntax measures into an agent-based model (Kim & Piao, 2017; Omer & Kaplan, 2017; Penn & Turner, 2001). However, what kind of pedestrian movement flow patterns will develop if the pedestrian only used space syntax measures for their wayfinding on the urban network?

1.3 Research objective

In this thesis, I explore the potential of space syntax to represent human movement behaviour and how the space syntax measures can be used to define an agent-based model. A real-world case study in Amsterdam is used to test the model.

This thesis aims to answer the following research questions:

1) Which different space syntax theories and analysis models are there that quantify relations between space syntax and human movement behaviour?

2) How can the relation between space syntax and human movement behaviour be translated into individual agent decisions?

3) How can the relation defined in question 2 be implemented into an agent-based model?

4) What are the effects of different space syntax measures on (simulated) human movement behaviour patterns and how do these compare to each other?

2 Theoretical framework

2.1 Space Syntax measures

The original space syntax theory proposed by Hillier and Hanson in 1984 introduced two different representations of space; the convex map and the axial map (Hillier & Hanson, 1984). While the convex map is needed for the creation of the axial map, all mobility measures are calculated using the axial map. Therefore, Only the axial map will be used in this thesis. The definition of the axial map is as follows:

> "An axial map of the open street structure of the network will be the least set of axial lines which pass through each convex space and makes all axial links"
> (B Hillier & Hanson, 1984)

While originally drawn by hand, an algorithm was later found to create the axial maps (Alasdair Turner, Penn, & Hillier, 2005). An example of a network and its corresponding axial map can be found in figure 1. The difference between an axial map and a network map is that the axial lines are the longest possible lines in a network system which means a new axial line starts when a road ends or makes a turn. A network map uses lines situated in the road centers which are broken across junctions, and therefore the corresponding networks tend to vary with physical distance rather than the changes of direction as measured with axial networks (Turner, 2007).
To show the relations between discrete spatial elements such as axial lines, the network is sometimes represented as a graph, with each spatial element denoted as a small circle or node and its relationship with other elements denoted as a line between the circles.

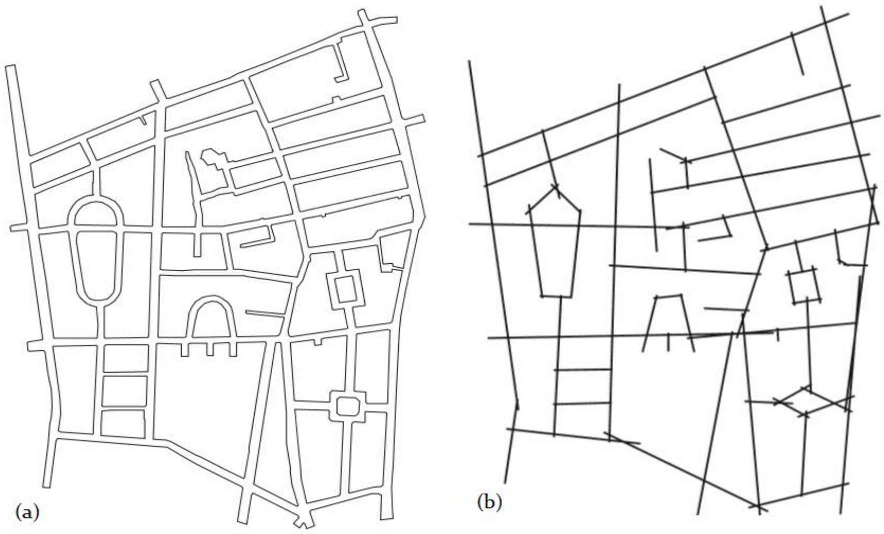


Figure 1. (a) map of an area in Barnsbury (b) and the corresponding axial map (B Hillier & Hanson, 1984)

2.1.1 *Original measures*

Hillier and Hanson introduced analysis measures to gather more information about the network using the indices created with the axial map.

The central concept of space syntax is integration (Teklenburg et al., 1993). Integration (also called closeness centrality) represents the measure of distance from an axial line to all other axial lines within the system. In general, it calculates how close each line is to all other lines. It is affected by the entire spatial configuration (Bafna, 2003). This is also called the to-movement potential or topological depth, where the shallowest axial line is closest to all other axial lines and the deepest is the furthest one (Al_Sayed et al., 2014; Omer & Kaplan, 2017). An example of a shallow line is a main street running through a city center and an example of a deep line is an alley in the suburbs. The depth in space syntax analysis is measured in steps between two lines instead of using metric distance values. It is calculated by the sum of the number of lines between two streets and adding 1 (Teklenburg et al., 1993). This means that two axial lines intersecting will have a depth of 1 between them. The sum of all depths from one axial line to every other axial line in the network is called the total depth (TD). Since these values can get very high in a large network, space syntax also works with the mean depth (MD) of an axial line. This is defined in equation 1:

$$\mathrm{MD} = \frac{\mathrm{TD}}{n-1} \tag{1}$$

Where the total number of axial lines is denoted as *n*. However, since the MD increases when the number of axial lines in the network increases, it is impossible to compare different sized networks. Therefore, a standardization is needed (Teklenburg et al., 1993) The first value is called relative asymmetry (RA). RA represents the centrality of an axial line by comparing its actual MD with the theoretically highest and lowest values that MD could have. The RA is calculated using equation 2:

$$\mathrm{RA}_{(i)} = \frac{2(\mathrm{MD} - 1)}{n-2} \tag{2}$$

With $\mathrm{RA}_{(i)}$ being any axial line within the network. The advantage of RA over the MD is that the result is a value between 0 and 1, which makes it easy to compare with RA values of other networks. High RA values indicate a low integration and vice versa (Teklenburg et al., 1993). Another measure introduced was the real relative asymmetry (RRA), which is a ratio of the RA value of an axial line and the RA value of the central node of a diamond graph with the same number of nodes as the number of axial lines in the system (Bafna, 2003). The diamond graph is characterized by an almost normal distribution of nodes across its levels and so had been found to represent a more realistic way to compare spatial networks of different sizes (Bafna, 2003). This diamond graph is called D_n and is calculated using equation 3:

$$D_n = 2\left\{n\left[\log_2\left(\frac{n+2}{3}\right) - 1\right] + 1\right\} / (n-1)(n-2) \qquad (3)$$

With the RRA of an axial line is defined in equation 4:

$$\mathrm{RRA} = \frac{\mathrm{RA}}{D_n} \qquad (4)$$

Again, the values of RRA are above zero just as RA and a high value means a low integration (Teklenburg et al., 1993). Current space syntax studies report integration values instead of MD or RRA values, which are the inverse of RRA values and are calculated using equation 5:

$$\mathrm{INT} = \frac{1}{RRA} \qquad (5)$$

Another important space syntax measure is choice (or betweenness centrality). Choice (CH) represents the extent to which a street functions as an intermediate location within the network. This can also be seen as it through-movement potential (Omer & Kaplan, 2017). The choice measure counts the number of times a street segment lies between all the origin-destination pairs in the network. Choice values in a network are distributed exponentially. Most axial lines render low values, while a minority of axial lines have higher than average values. These streets are most picked going from origin to destination on a local and global scale (Al_Sayed et al., 2014).

The last important space syntax measure is connectivity, which is defined as the number of other streets directly connected to a street. The degree of correlation between connectivity and integration can be used as a measure for the intelligibility of the entire network (Bafna, 2003).

Some research shows that the integration measure (closeness) and the choice measure (betweenness) can be used to predict human movement (Hillier & Hanson, 1984; Kim & Piao, 2017; Turner, 2001). However, many researchers have noted that choice seems to be a more intuitive model for movement (Turner, 2005). One of the most promising arguments for this is that people prefer to take routes that minimize trip length or maximize trip efficiency (Hillier, 1999) which translates to streets with better axial measures compared to the rest of the network.

2.1.2 *Angular segment analysis*

After the space syntax literature was introduced, it was soon proven that there is a relation between the space syntax measures mentioned in chapter 2.1.1 and human movement behaviour (B Hillier & Hanson, 1984; B Hillier et al., 1993; Penn & Turner, 2001; Alasdair Turner, 2007). After this correlation was found, a large body of research was done on this subject to further improve the correlation. In 2000 Turner proposed angular analysis; first as an extension of visibility graph analysis (see chapter 2.3.1) but later it was also found to improve on the current axial analysis (Turner, 2000). In original space syntax measurements, the measure of depth is in terms of the number of turns from origin to destination, biasing all turns equally. While on itself this method works, it has no regard for human nature. Research shows that some turns are more important than others to pedestrians (Turner, 2001). Dalton (2001) shows that when taking a route, people linearize it, taking shallower turns towards their goal and even prefer to walk in a straight line when going to their destination (Dalton, 2001). Instead of treating every turn the same, Turner (2001) included the angles into the integration measurements. Angular analysis uses a weight graph based on the angle (in radians) of connecting axial lines to calculate space syntax metrics rather than the standard measures. The angular sum is threated as the cost of a supposed journey through the network graph, and from it a shortest (least cost) path from one segment to another across the system can be calculated (Turner, 2007). To calculate the total angular depth, denoted as TD_α, the shortest angular path from each line to every other line in the system is calculated. The shortest angular path is the path from origin to destination with the lowest angular sum. The mean angular depth, denoted as MD_α, of a line is the sum of the shortest angular paths over the sum of all angular intersections in the system (Turner, 2000).

The MD_α is calculated using equation 6:

$$\mathrm{MD}_\alpha = \frac{\sum_{b \in V(L)} l_{ab}}{\sum_{e \in E(L)} w_e} \tag{6}$$

with l_{ab} the shortest angular path between lines a and b, $V(L)$ the set of all axial lines in the system, $E(L)$ the set of all connections between the axial lines in the system and w_e the angle, and therefore weight, of each individual connection.

Figure 2 shows an example of the calculation of mean depth and angular depth for a simple axial system. In this figure the network is shown as a graph in which each

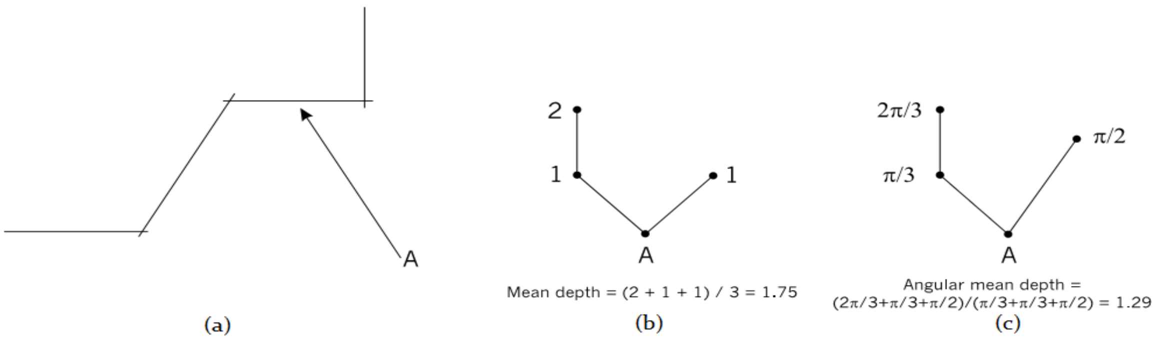


Figure 2. (a) Simple axial system, (b) mean depth calculation for line A in graph format, (c) angular mean depth calculation for line A in graph format (Alasdair Turner, 2000).

axial line is represented by a node and its relationship with the other axial lines is denoted as a line which joins the nodes.

While the axial map can be used to calculate the original space syntax measures, the axial map proved less useful for angular analysis. The method has trouble trying to detect semi-continuous lines (lines with a curvature over time) in a system and dealing with axial lines which were cut in the middle by another line (Conroy, 2001; Turner, 2001). As a solution a segmented map was proposed, with each axial line being split into segments between points where axial lines cross. This would not affect the total- or mean angular depth in a system, since the angles between the axial lines stay the same but result in a finer grained result of the axial map. This method is called angular segment analysis (ASA) (Turner, 2007). ASA creates a finer grained representation of the axial map and will generate better results than an angular analysis done on an axial map (Al_Sayed et al., 2014). Using this representation, the integration and choice values can be calculated for each separate segment.

To enable comparison between two systems of different sizes, two normalized values were introduced; the normalization of angular choice (NACH) and angular integration (NAIN) measures (Hillier et al., 2012). These two methods make it easier to expose the inner structure of a city and making it possible to compare street configurations in different cities based on their maximum and mean values (Al_Sayed et al., 2014). The calculation of NACH and NAIN are shown in equation 7 and 8:

$$\text{NACH} = \frac{\log CH + 1}{logTD + 3} \tag{7}$$

$$\text{NAIN} = \frac{n^{1.2}}{TD} \tag{8}$$

Using the ASA method, Turner (2007) found a correlation $R^2 = 0.82$ between the angular measures and pedestrian flows, which is an increase compared to the correlation found using the original space syntax measures.

2.1.3 PageRank

Most literature uses the integration values as a default indicator for pedestrian movement flows. However, research done by Jiang states that PageRank values were found to better correlate with movement flows than integration values (Jiang & Jia, 2009).

PageRank is a key technology behind the Google search engine that decides the relevance and importance of certain web pages (Jiang, 2008). The PageRank value is justified as a random surfer who jumps from page to page by clicking hotlinks. The visiting frequency of an individual web page can be characterized as the PageRank value (Jiang & Jia, 2009). This technology can also be used on a space syntax network, where the surfers are the pedestrians and the webpages the nodes. It is important to note that the PageRank algorithm calculates values for the nodes (junctions) in the network instead of the street segments. The propagation of the PageRank follows an uneven rule, which means that more popular nodes tends to attract more visitors. In other words, it is proportional to the connectivity of the linked nodes.

The PageRank is calculated using equation 9:

$$PR(i) = \frac{1 - d}{n} + d \sum_{j \in ON(i)} \frac{PR(j)}{n_j} \tag{9}$$

Where n is the number of nodes, *ON(i)* the number of nodes that point to node *i*, *PR(i)* and *PR(j)* are the rank scores of nodes i and j, d is the damping factor and n_j denotes the number of outlink nodes of node j.

Using the PageRank equation, a pedestrian gives a higher priority to nodes that are better connected within a network. Previous research also found that the correlation between PageRank and movement gets better when the damping factor increases, (Jiang, 2008).

2.2 Pedestrian modelling

Since the increase in more powerful technology and better measurement tools, models have been used to help understand and predict pedestrian movements. These models can be divided into three approaches with the difference being in the scale of application (Raford & Ragland, 2005):

1. Sketch plan models
2. Network analysis models
3. Microsimulation (agent-based) models

In this thesis sketch plan models are not used and therefore will not be discussed. Like already stated in the introduction, an example of the second approach is space syntax which is discussed in chapter 2.1. This chapter will discuss the third approach and how agent-based models are being used for pedestrian modelling. ABMs have been used in different ways to explore pedestrian movement problems such as navigation in confined spaces or how people evacuate a building during a fire (Crooks et al., 2015). The quality of these simulations highly depends on the model used and the behavioral attributes of the agents.

ABMs can have different definitions based on the agent properties used. An individual ABM for instance is an ABM in which the agents are represented as individuals with in advance set characteristics. In an adaptive ABM the agents are interacting and autonomous and they change their behavior during the simulation (Macal, 2016). While many pedestrian movement models exist and are being used in different science fields, only a few of these models have been validated against actual human data since there is a lack of quantitative data of human crowds at a resolution that would prove useful for accurate modelling (Fridman & Kaminka, 2010). Therefore, most agents are given a set of behaviors that is based on empirical or qualitative data (Crooks et al., 2015). Schelhorn et a. (1999) divides the characteristics values of an agent into two categories: socio-economic (income, gender) and behavioral (walking speed, fixation). These values can be the same for each agent or based on statistically determined distributions.

In the ABMs used for pedestrian movement modelling two types of agents are used. The first type of agent is goal oriented, also called a purposive walker. This agent has a clear goal, or multiple goals, in mind as it walks through the network (Jiang & Jia, 2009; Omer & Kaplan, 2017). These agents choose their route usually using a shortest path algorithm. The other type of agent is a random walker, also called an exploratory walker. This agent has no clear goal and walks along the street until an intersection. At the intersection a decision will be made on the next street to traverse on. This can be totally random or with a small preference to certain streets. This preference can be based on anything, such as a preference for better sight lines or higher connected streets (Jiang & Jia, 2009; Penn & Turner, 2001).

2.3 Space syntax implementations

Before creating an agent-based model using space syntax measures as rules for agents, it is good to see what kind of research has already been done on space syntax implementations in ABMs as a frame of reference.

2.3.1 Exosomatic visual architecture

One of the key elements of the relation between spatial configuration and movement is natural movement. The basics of natural movement is that movement in an urban grid is determined by the street network structure itself (B Hillier et al., 1993). Turner and Penn (2002) wanted to improve on pedestrian models by encoding natural movement as an agent-based system and therefore creating a more 'natural' way the agents interact with their environment. To do this, they introduced an Exosomatic visual architecture (EVA) to help the agents guide them through the network (Turner & Penn, 2002). Underlying an EVA is a visibility graph. A visibility graph is computed by overlaying a two-dimensional grid over a network and calculating which points within the grid can see which other points. The set of visible locations for each point are stored and thus the visibility graph can be used to calculate the isovist from each point on the grid (Turner et al, 2001). The isovist can be divided into angular segments, which are called bins. Each bin holds a set of locations an agent can see, as shown in figure 3. It is important to note that this analysis method can only be used on a polygon plan structure, in which agents can walk and look around.

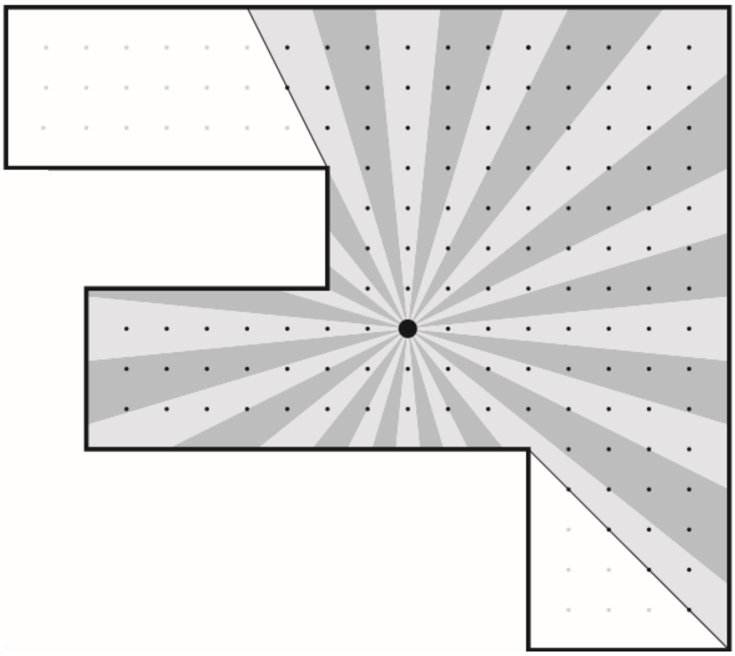


Figure 3. All visible locations on the visibility graph are split into 32 bins, which results in a set of viewable locations divided between 11.25° angular segments. (Turner et al., 2001)

The visibility graph gives the agents a degree of vision. Each bin can store information such as the space syntax measures which the agents can access. By limiting the bins to which the agents have access, it is possible to give them different visual field. In the example mentioned above this is between the 11.25° and 360° (Penn, 2003).

Space syntax research has established that length of line of sight is important to natural movement (B Hillier et al., 1993; Peponis & Zimring, 1990). Using the EVA, a strong correlation has been found between observed human movement and agent movement. When given a visual field of 170° and using a next-step rule (first the agent takes three steps before deciding on its next destination), this correlation is at a maximum of $R^2 = 0.75$ (Penn, 2003). The other 25% of variance could be explained by other factors, such as goals, motivations and attractors in the environment.

2.3.2 Modelling pedestrian volume distribution

Another attempt to implement space syntax measures in an ABM was done in 2017 by Omer and Kaplan. The model was developed to estimate pedestrian volume distribution over the network when considering the street network's structure and agents using an origin-destination pair (Omer & Kaplan, 2017). The model was constructed by transforming two main components of space syntax modelling into agent's movement choice options:

1. *Scale/radius for movement*: this represents the maximum distance available for which movement between origin-destination pairs can be defined. By choosing the shortest route between an origin and destination, the agents determine the pedestrian volume for each street segment and therefore measure the number of times a segment is used on the shortest path. This number is equivalent to the space syntax measure of choice.

2. *Distance type*: Three types of agents were defined based on space syntax modelling, each with their own distance type: metric, topological and angular. Each agent uses their distance type to choose the shortest path between origin-destination pairs. The three different distance types can be defined as follows:
    a. *Metric (least length):* The distance cost of routes is measured as the sum of the segments lengths, defining length as the metric distance along the lines of two adjacent segments (Hillier & Iida, 2005).
    b. *Fewest turns (topological):* The distance cost is measured as the number of changes of direction that are taken on a route (Hillier & Iida, 2005).
    c. *Least angle change (geometric):* The distance cost is measured as the sum of angular changes that are made on a route. This is done by assigning a weight to each turn based on the angle of incidence (Hillier & Iida, 2005).

The resulting ABM therefore enables simulation of movement potentials at different scales and of different distance types. The movement potential will be different with each distance type (Omer & Kaplan, 2018). As a third component, Omer and Kaplan assumed that the movement destinations were affected by the relative weight of different land uses. This was done by classifying the non-residential land-uses into different categories and estimating the relative weight of each category within a the range 0.1-1.0, where 1.0 indicates maximum attractiveness (Omer & Kaplan, 2017).

As a result, they found that the ABM gave better results compared to the space syntax model when it came to the predictability of movement flows. The ABM is better at estimating movement flows even when the connection between street network structure, land-uses and pedestrian movement is low. In this respect, the proposed ABM aligns itself with the core concept of natural movement, which is dynamic in nature.

2.3.3. Using space syntax measures for predicting traffic flow

Jiang & Jia (2009) did research on which space syntax measure was best correlated with to human movement behavior. They compared the aggregated movement flow of two types of agents (purposive walker and a random walker) with seven different space syntax measures: weighted PageRank, PageRank, connectivity, control, betweenness, local integration and global integration. The correlation between the measures and the movement flow were examined to see how aggregated movement flow can be captured by the measures or to see how the underlying street structure can shape the movement patterns and to check which measures are best used for predicting traffic flow.

They discovered that the movement patterns formed by the two different types of agents were the same and that therefore higher cognitive abilities are not needed to form movement patterns at a collective level. The movement patterns were formed through the interaction between the agents and the underlying street structure and not by the human element. This is opposite to Hillier's view who believed the movement patterns were more formed by human cognitive abilities (Hillier et al., 1993). Furthermore, they found that PageRank (R^2 = 0.86) and choice (R^2 = 0.86) were better correlated with human movement flow compared with integration (R^2 = 0.82).

2.4 Implications for methodology

As stated by research, the space syntax measures which correlate the best with human movement behaviour are integration and choice, both from the axial map and the angular segment analysis(Omer & Kaplan, 2017; Penn & Turner, 2001; Teklenburg et al., 1993; Alasdair Turner, 2000). These values will be used in an agent-based model as the way agents navigate over the network. While PageRank is a different way to analyse a network, literature states that the PageRank values are better correlated with movement patterns then the integration measure of the space syntax (Jiang & Jia, 2009). Therefore, these values will also be used by the agents to navigate over the network. Afterwards, these values can be compared to each other. The agents will display exploratory behaviour and translate space syntax- and PageRank values into wayfinding decisions using a probability equation.

While an EVA is an interesting approach to improve pedestrian models by giving agents a sense of vision, it diverges from the scope of the research questions, which is to use the space syntax measures for the decisions of the agents. Including EVA would mean that the agents still access the space syntax and PageRank values of the network at each step they take, which would increase the computation time substantially and even impossible for the size of the research area chosen. For a larger network, such as Amsterdam, it would be better for agents to access the space syntax and PageRank values of the network at each junction to make them base their decision on the network characteristics and create movement flow density maps based on their decisions. Therefore, EVA will be excluded from this research.

3 Methodology

This chapter presents the methodology, research design and details of the ABM. In this research we want to implement space syntax and PageRank measures as a way for agent to traverse the street network of Amsterdam and compare and analyse the resulting movement density patterns. The different measures need to be calculated first using a combination of ArcGIS, DepthmapX and Python. After that, the agents will use the measures for their wayfinding over the network. GAMA is used for the creation of the ABM. For the description of the ABM, the ODD protocol is followed (Grimm et al., 2006). The ODD protocol consists of three blocks (overview, design concepts and details) which makes it easier for readers to understand the purpose and focus of the model.

3.1 Concept

The goal of this study is to see how the space syntax measures identified in chapter 2 can be translated into individual agent decisions and implemented into an ABM. The resulting movement patterns should be analysed and compared to see if the effect of the different measures on the patterns. The conceptual approach of this thesis is illustrated in figure 4.

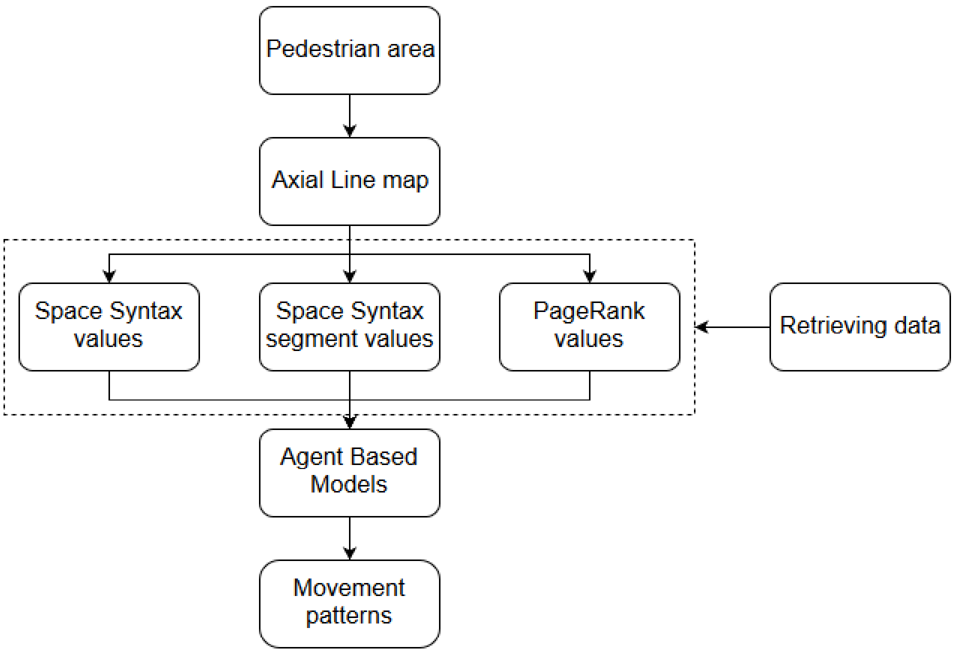


Figure 4. Conceptual approach.

To achieve the goal, the pedestrian area needs to be transformed into an axial line map using the fewest axial line method of DepthmapX. This process is further explained in chapter 3.3. Using the axial line map, three sets of measures are calculated: the space syntax values, the space syntax segmented values and the PageRank values. This step is named the 'retrieving data' step. These values are the basis for the ABM and used by the agents in the decision-making process on junctions with multiple path options. Since the goal is to see how the space syntax and PageRank values and therefore the street network structure can influence the agent's movement behaviour, the agents' goal is to walk around the street network of Amsterdam with no destination other than going for axial line to axial line. This means that the agents display exploratory behaviour. The resulting movement patterns from the ABM will be compared to each other and to a non-preference model. Differences between the models might draw interesting conclusions on the plausibility of the movement patterns and the differences between themselves.

3.2 Study area and data

As a study area the city centre of Amsterdam, including the Museumplein, Vondelpark and the Pijp was chosen. Figure 5 shows the study area. This study area was chosen because it includes many of the popular tourist attractions and shopping

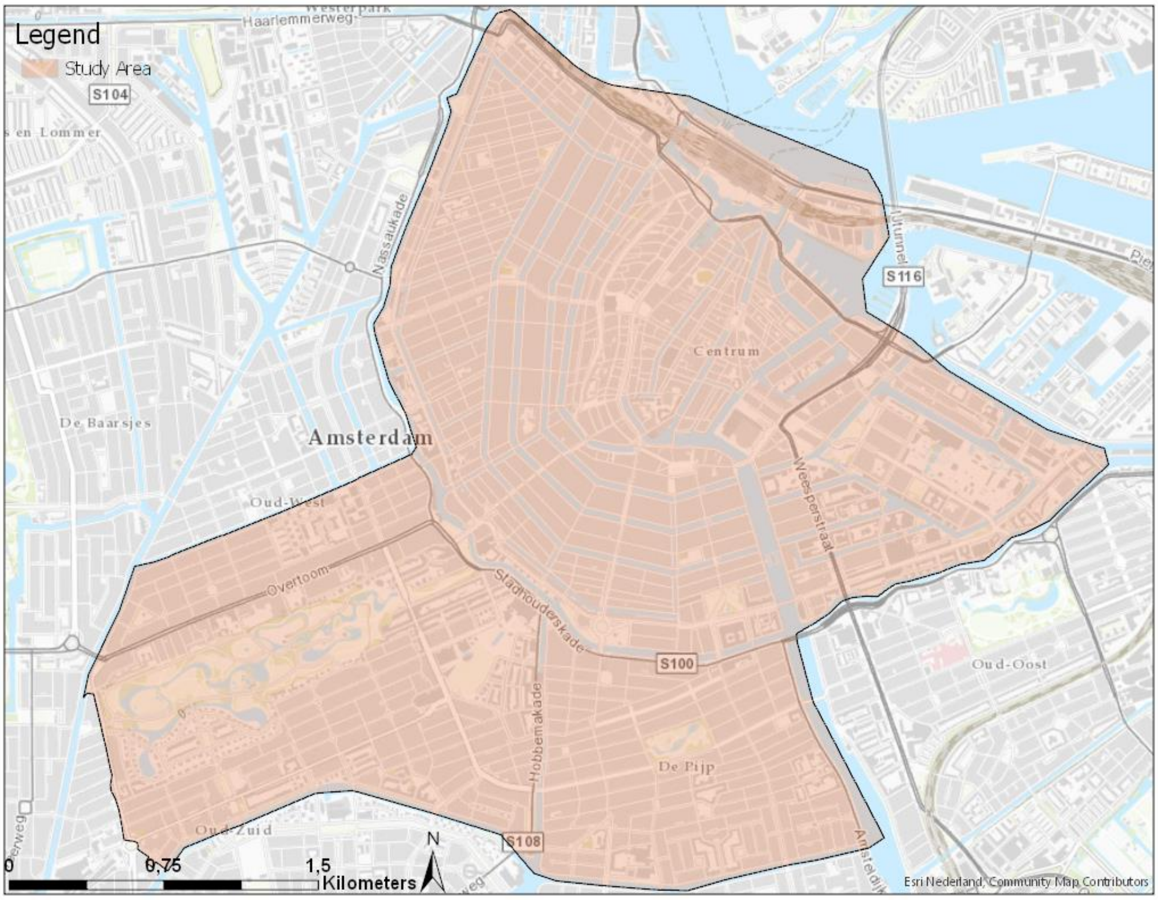


Figure 5. Study area.

streets and the behaviour of exploratory agents corresponds best with tourist behaviour who wants to enjoy and discover a city. The street and building data were retrieved from Basiskaart Grootschalige Topografie (BGT), which is a digital topographical map from the Netherlands. The street data is modified in ArcGIS 10.3 to merge the street polygons into a single polygon which represents the pedestrian walkable area within the study area. This is done to create a geometrically closed polygon of the pedestrian area.

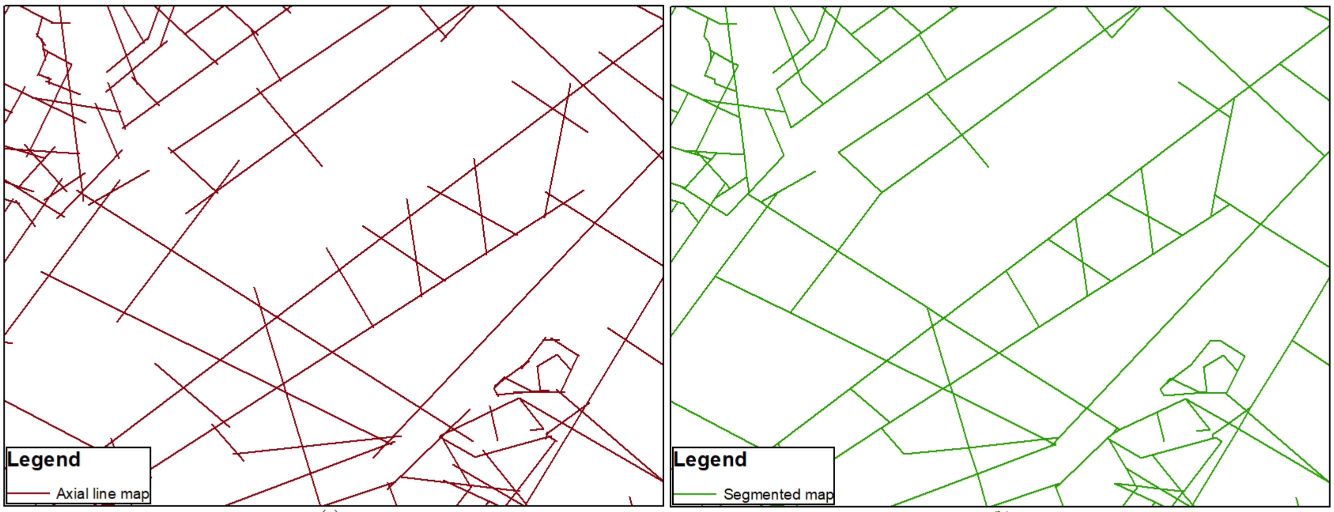
3.3 Space Syntax en PageRank values

3.3.1 Axial map and segmented map

To create the axial map, the program DepthmapX was used (Varoudis, 2018). DepthmapX is a multi-platform software for network analyses designed by Tasos Varoudis. It is a program used for the automatic calculations of the axial map and segmented map.

An all-line-axial-map was created using the geometrically closed polygon of the pedestrian area. The all-line-axial-map is the map with all the possible axial lines within the polygon. This map was then reduced to a fewest-line-map, which represents the axial line map with the fewest axial lines used where every area with the polygon was still covered. Since the study area is too large for the complex computations of DepthmapX, the study area was divided into eighteen parts, using the canals and streets as boundaries. After the calculations were completed and each part had their fewest-line-map created, they were joined together in ArcGIS 10.3 to create the complete axial map for the study area.

To create the segmented map and calculate the space syntax values, the plugin Space Syntax Toolkit was used in QGIS 2.18 (Gil et al., 2015). The plugin is a front-end for the DepthmapX software which is easier to use and more user friendly then DepthmapX itself for the analysis. To create the segmented map, stubs removal was set at 40%. This setting removed a part of an axial line if the part was a dead-end and less than 40% of the total length of the axial line. An example is illustrated in figure 6. After this, the segmented map was manually altered to change double junctions, multiple axial lines in one street and axial lines which add nothing to the network. This was done to improve the movement capabilities of the agents and make sure they would not get stuck in one part of the axial map due to the presence of too many axial lines. An example is illustrated in figure 7. After these pre-processing steps, the ASA was done on the final segmented map using the space syntax plugin in QGIS. The axial space syntax values were joined to this final map using the reference number of each line to create a final map, which includes both the axial map measures and the segmented map measures.



(a) (b)

Figure 6. difference between the axial map (a) and the segmented map (b) with a stub removal of 40%.

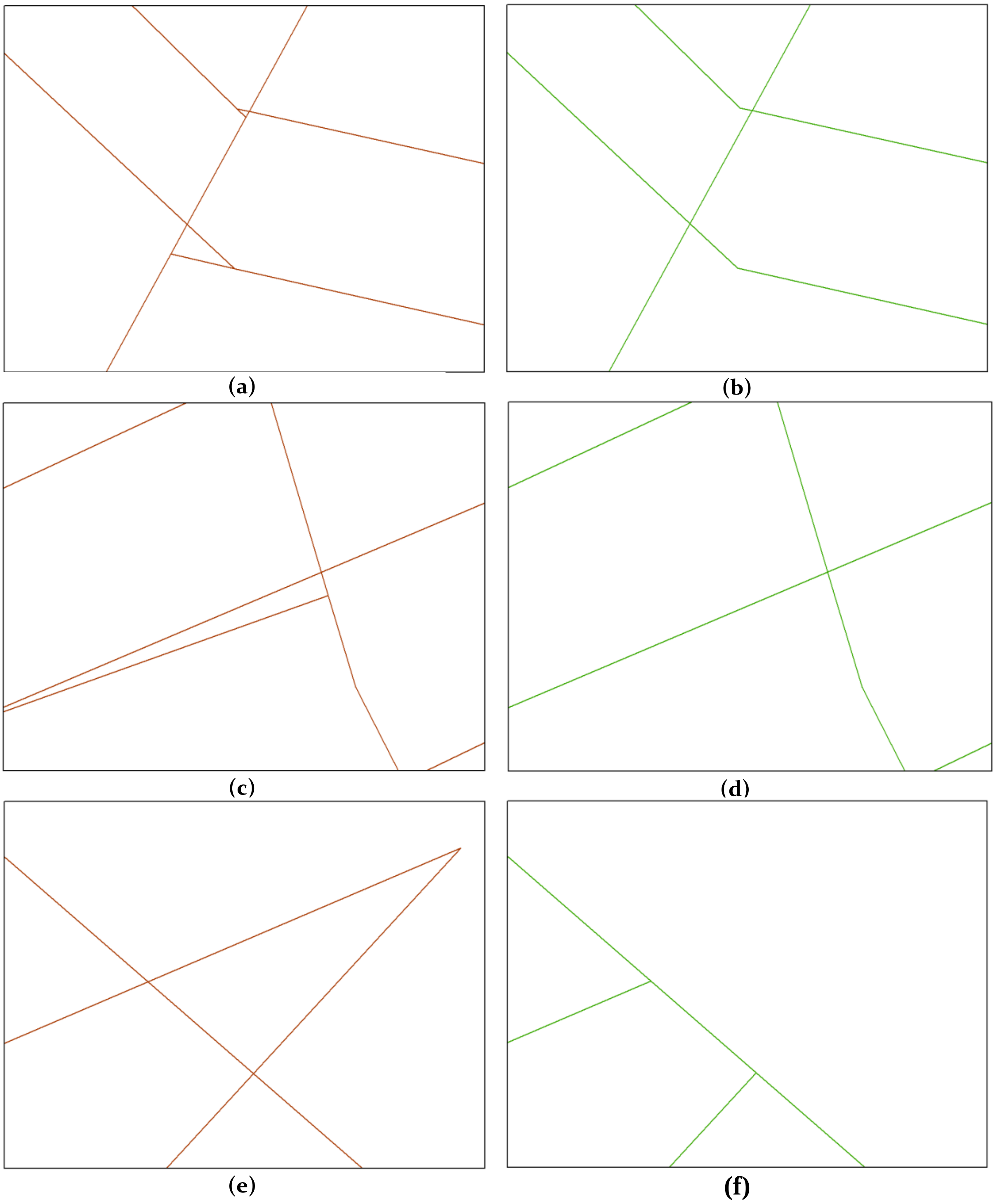


Figure 7. manually altering of the segmented map for easier movement. Before (a) and after (b) of double junctions, before (c) and after (d) of double lines and before (e) and after (f) of removing axial lines who add nothing to the network.

3.3.2 *Space Syntax Values*

Research showed that the integration and choice values of both the axial and the segmented map showed the best correlation with human movement patterns. Therefore, these values were selected. Table 1 shows the full overview of all the selected space syntax values, each with a connection to either integration or choice.

Table 1. Space syntax values selected for the ABM.

Space syntax value	Type of map	Details
RA	Axial map	The relative asymmetry of the axial line. For usability they are inverted.
Integration	Axial map	Integration of the axial line. calculated using: $\frac{1}{RRA}$
Choice	Axial map	The choice of the axial line.
Choice normalised	Axial map	The choice normalised based on length of the axial line. Value between 0 and 1.
Integration	Segmented map	The angular integration of the segmented line. Calculated using: $\frac{n}{MD_{\alpha}}$
Choice	Segmented map	The angular choice of the segmented line.
NACH	Segmented map	The normalised angular choice.
NAIN	Segmented map	The normalised angular integration.

The RA values are ranged from 0 till 1, where high values mean a low integration. The other space syntax values work the other way around: a high value means a high choice or integration. To make sure the RA values are in line with the rest of the space syntax values, the RA values were inverted.

3.3.3 *PageRank values*

The PageRank values are calculated in Python 3.6, using the Fiona library (Gillies, 2016) to load in the segmented map and using NetworkX to create a NetworkX graph, which can be used to calculate the PageRank value for each node in the network. These nodes are then imported into ArcGIS to transfer them into a shapefile. Table 2 shows the PageRank values created.

Table 2. PageRank values selected for the ABM.

PageRank value	Type of map	Details
PR85	Segmented map	PageRank values with damping factor 0.85.
PR99	Segmented map	PageRank values with damping factor 0.99.

Literature states that the higher the damping factor, the higher the correlation with human movement patterns (Jiang, 2008). Therefore, two values are calculated; the default value with a damping factor of 0.85 and one with a higher damping factor of 0.99.

Using these values together with the non-preference walk will result in 11 different model runs.

3.4 Agents

To explore the various space syntax measures three different ABM's need to be made for this research. These models work mostly the same, except for a few differences.

- The first model works with the edge data and uses the space syntax and the segmented space syntax measures.
- The second model uses the PageRank nodes as input values in which the PageRank node values are attached to the corresponding axial line segment.
- The third model is a non-preference ABM. In the non-preference walk no values are used and each road is treated equally.

More details on the differences between these models will be discussed further in this chapter. If no differences are mentioned, then it means that the statements are true for all three models. To avoid terminology confusion, the term 'edge value' is used for the collective of all different space syntax and PageRank values in the models. The ABM's are built using GAMA 1.7. The GAMA code used can be seen in Appendix A.

3.4.1 Overview of the ABM

3.4.1.1 Purpose

The purpose of the model is to gather movement patterns in the city of Amsterdam and to compare the effect of the different edge values on these patterns. The simulated agents depict tourists who are interested in sightseeing the city with no goal in mind, therefore displaying exploratory behaviour. This is important, since we use the network characteristics for the wayfinding of the agents. Would the agents have a destination programmed, models using another wayfinding algorithm, such as the shortest path algorithm, would prove to be superior. To see the impact of the network characteristics on the movement behaviour of the agents, it is important that their decisions are only based on these characteristics, and therefore they should walk around using exploratory behaviour.

3.4.1.2 State variables and scales

The model is created using both the individual tourists and the simulated environment. The size of the simulated is equal to the study area depicted in figure 5. Included in the simulated area are the buildings and the roads. These are static and modelled as agents in the model. The buildings have the following variables: function, which shows if it is an Airbnb or not, the maximum number of tenants and the current number of tenants. The roads have the space syntax values (figure 1) as variables. In the PageRank and the non-preference model no values are attached to the roads. In the PageRank model the PageRank nodes are included in the simulated area. The road agent also has a variable called times walked on, which states how many times the road was visited by the tourist agents.

State variables of the individual tourist agents are: starting location, walking speed, walking time, current node, next node and a list of walked roads.

The timestep is set at 10 seconds. This translates to 360 ticks per hour. This timestep is high enough for a fast-enough run but also makes sure that small road segments are not skipped by the agents which would happen with a higher timestep.

3.4.1.3 Process overview and scheduling

When the model is initialised, each agent receives a walking speed, a walking time and a starting location. When an agent is in the starting location, he finds the closest junction (from now on called node) and walk towards it over the street network. When the agent arrives at the node, he gathers all connecting streets (from now on called edge) as new possible destinations. The agent retrieves the corresponding edge value of each edge and calculates the probability of that edge being the next walking direction. This probability is based on the values of all possible edges. The higher the value, the higher the probability. When an edge is picked, the agent walks towards the next node and adds the edge to the list of walked roads. This process is repeated at each node. To avoid backtracking, the agent uses the walked roads list at each node to see if the agent already visited one of the edges. If so, the value of this edge gets decreased by a weight factor. When an edge is picked by a tourist agent, the variable times walked on of the road agents gets incremented by 1.

Time is modelled using discrete timesteps. After the current runtime of the model reaches the walking time of an agent, it dies. After all the agents are done, the model stops. A flow diagram of the model processes is shown in figure 8.

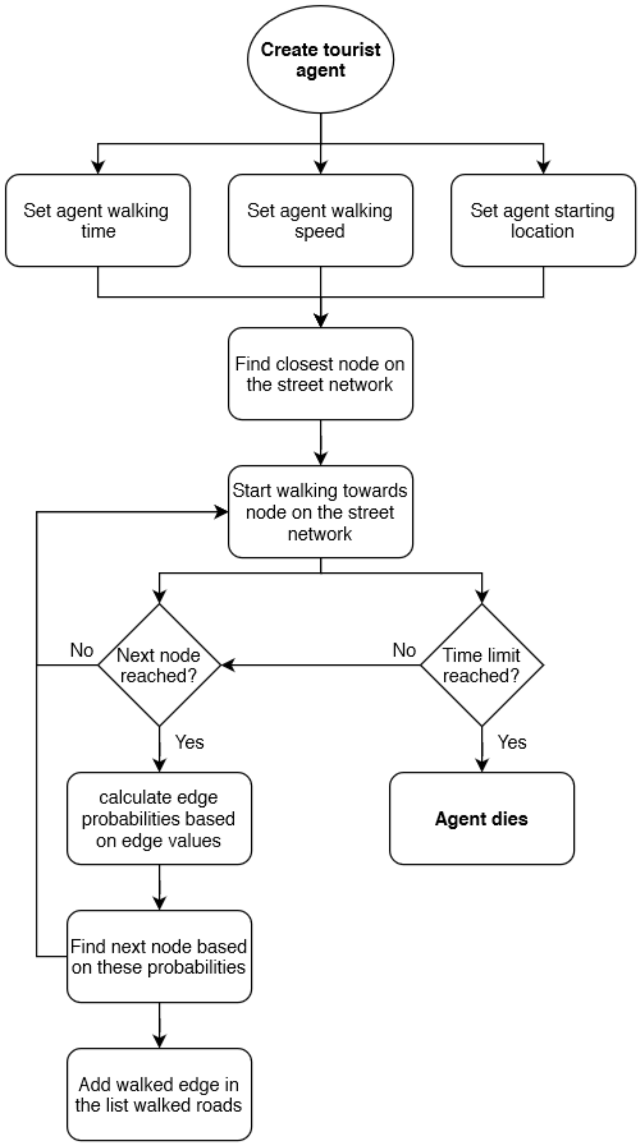


Figure 8. flow diagram of the model processes.

3.4.2 Design concepts

The models used in this thesis are not complicated because the goal of the model is to see how the edge values influence the movement patterns. The individual agents do not need to interact with each other, adapt or change their behaviour over time or create new generations. Therefore, many of the design concepts explained in the

ODD protocol do not apply to this research. Those who do apply, are described below.

Interaction: While the individual tourist agents do not interact with each other, they do interact with the road network. At each decision point, they ask the assigned values for each possible edge and use those to decide on their next path to follow.

Stochasticity: The values of the simulated environment are always the same, however, there is randomness involved in the starting locations of the agents. The agents also use a probability to decide their next path, which means that higher values tend to be picked more, but there is always a change a lower value edge is picked. The walking speed en walking time is also picked random out of a range of numbers.

Observation: The model results observed are how many times each road is visited, which then could be used to create movement density maps of the research area and compare the different values. The perspective of the observation is omniscient.

3.4.3 Details

3.4.3.1 Initialization

A model run starts by assigning each agent a walking speed, walking time and starting location. The speed is a float value between 2 and 5 km/h. The walking time is a float value between 2 and 4 hours. The starting location is randomly picked out of a list of possible starting locations, which is a list of buildings with the function Airbnb. Once a starting location is picked, the current number of tenants of that building is incremented by 1. If the number of tenants is equal to the maximum number of tenants, the building is removed from the list of possible starting locations.

In each run 5000 individual tourist agents are used (N=5000). Each space syntax and PageRank value are run 10 times. The non-preference model is also run 10 times.

3.4.3.2 Input

The input of the models are the space syntax and PageRank values. How these values are calculated is explained in chapter 3.2, which is based on the theoretical framework described in chapter 2. Another input is the weight factor used to decrease the value of edge already walked upon. This weight factor is explained in chapter 3.4.3.3.

3.4.3.3 Submodels

Types of junctions

To improve calculation time of the model and to avoid as much backtracking as possible, each agent checks at each node what kind of junction it is. An example of each junction is shown in figure 9. If it is a dead end, which is a node with only one edge option, the agent picks that option and adds the edge to the walked roads list with no probability calculations involved.

If it is a node with two edge options, which mostly relate to corners in streets, the agent picks the edge which he did not just came from. An example with figure 9b. If the agent is at N1, and just came from N3, it will automatically pick N2 as its next destination. The edge picked will be added to the walked roads list and no probability calculations are done.

If it is a node with more than 2 options, probability calculations need to be done to decide the agent's next path.

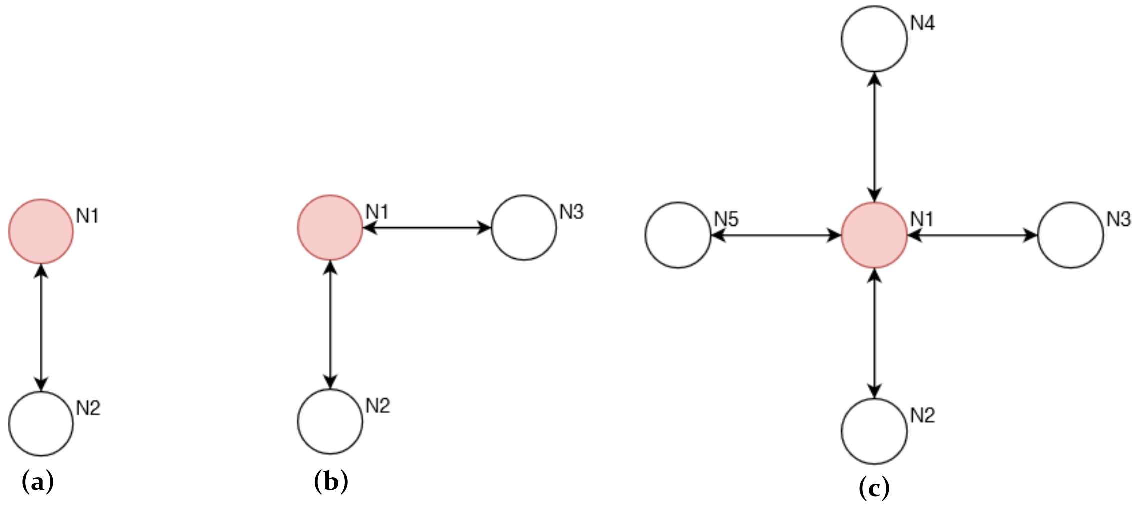


Figure 9. different types of junctions. A dead end (a) with one edge option, a corner (b) with two edge options and a junction (c) with 2 or more edge options.

Edge probability

If an agent arrives at a node with more than 2 edge options, the edge probability is calculated. First the total of all the edge values (E) directly connected to the node is calculated using equation 10:

$$E_1 + E_2 + \ldots + E_n = E_{total} \tag{10}$$

Then, the probability (P) of each edge is calculated using equation 11:

$$\frac{E_1}{E_{total}} = P_1, \frac{E_2}{E_{total}} = P_2, \ldots, \frac{E_n}{E_{total}} = P_n \tag{11}$$

The model then picks an edge based on all the probability values. For example, an edge with a probability of 0.40 has 40% change it will be picked as the next edge the agent will follow.

Weight factor

To avoid backtracking as much as possible, without removing the edge as a possible walking destination, the model uses a weight factor to reduce the value of an edge if the agent has already visited that edge. To see the impact of the weight factor, three different weight factors are introduced. Each value was run with all three weight factors. If an agent arrives at a node with more than 2 edge options, it loops each edge over the walked roads list. The agent counts how many times he visited each edge. The edge value is divided by the weight factor, which in turn reduces the probability of that edge.

The first weight factor (Wf) I named the linear weight factor, which is calculated as shown in equation 12:

$$W_f = nb\_walked + 1 \tag{12}$$

For example, if E1 has a value of 6, and the agent walked 3 times on that edge, the weight factor is 4 which means that the new value of E1 will be 1.5.

The second weight factor is named the double linear weight factor, which is calculated as shown in equation 13:

$$W_f = nb\_walked + nb\_walked \tag{13}$$

In our example this means that the weight factor is 6 (3 + 3), which translates to a new value of E1 of 1.

The third and last weight factor is the exponential weight factor, which is calculated as shown in equation 14:

$$W_f = nb\_walked * nb\_walked \tag{14}$$

In our example this means that the weight factor is 9 (3 * 3), which translated to a new value of E1 of 0.6667.

3.5 Additional model

To see the effects of the edge values on a single agent with a fixed starting location, another model was created which instead of 5000 individual agents with different starting locations and walking times, uses 1 individual agent, with a fixed starting location and a fixed walking time of 4 hours. The starting location used is an Airbnb in the city centre of Amsterdam. Figure 10 shows the starting building. Since the run time of this model is far less than the previous models, this model is run 100 times instead of 10 times to decrease uncertainty. For this model, all three weight factors are used per edge value.

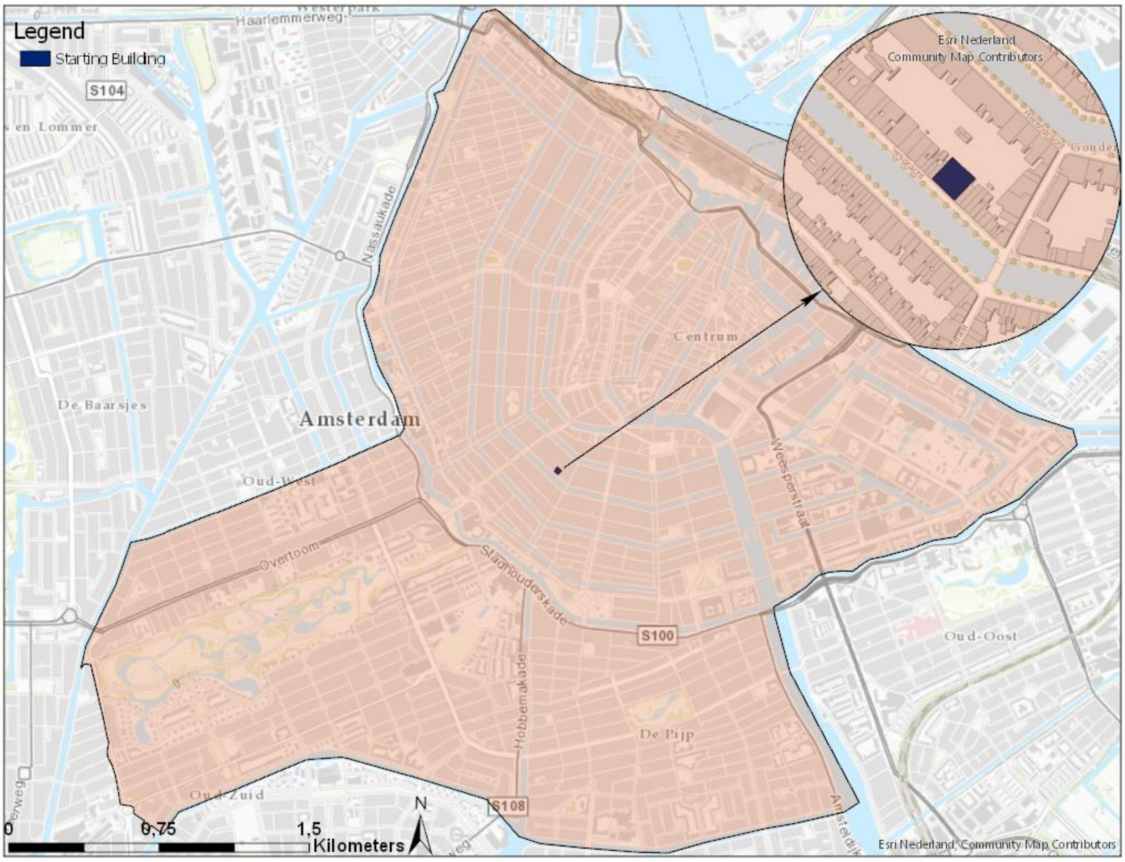


Figure 10. Map showing the starting location of the agent.

3.6 The result design

The resulting movement patterns will be displayed on the study area network using a graduated colour ramp. The values displayed will be the total amount each street is visited.

In total, there are 11 different edge values (8 space syntax values, 2 PageRank values and one non-preference walk). There are also 3 different weight factors (linear, double linear and exponential). Each edge value will be paired up with all three of the different weight factors, resulting in 33 pairs. Each pair is run 10 times.

Since 33 different pedestrian movement patterns will be generated for N = 5000 and 33 different patterns for N = 1, it is key to eliminate some of the resulting movement patterns before analysing the maps. The movement patterns which can be eliminated are the patterns which show hardly to none differences between each other and therefore can be stated that the values produce the same results and can be grouped as one.

First the different weight factors will be compared to select the weight factor which will be used to compare the edge values and therefore reduce the number of movement patterns by three. Then the edge values will be compared using the total sum visited of each axial line

4 Results

In this chapter the results of the thesis are presented. First, the three different weight factors and their impact on the movement patterns are compared to each other and one of them selected. Then, the different movement pattern maps using the selected weight factor are compared. The maps are discussed using their similarities and differences.

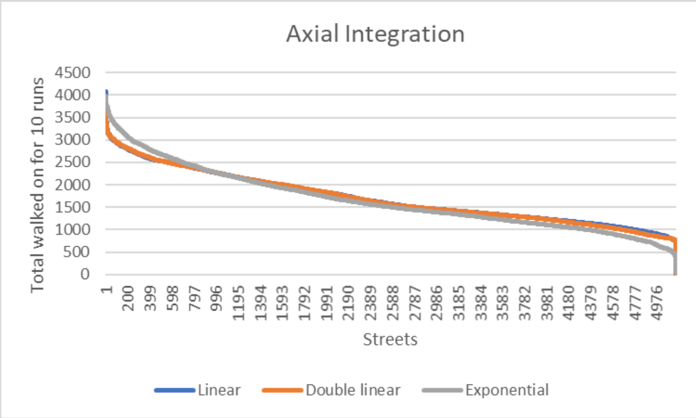
4.1 Weight factor impact

Line graphs are used to compare the weight factors. The weight factors are plotted per edge value, with on the horizontal axis the streets and on the vertical axis how many times each street is visited for the total amount of runs. The streets are sorted from high values to low values. Figure 11 shows some of the line graphs.

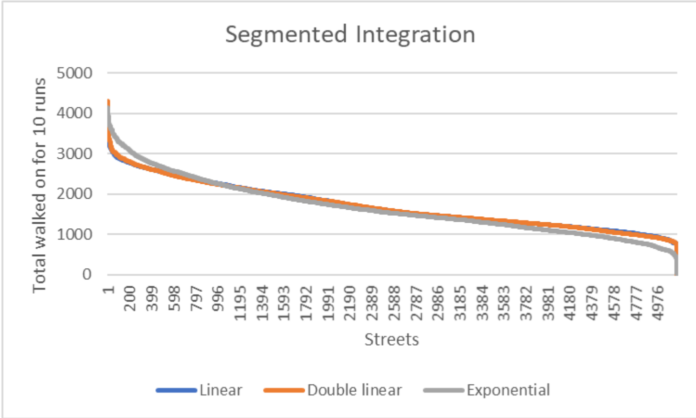
Two different kind of line graphs can be identified among all the edge values. The first and most prominent is an almost linear graph, with two small exponential curves: one at the most visited streets and one at the least visited streets. This graph is the same for almost all edge values, including the PageRank and the non-preference walk. Every road in the network gets visited at least once over 10 runs. There is hardly any difference between the different weight factors, except for a small difference at exponential curves in the beginning and the end.

The second graph is an exponential graph, which only shows when plotting the axial choice or the segmented choice. This curve changes into the almost linear graph when plotting the normalized axial choice or the NACH. In the exponential graph some streets are not visited over 10 runs. There is also hardly any difference between the different weight factors in the exponential graphs. An interesting observation is that a few outliers show which skew the resulting graph. These outliers are in both cases the same two streets (with reference ID 2136 and 2137). With the two streets removed a better overview of the graph curve can be shown (figure 12). The weird thing is that these two streets only display these outliers on the choice edge values, and they do not behave as outliers when the other edge values are used. Since the same model is used for all the values, this is probably not the result of a fault in the model or the network

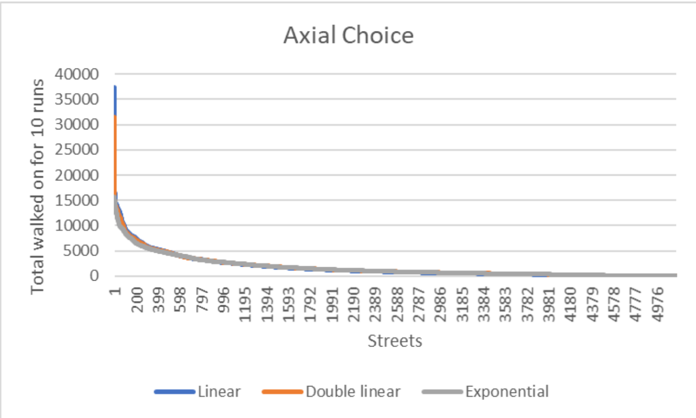
connectivity, but due to an error in how these choice values are calculated in DepthmapX. These two outliers were removed from further results.



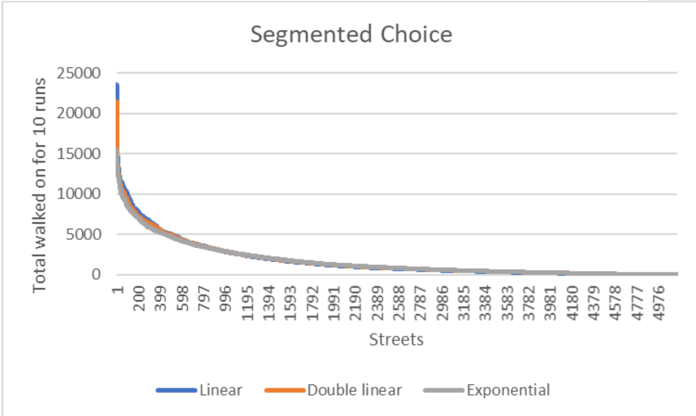
(a)



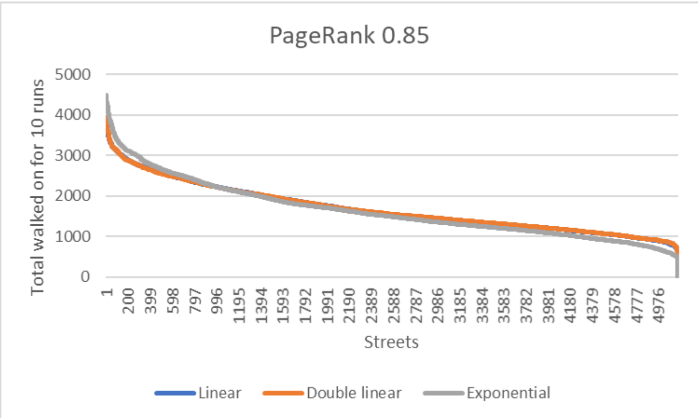
(b)



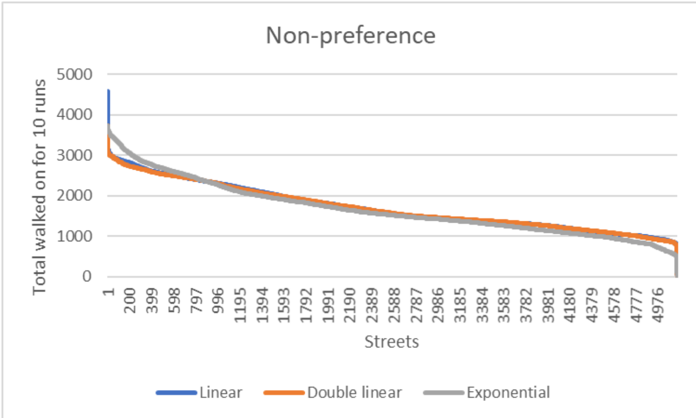
(c)



(d)



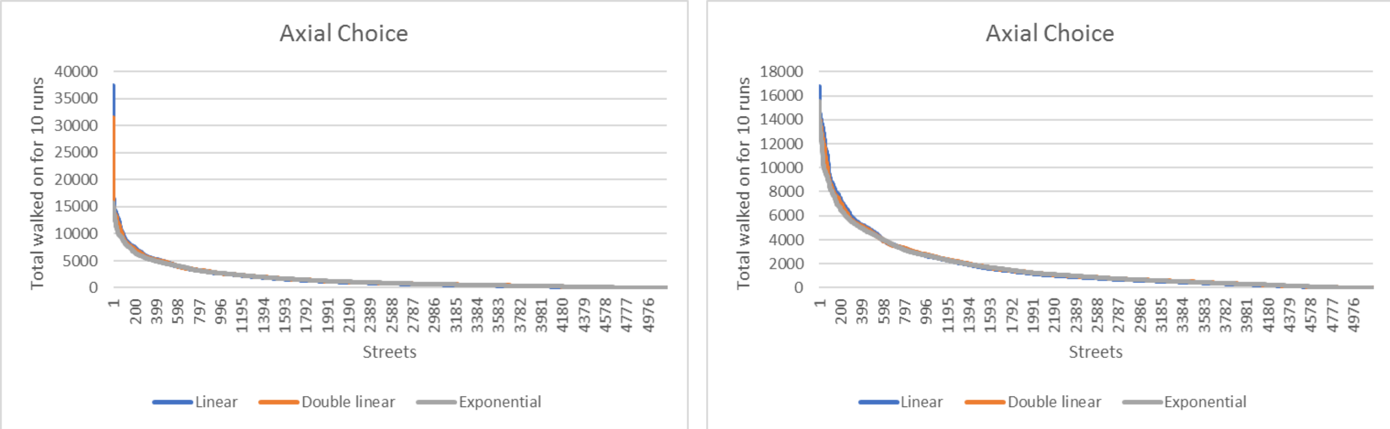
(e)



(f)

Figure 11. 6 different line graphs of the different weight factors. The edge values shown are: axial integration (a), segmented integration (b), axial choice (c), segmented choice (d), PageRank with damping factor 0.85 (e) and the non-preference walk (f).

Since there hardly is any difference between the different weight values and their impact on the movement pattern, the double linear weight factor was chosen to be used as weight factor when comparing the different maps. This method represents the average between all the different weight factor values.



(a) (b)

Figure 12. The difference of the exponential graph of axial choice when the outliers are included (a) and excluded (b).

4.2 Movement pattern maps

The movement pattern maps are created by calculating he number of times each street is visited in each run, creating a sum per street for 10 runs (N = 5000) and 100 runs (N = 1). The maps are made using 20 natural breaks, defined by ArcMap. All the movement pattern maps can be found in appendix B.

4.2.1 N = 5000

Three different types of movement patterns maps can be distinguished by looking at the difference between the patterns. The first is the movement pattern created when using the following values (figure 13):

- Axial choice
- Axial choice normalised
- Segmented choice

The second movement pattern is created when using the following values (figure 14):

- Axial integration
- Axial RA
- Segmented integration
- Segmented NACH
- Segmented NAIN
- Non-preference walk

The third movement pattern is almost the same as the second movement pattern, however, dead-end streets and corner streets are almost unvisited (figure 15):

- PageRank 0,85
- PageRank 0,99

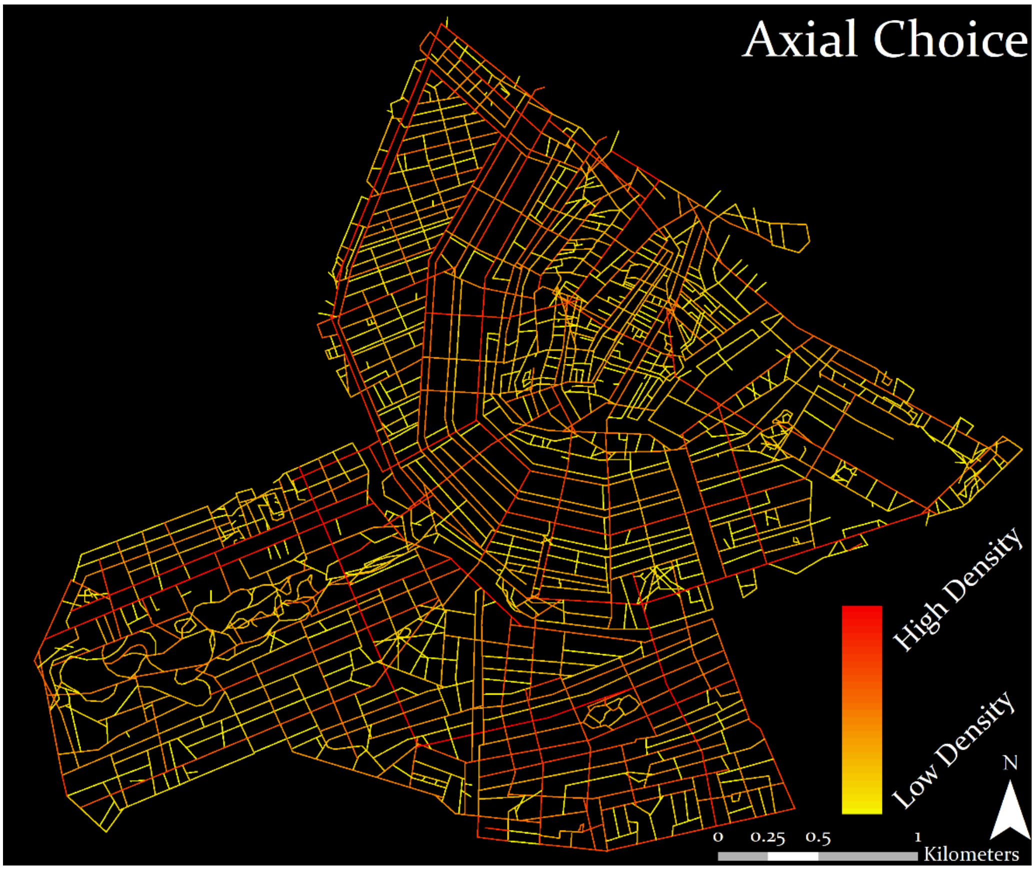


Figure 13. Example of the first movement pattern, the axial choice movement pattern map.

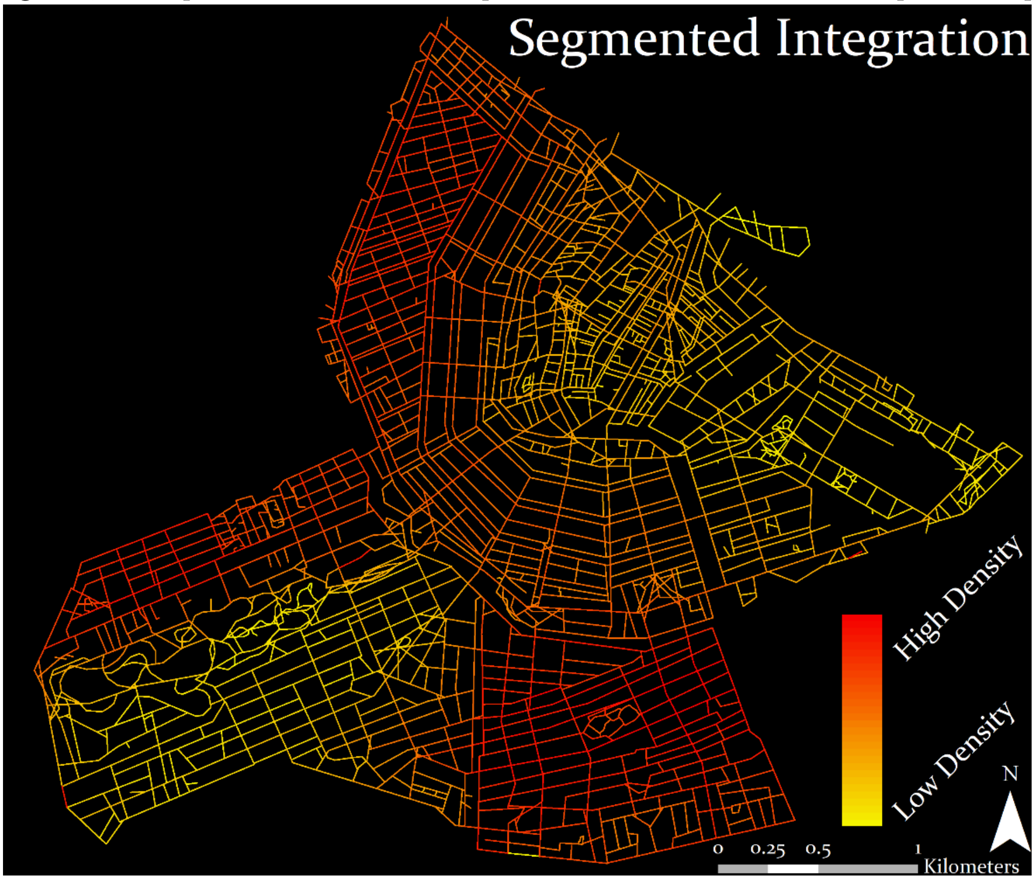


Figure 14. Example of the second movement pattern, the segmented integration movement pattern map.

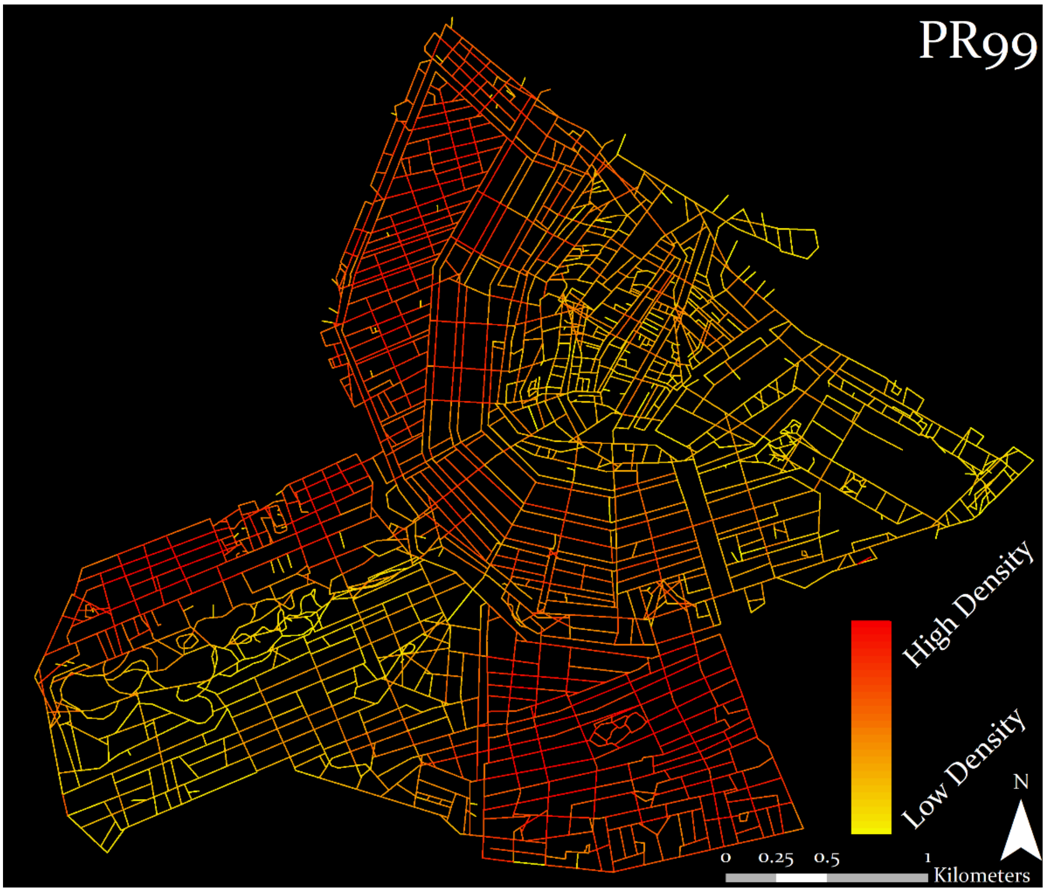


Figure 15. Example of the third movement pattern, the PR99 movement pattern map.

So, what is the reason for these three types of movement patterns? If figure 13 and 14 and their corresponding values are compared, their values can be grouped: (1) the choice space syntax values, which are exponential and (2) the integration space syntax values, including the NACH and the non-preference walk.

The choice values are also seen as the through-movement potential of a street, or the potential a street is passed when going from a random origin to a random destination. In figure 13, the main roads in the study area have the highest number of visitors. The integration values are the to-movement potential of a street or the topological depth. The shallowest lines have the best integration and should be visited the most. However, it seems that the agents are getting stuck in certain neighbourhoods in the study area. This is confirmed by the non-preference movement pattern, which shows the same result as the integration values. The NACH, while in theory a choice value, also shows the same results as the integration values. This is properly because the value is divided by the logarithmic TD, which creates a linear scale instead of an exponential scale. This difference is more obvious when taking the top 10% streets of a choice and an integration value and overlaying the results (figure 16). The segmented integration (blue) has three main neighbourhoods in which the agents walked the most (one to the north-west, one to the west and one to the south). The segmented choice (green) creates a core network which crosses almost the entire study area. Some streets are in the top 10% of both the integration and the choice values.

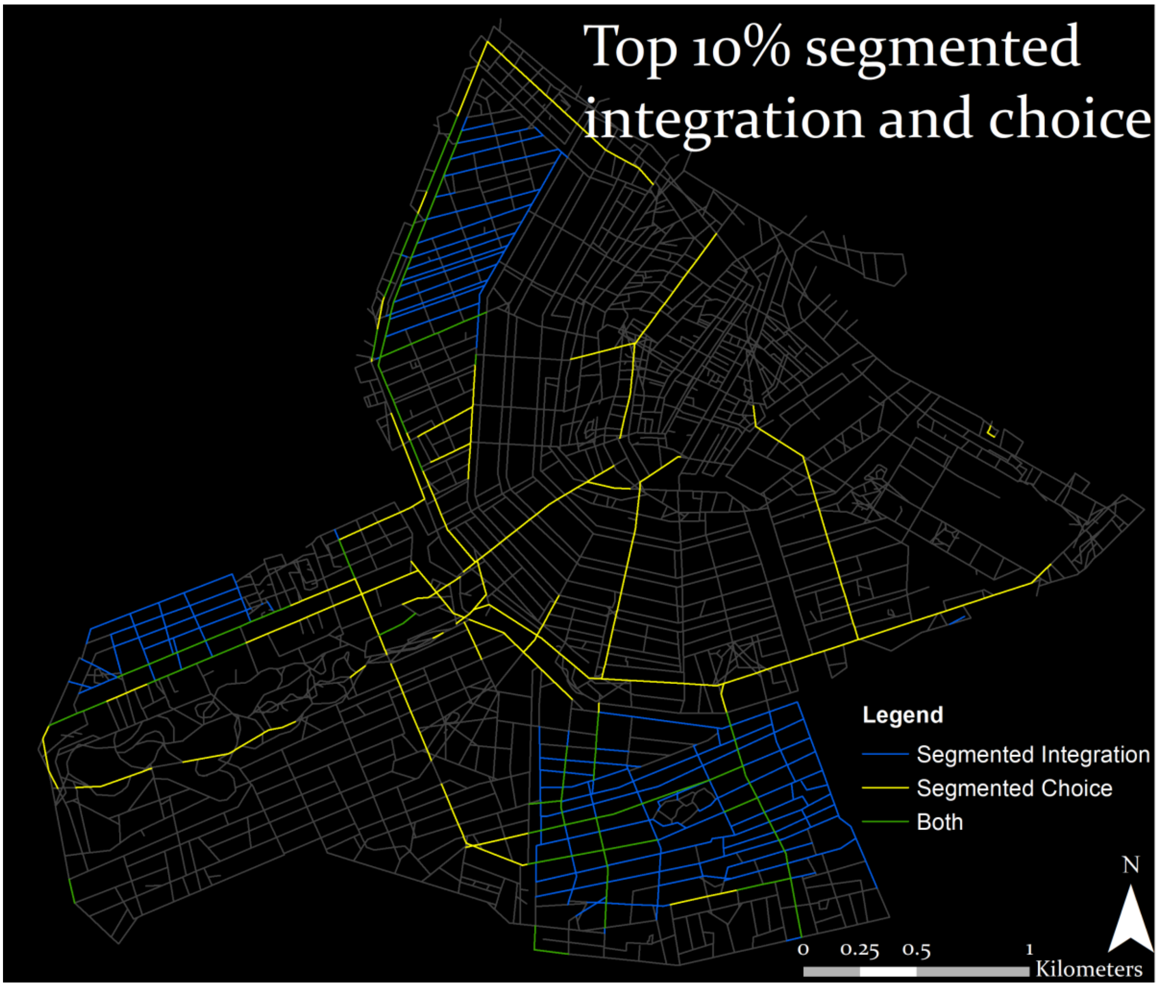


Figure 16. Top 10% visited roads of the segmented integration and the segmented choice.

Figure 16 does reveal some more problems with the space syntax model. Some of the highlighted streets are disconnected with the rest of the network (e.g. bottom left). This is because some errors in the DepthmapX analysis, in which some streets received the wrong values. This caused the agents to prioritize these streets above others.

The reason why the integration values are concentrated in these three neighbourhoods can be explained by looking at what these values represent. The integration represents the streets which are accessible and easy to reach. Within these neighbourhoods, the shallowest lines get the most visitors. Outside of these neighbourhood 'cores', the values are much lower, which creates an uninteresting scenario for most agents to traverse to other parts of the study area. It generates islands within the study area. If instead the top 10% roads inside one of these neighbourhoods is selected, the central road system (most integrated roads) is visible (figure 17). The resulting road network looks more like the choice value road network. This means that integration values should not be used on a large scale, to prevent agents from staying in a small area for the duration of the simulation.

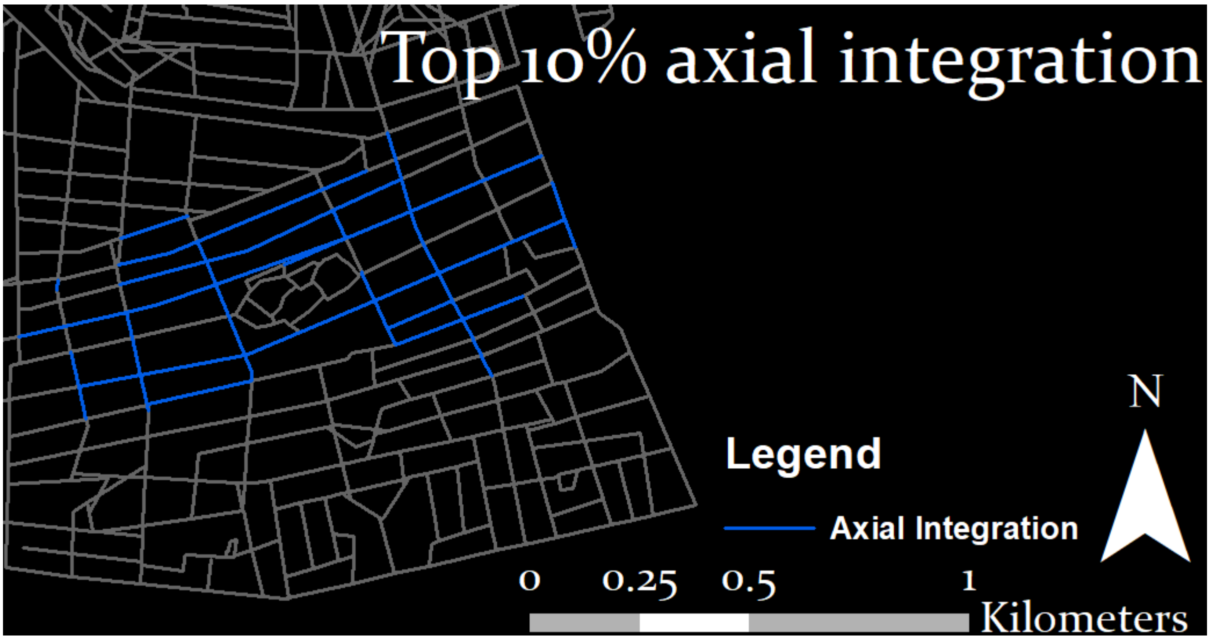


Figure 17. Top 10% of the axial integration roads in one of the neighbourhoods.

The non-preference and integration values movement patterns are corresponding, is because of what the integration values represent. The better integrated, and therefore connected a street, the higher the change this street will be selected by a non-preference walk agent. These agents are mostly located in the same neighbourhoods as the agents using the integration values, as a result of bad connection between these neighbourhoods and the rest of the study area.

The final values not yet discussed are the PageRank values. While their movement pattern maps (figure 15) correspond with the integration value movement patterns, there are some differences. Instead of a gradient change from high density to low density, the PageRank networks represents more of a patchwork, with less visitors in the dead-end and corner streets. This is because of the way these PageRank node values are calculated. Since these types of nodes have a low popularity, the number of visitors is also lower when compared to the integration values.

Resulting, three types of movement pattern maps can be identified using the 11 different edge values and the double linear weight factor. There is no significant difference between using the original space syntax measures and the angular segment analysis and no difference was found between the different integration, choice or PageRank values. This is likely due to the calculated edge values. There were not any substantial differences between the calculated values, which resulted in the agents taking the same routes over the course of a model run.

4.2.2 N = 1

Using one agent in a fixed location instead of 5000 agents in random starting locations, did not deliver unexpected results, except that the PageRank movement patterns now correspond more with the integration movement patterns, resulting in two different maps instead of three. Figure 18a shows the segmented choice map and figure 18b the segmented integration map.

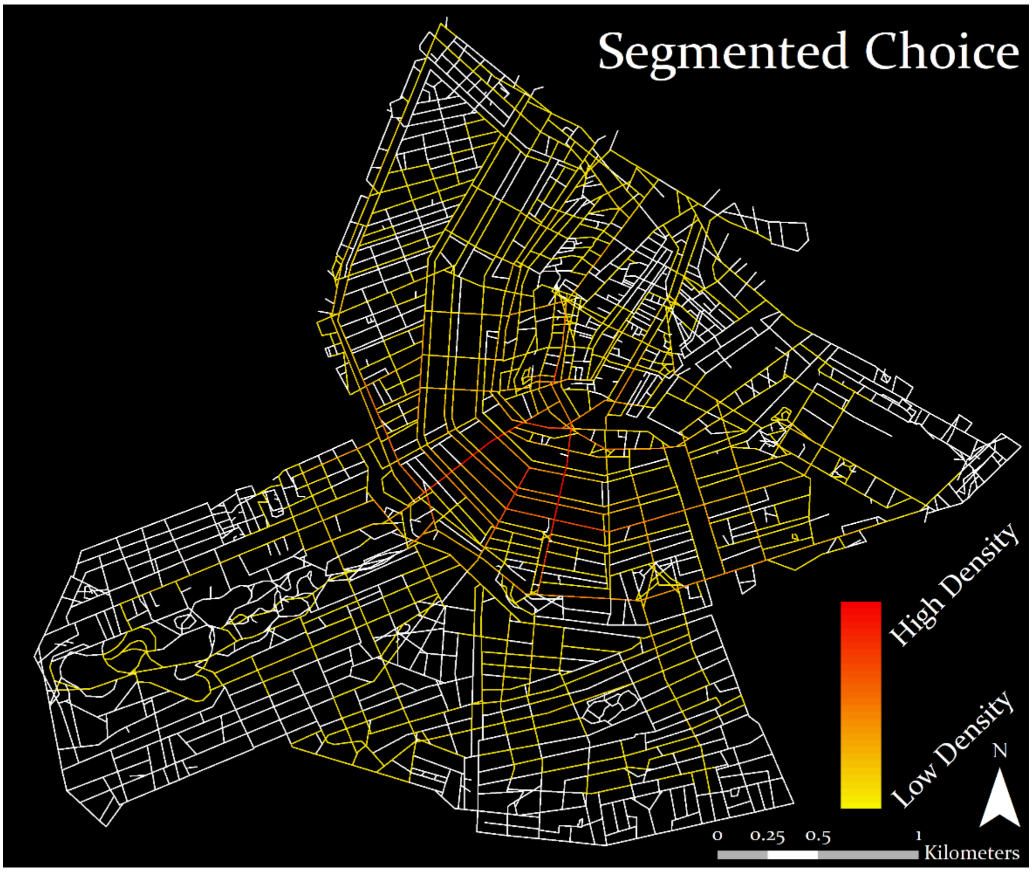


Figure 18a. Segmented choice movement pattern with N=1.

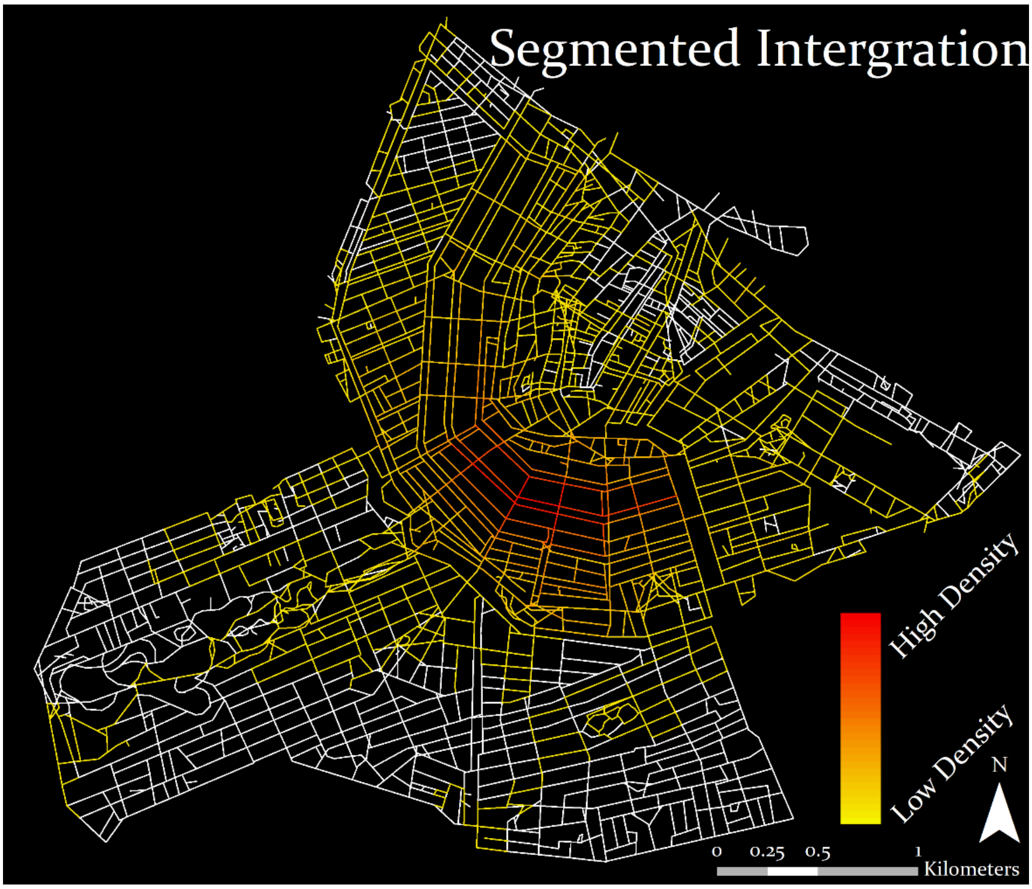


Figure 18b. Segmented movement pattern with N=1.

The choice map shows a street network in which the main streets in the study area have the highest density, and some streets in between have no visitors in 100 runs. The integration map shows a gradient flow from the starting location outwards, with almost all streets in between visited at least once. The reason for this difference in pattern is already discussed in chapter 4.2.1.

4. Conclusion

This thesis explored the potential of space syntax to represent human movement behavior and how the space syntax measures can be used to define an agent-based model. The proposed model was then used in a case study in the city center of Amsterdam.

Which different space syntax theories and analysis models are there that quantify relations between space syntax and human movement behaviour?

The central concept of space syntax is integration, which is also called the to-movement potential and represents the average depth of an axial line from all other axial lines within the system. Another important space syntax measure is choice, which is the through-movement potential and represents the extent to which an axial line functions as an intermediate location within the network. Integration (closeness) and choice (betweenness) can be used to predict human movement, with choice being the more intuitive measurement. An improved space syntax method with more regard for human nature is the angular segment analysis, which uses a weight graph based on the angle of connecting axial lines to calculate the space syntax metrics.

Other research states that PageRank values were a better correlation with movement flows than integration values. PageRank is the technology behind search engines which decides relevance and importance of web pages. This technology can also be used on a space syntax network, where the nodes between axial lines are the web pages. Pedestrians will give a higher priority to nodes that are better connected within the network. In other words, these values are comparable to the connectivity of the axial lines.

How can the relation between space syntax and human movement behaviour be translated into individual agent decisions?

To translate the relation between space syntax and human movement behaviour, the first thing that should be done is transform the pedestrian area to an axial line map. In this case, the city center of Amsterdam was used. Using a program called DepthMapX (Varoudis, 2018), a fewest axial lines was created. The segmented map and the space syntax values were created using the Space Syntax Toolbox (Gil et al., 2015). Some alterations should be made to the segmented map (e.g. remove double junctions) to assure a better movement flow for the agents. The PageRank values were calculated using the segmented map, transformed into a node network.

These calculated edge values are used by the agents to make their decisions. While traversing a street segment, the segment will have the same value until an intersection. Here, the agent should decide the next destination based on the edge values alone. The better values (higher integration, choice or PageRank) should have a better change to be picked as the new destination. Since network characteristics are used for the wayfinding of agents, they display exploratory behaviour.

How can the relation defined in question 2 be implemented into an agent-based model?

To implement these values in an agent-based model, three different models were made. (1) A model that works with edge data (space syntax measures), (2) a model that works with nodes (PageRank) and (3) a non-preference model, in which the agent treats all roads equally. Since the goal of the agent-based model is to see how the edge values influence the movement patterns over the network, the agents do not need to interact with each other or adapt their behaviour over time. Their only goal is to explore the city for a certain time and decide based on the edge values at each intersection between two axial segments. However, to avoid an agent getting stuck on an axial line segment with high space syntax values, a weight factor is introduced. The weight factor is used to reduce the edge value when an agent has already walked over it.

What are the effects of different space syntax measures on (simulated) human movement behaviour patterns and how do these compare to each other?

Three different movement pattern maps can be distinguished after running the agent-based models:

- Choice map (axial choice, axial choice normalised, segmented choice).
- Integration map (axial integration, axial RA, segmented integration, segmented NACH, segmented NAIN, non-preference walk).
- PageRank map (PageRank 0.85, PageRank 0.99).

The choice map gives the most logically human movement pattern. The main streets in the study area are also most visited. When selecting the top 10% of streets visited, a framework emerges which crosses the entire study area. There is no difference between using the axial map values or the segmented map values.

The integration map does not give a logical human movement pattern. The agents are mainly found in three neighbourhoods in the study area and do not leave them. This is happening due to poor connectivity between these neighbourhoods and the rest of the network, and not enough meaningful differences between space syntax values to force the agents out. This is supported by the non-preference walk, which gives the same human movement pattern with each edge value being valued the same. Changing the scale of the study area to a neighbourhood level, the results do change, and more relevant movement patterns emerge.

The PageRank map is almost the same integration map, except that axial lines which are dead-ends or have a low connectivity to the rest of the network, are less frequent visited. This results in the more connected streets receiving more visits. There is no difference between using PageRank 0.85 or PageRank 0.99.

In conclusion, the best method to simulate plausible human movement patterns is by using one of the choice map values. The PageRank are indeed a better correlation with human movement than integration, since a more logical map was created using these values. However, PageRank also deals with the same connectivity and scale issues as the integration map. The integration map does not show any plausible human movement patterns and a non-preference walk will give the same results.

6. Discussion

6.1 Agent behaviour

In this study, it is assumed that the agents who traverse the space syntax network have (1) no previous knowledge about the area, (2) uses the network characteristics (space syntax measures) for its path finding and (3) has no destination. In this case, a very simple, exploratory agent. However, when the resulting movement patterns and comparing them to reality, a few problems arise with this method.

It is first and foremost important to realise that integration, choice and PageRank are correlating with movement due to the way pedestrians move around the system. The literature says there is a correlation between movement and the mentioned values, but not that these values are the causal factor for their movement (Turner, 2001).

The study also ignores other important network characteristics, such as land use, road width and building height. These characteristics are also very important in the way an individual perceives their environment and makes decisions about where to go next. Penn et al. (1998) states that, if these kinds of factors are also considered, there is no doubt that the movement pattern prediction will improve. However, these kinds of values are hard to incorporate, since there is currently no research available on how these characteristics interact and what pedestrians deem important.

The other thing to keep in mind is the cognitive capabilities and knowledge level of humans. If one knows nothing about the network and has no destination, the model used in this study might give insight in the way pedestrians move around the network. However, the network characteristics lose their power as a wayfinding tool the more experience and knowledge a pedestrian gains about the network. Therefore, the model used is only usable when working with exploratory individuals, and in reality, this will never be the case. This method alone is not a good indication for predicting movement flows, however it might improve existing rout-finding algorithms. If the space syntax values, especially the choice values, are integrated as a cost into the weight graph used in rout-finding, it might improve the results.

Most research notes the correlation between the space syntax measures and human movement patterns (Bafna, 2003; Kim & Piao, 2017; Penn & Turner, 2001), but these correlations were found using observed pedestrian flows and then comparing them to the space syntax values. It was almost never used the other way around; using the space syntax values to simulate the pedestrian movement flows. One research of Omar and Kaplan (2017) did create an ABM, but instead of using the space syntax measures directly, they used a model that enabled simulation of movement potentials at different scales and distance types, which in turn represented the centrality measures. While this ABM does aligns itself with the core concepts of space syntax modelling, it does not directly use the space syntax as the wayfinding method of the agents.

6.2 The network

The study area used (Amsterdam) and the resulting movement pattern maps present some questions about the space syntax values used. Using the space syntax values in the agent-based models, results in three different kind of movement pattern maps. One of these, is the result of 6 different values, including the non-preference walk. While literature states these values, all tell different information about the network, it seems like they produce the same result (Teklenburg et al., 1993, Turner 2000, Turner 2007, Hillier et al., 2012). This can have multiple explanations.

The first is that these values are strongly correlated. All the space syntax values use mean or total depth of the axial line in their calculations. This results in hardly any differences in the calculated values for each space syntax measurement. In turn, the agents that use these values in their path decision will arrive at an intersection with each edge having almost the same edge probability (e.g. 31%, 33% and 36%). This will result in the agents acting like they have no preference, since the probabilities are too close together. This explanation is one of the main reasons the movement patterns are similar, however does not explain why the agents end up in the same neighbourhoods 'islands' every run, even though the starting locations are spread around the study area.

The second explanation is the scale of the study area. When the scale is changed to one of these neighbourhoods, the streets with the most visitors when using one of the integration values form a central network pattern, from which the rest of network is reachable. Other studies employing space syntax techniques to evaluate pedestrian movement has found that cities and smaller urban areas may differ in the type of measurements and in the scale needed to best capture the particular movement distributions (Omer & Kaplan, 2017). In this case, the scale of the research might have been too large for substantial differences between the axial and segmented space syntax measures.

It does however explain why certain neighbourhoods get more visitors. If the centre of these neighbourhoods has high integration values, the agents will go there first. After they reach those axial lines, they will stay there until the values are so low due to the weight factor, they will go to the surrounding axial lines. This will continue until the time runs out, resulting in the movement pattern shown in the result. This might have not happened on a smaller scale.

6.3 Software used

Because of the software, the study area had to be divided into eighteen parts, using streets and canals as boundaries. Afterwards, the segmented map had to manually altered to avoid agents getting stuck at certain intersections due to the presence of too many, or double, axial lines. The result was that certain lines did not connect properly in the study area and therefore gave outliers in the result. Sadly, these outliers were found after the analysis was already completed. This working method is not ideal for movement pattern analysis, since you do not want any of these errors in your data. Currently, DepthMapX and the Qgis Space Syntax Toolkit are the only open source tools who can run this kind of analysis. After running the analysis for the city centre of Amsterdam, it can be concluded that these tools are not suitable for an analysis on a larger scale.

6.4 Model runs

There were some time restrictions during the study. Each run of 5000 agents with one of the weight factors and one of the space syntax measurements took between the three or four hours. Therefore, it was chosen to only do 10 runs per measurement and weight factor. This still resulted in a long computation time. However, if possible, the model results would contain less uncertainty if the model was run more times, giving better results.

Another improvement on the model would be to improve the wayfinding. Right now, the agents simply see an axial line as less suitable, but still as an option. An improvement would be to give the agents some kind of 'memory', in which they store a certain amount of axial lines. If one of the options at a decision point is an axial line inside the memory, it gets discarded from the options. This will result in forcing agents to explore the network more and might give more significant results.

7. Recommendations

Based on the previous chapters, I conclude my thesis with recommendations for future work. These include both starting points for future research and improvements on the research done in this thesis.

- The current study area proved not sufficient for significant results of the integration and PageRank values, while research does state these values have a correlation with human movement flows Doing the same research in a different study area might give different results and therefore more insight in what characteristics of a street network are important for an ABM which uses space syntax values. Besides changing the study area, taking a smaller study area might also improve the results since the current scale might have been too large for substantial differences between the axial and segmented space syntax measures.
- This thesis also ignores other important network characteristics, such as land use, road width, building height and landmarks. Research found that these characteristics are also very important in the way an individual perceives their environments and subconsciously considers when deciding on their next destination. If these values were included, there is no doubt that the movement patterns will improve. Research should be done on how these characteristics interact and which are deemed important by pedestrians.
- Integrate the space syntax measures with rout-finding algorithms. While this thesis uses agents with exploratory behaviour, it could also improve the algorithms used to go from origin to destination. Most of these use a weighted graph to calculate the path and then using the least cost path to reach the destination. The space syntax measures could improve these algorithms when they are added as a cost to the weight graph, making some path less desirable to walk on than others.
- Next to network characteristics, other external and internal effects also influence the decision-making process of pedestrians. Examples are speed, safety but also crowds and other traffic. An internal effect could be personal preference for certain roads, landmarks and even sounds. While these are hard to impossible to model in an space syntax ABM, the space syntax measures could be implemented in current decision-making or social processes models, which is a task that is easier to do.

Glossary

Al_Sayed, K., Turner, A., Hillier, B., Iida, S., & Penn, A. (2014). Space Syntax Methodology. *Book*, 1–117. https://doi.org/10.1017/CBO9781107415324.004

Bafna, S. (2003). Space syntax: A brief introduction to its logic and analytical techniques. *Environment and Behavior*, *35*(1), 17–29. https://doi.org/10.1177/0013916502238863

Bonabeau, E. (2002). Agent-based modeling: methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences*, *99*(suppl. 3), 7280–7287. https://doi.org/10.1073/pnas.082080899

Castle, C. J. E., & Crooks, A. T. (2006). Principles and concepts of agent-based modelling for developing geographical simulations. *CASA Working Paper Series*, *110*(0), 60. https://doi.org/ISSN: 1467-1298

Conroy Dalton, R. (2011). The secret is to follow your nose: Route path selection and angularity, (May), 7–11.

Conroy, R. (2001). Spatial navigation in immersive virtual environments.

Crooks, A., Croitoru, A., Lu, X., Wise, S., Irvine, J., & Stefanidis, A. (2015). Walk This Way: Improving Pedestrian Agent-Based Models through Scene Activity Analysis. *ISPRS International Journal of Geo-Information*, *4*(3), 1627–1656. https://doi.org/10.3390/ijgi4031627

Dalton, N. (2001). Fractional Configurational Analysis and a solution to the manhattan problem. *Proceedings of the 3rd International Space Syntax Symposium*, (January 2001), 26:1-13.

Fridman, N., & Kaminka, G. A. (2010). Modeling pedestrian crowd behavior based on a cognitive model of social comparison theory. *Computational and Mathematical Organization Theory*, *16*(4), 348–372. https://doi.org/10.1007/s10588-010-9082-2

Gil, J., Varoudis, T., Karimi, K., & Penn, A. (2015). The space syntax toolkit: Integrating depthmapX and exploratory spatial analysis workflows in QGIS. *Proceedings of the 10th International Space Syntax Symposium*, (July), 148. Retrieved from http://www.sss10.bartlett.ucl.ac.uk/wp-content/uploads/2015/07/SSS10_Proceedings_148.pdf

Gillies, S. (2016). Fiona. Retrieved from http://toblerity.org/fiona/manual.html

Grimm, V., Berger, U., Bastiansen, F., Eliassen, S., Ginot, V., Giske, J., … DeAngelis, D. L. (2006). A standard protocol for describing individual-based and agent-based models. *Ecological Modelling*, *198*(1–2), 115–126. https://doi.org/10.1016/j.ecolmodel.2006.04.023

Helbing, D., & Johansson, A. (2013). Pedestrian, Crowd and Evacuation Dynamics. *Encyclopedia of Complexity and Systems Science*, *16*(4), 6476–6495. https://doi.org/10.1007/978-3-642-04504-2

Hillier, B. (1999). The hidden geometry of deformed grids: Or, why space syntax works, when it looks as though it shouldn't. *Environment and Planning B: Planning and Design*, *26*(2), 169–191. https://doi.org/10.1068/b4125

Hillier, B., & Hanson, J. (1984). *The Social Logic of Space*. Cambridge university press. https://doi.org/https://doi.org/10.1017/CBO9780511597237

Hillier, B., & Iida, S. (2005). Network and psychological effects in urban movement. *International Conference on Spatial Information Theory*, 475–490.

Hillier, B., Penn, A., Hanson, J., Grajewski, T., & Xu, J. (1993). Natural movement: or, configuration and attraction in urban pedestrian movement. *Environment and Planning B: Planning and Design*, *20*, 29–66.

Hillier, B., Yang, T., & Turner, A. (2012). Normalising least angle choice in Depthmap. *Journal of Space Syntax*, *3*(2), 155–193.

Jiang, B. (2008). Ranking spaces for predicting human movement in an urban environment, 1–11.

Jiang, B., & Jia, T. (2009). agent-based simulation of human movement shaped by the underlying street structure, (September), 126–131.

Kim, M., & Piao, G. (2017). A Study on the Applying Concept of Metric Weighting to Space Syntax. *Asian Architecture and Building Engineering*. https://doi.org/10.3130/jaabe.16.447

Macal, C. M. (2016). Everything you need to know about agent-based modelling and simulation. *Journal of Simulation*, *10*(2), 144–156. https://doi.org/10.1057/jos.2016.7

Matthews, R. B., Gilbert, N. G., Roach, A., Polhill, J. G., & Gotts, N. M. (2007). Agent-based land-use models: a review of applications. *Landscape Ecology*, *22*(10), 1447–1459. https://doi.org/10.1007/s10980-007-9135-1

Omer, I., & Kaplan, N. (2017). Using space syntax and agent-based approaches for modeling pedestrian volume at the urban scale. *Computers, Environment and Urban Systems*. https://doi.org/10.1016/j.compenvurbsys.2017.01.007

Omer, I., & Kaplan, N. (2018). Structural properties of the angular and metric street network's centralities and their implications for movement flows. *Environment and Planning B: Urban Analytics and City Science*, 239980831876057. https://doi.org/10.1177/2399808318760571

Penn, A. (2003). Space Syntax And Spatial Cognition. *Environment and Behavior*. https://doi.org/10.1177/0013916502238864

Penn, A., & Turner, A. (2001). Space Syntax Based Agent Simulation.

Peponis, J., & Zimring, C. (1990). Finding The Building In Wayfinding, *22*(5).

Raford, N., & Ragland, D. R. (2005). Pedestrian Volume Modeling for Traffic Safety and Exposure Analysis : The Case of Boston , Massachusetts. *Transportation Research Board 85th Annual Meeting*, *M*.

Teklenburg, J. a F., Timmermans, H. J. P., & Wagenberg, a F. Van. (1993). Space syntax: standardised integration measures and some simulations. *Environment and Planning B: Planning and Design*, *20*(March 1991), 347–357. https://doi.org/10.1068/b200347

Turner, A. (2000). Angular analysis: a method for the quantification of space. *Casa*, 22.

Retrieved from http://discovery.ucl.ac.uk/1368/1/paper23.pdf

Turner, A. (2001). Angular Analysis.

Turner, A. (2007). From axial to road-centre lines: A new representation for space syntax and a new model of route choice for transport network analysis. *Environment and Planning B: Planning and Design*. https://doi.org/10.1068/b32067

Turner, A., Doxa, M., O'Sullivan, D., & Penn, A. (2001). From isovists to visibility graphs: A methodology for the analysis of architectural space. *Environment and Planning B: Planning and Design*, *28*(1), 103–121. https://doi.org/10.1068/b2684

Turner, A., & Penn, A. (2002). Encoding natural movement as an agent-based system: An investigation into human pedestrian behaviour in the built environment. *Environment and Planning B: Planning and Design*, *29*(4), 473–490. https://doi.org/10.1068/b12850

Turner, A., Penn, A., & Hillier, B. (2005). An algorithmic definition of the axial map. *Environment and Planning B: Planning and Design*, *32*(3), 425–444. https://doi.org/10.1068/b31097

Varoudis, T. (2018). DepthmapX. Retrieved from https://varoudis.github.io/depthmapX/

Appendix A. GAMA code

```
/**
* Name: SpaceSytnaxModel
* Author: Simon Veen
* Description: This model checks all possible edges to a list of all the streets walked upon. If a street is already walked upon,
the probability the street is picked is decreased. This is done by dividing the edge weight by a factor of number of times walked upon.
This probability is based on space syntax values*
Date: 04/07/2018
*/

model SpaceSyntaxModel

global {

	/** Name for the pre-fix of the output files */
	string runName <- "default_A";
	/** Load buildings shapefile */
	file shapefile_buildings <-
file("../includes/Buildings_Amsterdam.shp");
	/** Load bounding box shapefile*/
	file shapefile_bbox <- file("../includes/BoudingBox.shp");
	/** load road shapefile, including the space syntax measures */
	file shapefile_roads <- file("../includes/AxialSegmentData.shp");

	/** set the extent of the geometry and create road network graph*/
	geometry shape <- envelope(shapefile_bbox);
	graph road_network;

	/** variables to set the speed of the agents */
	float min_speed <- 2 #km / #h;
	float max_speed <- 5 #km / #h;

	/** variable to set how much time each step takes */
	float step <- 10 #seconds;

	/** variable to see the current time to the minute */
	int current_minute update: (time / #minute);

	/**variable list of all the junctions in the system */
	list junctions;
	/** Variable to set number of agents */
	int nb_toerists <- 1;
	/**variable to count how many toerists are back home */
	int backHome <- 0;

	init {
		write "model started";
		/** create the roads including the space syntax values. */
```

```
create roads from: shapefile_roads with:
[AngConn::int(read("angCONN")),
      SegConn::int(read("SegConn")),
      SegChoice::int(read("SegChoice")),
      SegChoice_len::int(read("SegChoice_")),
      SegInt::float(read("SegInt")),
      SegInt_len::float(read("SegInt_Len")),
      SegTD::float(read("SegTD")),
      SegTD_len::float(read("SegTD_Leng")),
      SegNach::float(read("SegNach")),
      SegNach_len::float(read("SegNach_Le")),
      SegNain::float(read("SegNain")),
      SegNain_len::float(read("SegNain_Le")),
      AxChoice::int(read("AxChoice")),
      AxChoiceNorm::float(read("AxChoiceNo")),
      AxControl::float(read("AxControl")),
      AxIntHH::float(read("AxIntHH")),
      AxIntPV::float(read("AxIntPV")),
      AxIntTK::float(read("AxIntTK")),
      AxMD::float(read("Ax_MeanDep")),
      AxRA::float(read("AxRA")),
      AxRRA::float(read("AxRRA")),
      AxRE::float(read("AxRelativi")),
      AxTD::int(read("AxTD")),
      AxRAInv::float(read("AxRAInvert")),
      AxRRAInv::float(read("AxRRAInver")),
      Oid::int(read("ref"))]{}
/**create a road network graph using the roads and get the junctions */
road_network <- as_edge_graph(roads);
junctions <- road_network.vertices;
/** create the buildings using the buildings shapefile and adding important
variables such as nrPeople and Function*/
create buildings from: shapefile_buildings with:
[type::string(read ("Function")),
      maxPeople::int(read("MaxNrPeopl")),
      nrPeople::int(read("NrPeople")),
      Oid::int(read("RefNumber")),
      totalPeople::int(read("TotalPP"))] {
            if type = "AirBNB" {
                  color <- #slategray ;
            }
      }
/**create a list of buildings with the function airbnb */
list<buildings> airBNB <- buildings where (each.type="AirBNB");
/** create the toerists using the # of toerists, each with a random walking
time, a random speed
 * and a living location using one of the AirBNB buildings and a max set equal
to max AirBNB*/
create toerists number: nb_toerists {
      speed <- min_speed + rnd(max_speed - min_speed);
      list<buildings> possible_locations <- airBNB where
(nrPeople <= maxPeople);
      possible_building <- one_of (possible_locations);
      if possible_building.nrPeople <
possible_building.maxPeople {
            start_location <- possible_building;
            start_location.nrPeople <- start_location.nrPeople
+ 1;}
      else {
```

```
						possible_locations <- possible_locations - 
possible_building;
						start_location <- one_of (possible_locations);
						}
					location <- any_location_in (start_location);
				}
		}
}

species toerists skills:[moving] {
	buildings possible_building <- nil;
	buildings start_location <- nil;
	point target <- nil;
	int walking_time <- rnd(120, 240);
	point current_node <- nil;
	point next_node <- nil;
	path pathToFollow <- nil;
	list<roads> potentialEdges <- nil;
	roads nextRoad <- nil;
	roads roadJustFollowed <- nil;
	list<roads> walkedRoads <- nil;
	rgb color <- #red;
	float distanceWalked <- 0.0;

	/** The walk reflex. will use the closest node as the starting node, after that it will use
	 * the probability of the space syntax values */
	reflex walk when:current_minute < walking_time {
		if current_node = nil {
			target <- junctions closest_to(location);
			do goto target: target on: road_network speed:speed;
			switch location {
				match target {
					/** set current node equal to current location */
					set current_node <- location;
					do determine_next_path;
					}
			}
		}
		else {
			do goto target: next_node on:road_network speed: speed;
			switch location {
				match next_node {
					/** set current node equal to current location */
					set current_node <- location;
					/** update distance walked for the agent */
					distanceWalked <- distanceWalked + 
nextRoad.shape.perimeter;
					/** add the walked road to the list of walked roads */
					walkedRoads <- walkedRoads + 
roadJustFollowed;
					/**increment the variable which shows how many times 
a road is walked on by an agent */
					loop fr over: roads {
						if fr = roadJustFollowed {
							fr.timesWalkedOn <- 
fr.timesWalkedOn + 1;
						}
					}
					do determine_next_path;
```

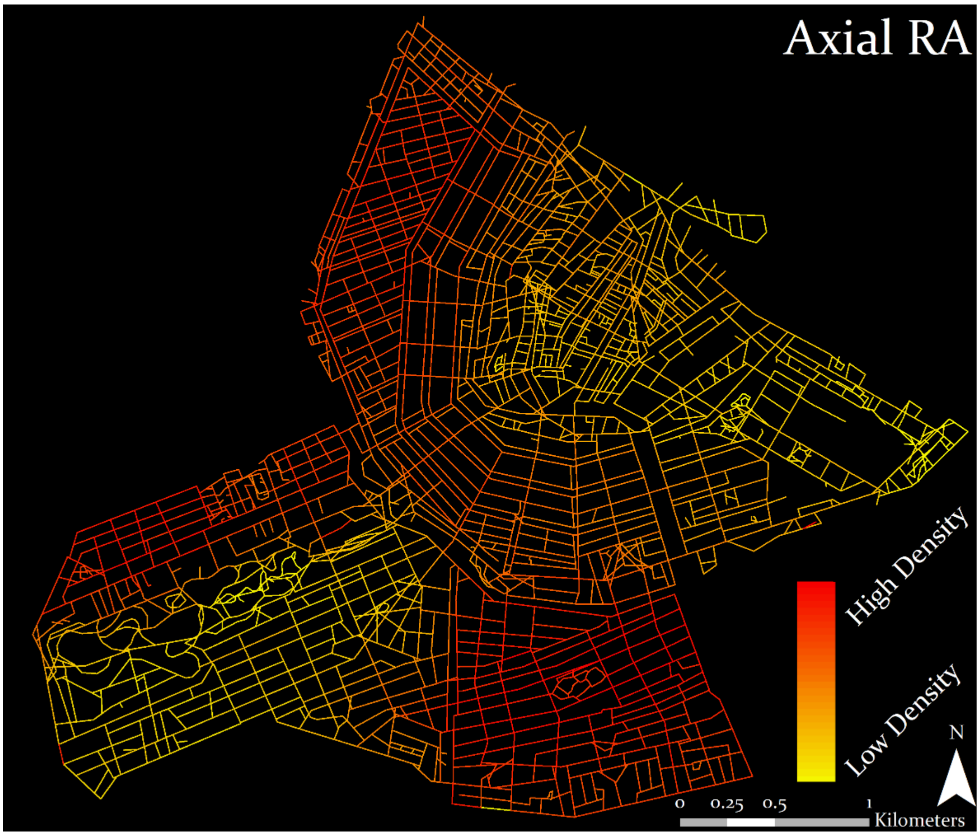
```
                        }
                    }
                }
            }
            /** the go home reflex. will work when the current , is the same as
             * the walking time. The agent will take the shortest route home.*/
            reflex go_home when: current_minute >= walking_time {
                backHome <- backHome + 1;
                do die;
            }

            /** this action will determine which path out of the possible edges the agent will
follow */
            action determine_next_path {
                float totalWeight <- nil;
                list<float> edge_prob <- nil;
                int picked_prob <- nil;
                /** all potential in and out edges are selected */
                potentialEdges <- road_network out_edges_of(current_node) +
road_network in_edges_of(current_node);
                if length(potentialEdges) = 1 {
                    nextRoad <- potentialEdges[0];
                    write "nextRoad is " + nextRoad;
                }
                else if length(potentialEdges) = 2 {
                    if potentialEdges[0] = roadJustFollowed {
                        nextRoad <- potentialEdges[1];
                        write "nextRoad is " + nextRoad;
                    }
                    else {
                        nextRoad <- potentialEdges[0];
                        write "nextRoad is " + nextRoad;
                    }
                }
                else {
                    /** loop to see if the edge is already walked upon. if so, the weight
gets divided by a factor, therefore makes it less likely to be picked */
                    loop i from: 0 to: length(potentialEdges) - 1 {
                        roads potentialEdge <- potentialEdges[i];
                        write "potentialEdge is " + potentialEdge;
                        int nb_walked <- 0;
                        loop walkedRoad over: walkedRoads {
                            if potentialEdge = walkedRoad {
                                nb_walked <- nb_walked + 1;
                            }
                        }
                        /** select the weight of the selected edge. if the edge is already
walked /on, the weight gets divided by a factor. if not,
                         * the weight is equal to the current value of the edge.*/
                        write "nb_walked is " + nb_walked;
                        if nb_walked >= 1 {
                            /** select the factor by which the weight is divided by.
nb_walked + nb_walked.
                             * If you want te factor to increase less, you should
change it to factor <- nb_walked + 1 */
                            int factor <- nb_walked + nb_walked;
```

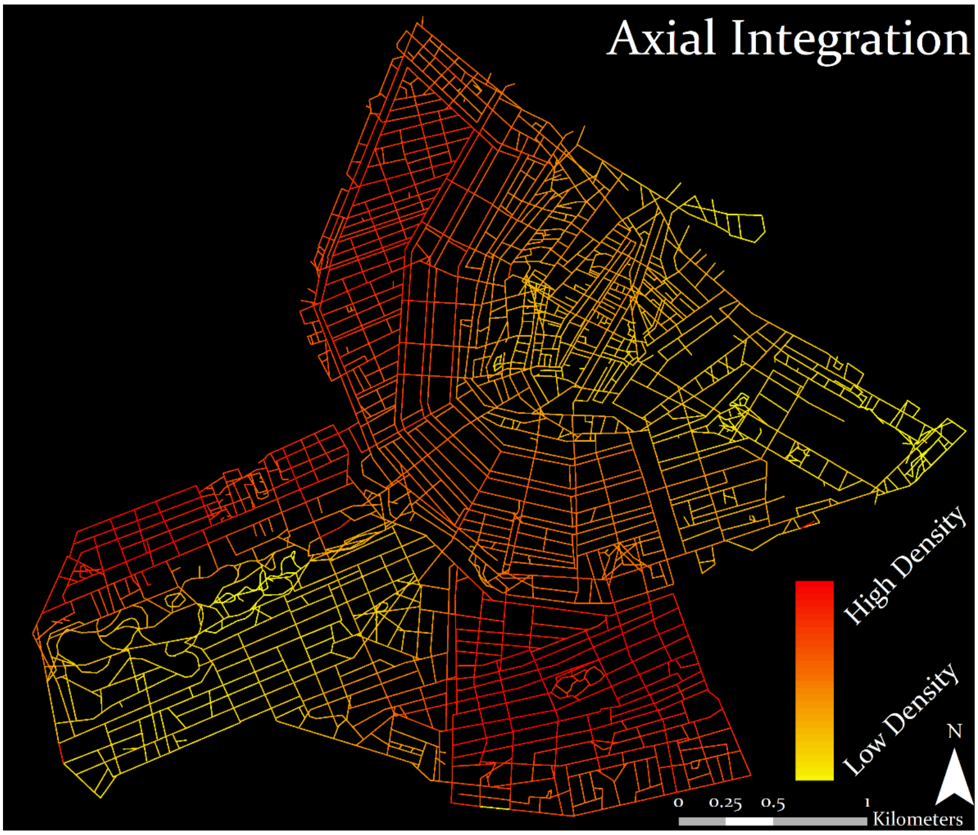
```
                                        potentialEdges[i].weight <-
potentialEdges[i].CV / factor;
                                }
                                else {
                                potentialEdges[i].weight <- potentialEdges[i].CV;
                                }
                                /** calculate the total weight of all the potential edges. */
                                totalWeight <- totalWeight +
potentialEdges[i].weight;
                                }

                        /** the probability of each edge is calculated and put into the list
edge_prob */
                        loop i from: 0 to: length(potentialEdges) - 1 {
                                potentialEdges[i].prob <- (potentialEdges[i].weight
/ totalWeight);

                                edge_prob <- edge_prob + potentialEdges[i].prob;
                        }
                        /** using the operator rnd_choice, one of the probabilities is selected
*/
                        picked_prob <- rnd_choice(edge_prob);
                        /** of the potential edges is selected using the index of the
picked_prob */
                        nextRoad <- potentialEdges[picked_prob];
                        write "nextRoad is " + nextRoad;
                        }
                /** roadJustFollowed is the same variable as nextRoad, however this name
makes it easier to see what the variable means.
                 * the nextRoad variable is still used to determine the next point the agent is
walking towards. roadJustFollowed is used to add to walkedRoads per agent
                 * and determine how many times a road is walked on. */
                roadJustFollowed <- nextRoad;
                /**the road is set to a path and from the two points the next node is selected
*/
                pathToFollow <- path(nextRoad.shape.points);
                /** the next node is selected based on the next road */
                point p1 <- pathToFollow.source;
                point p2 <- pathToFollow.target;
                if p1 = current_node{
                        next_node <- p2;
                }
                else {
                        next_node <- p1;
                }
        }

        aspect base {
                draw circle(2) color: color;
        }
}

species roads {
        int Oid;
        int AngConn;
        int SegConn;
        int SegChoice;
```

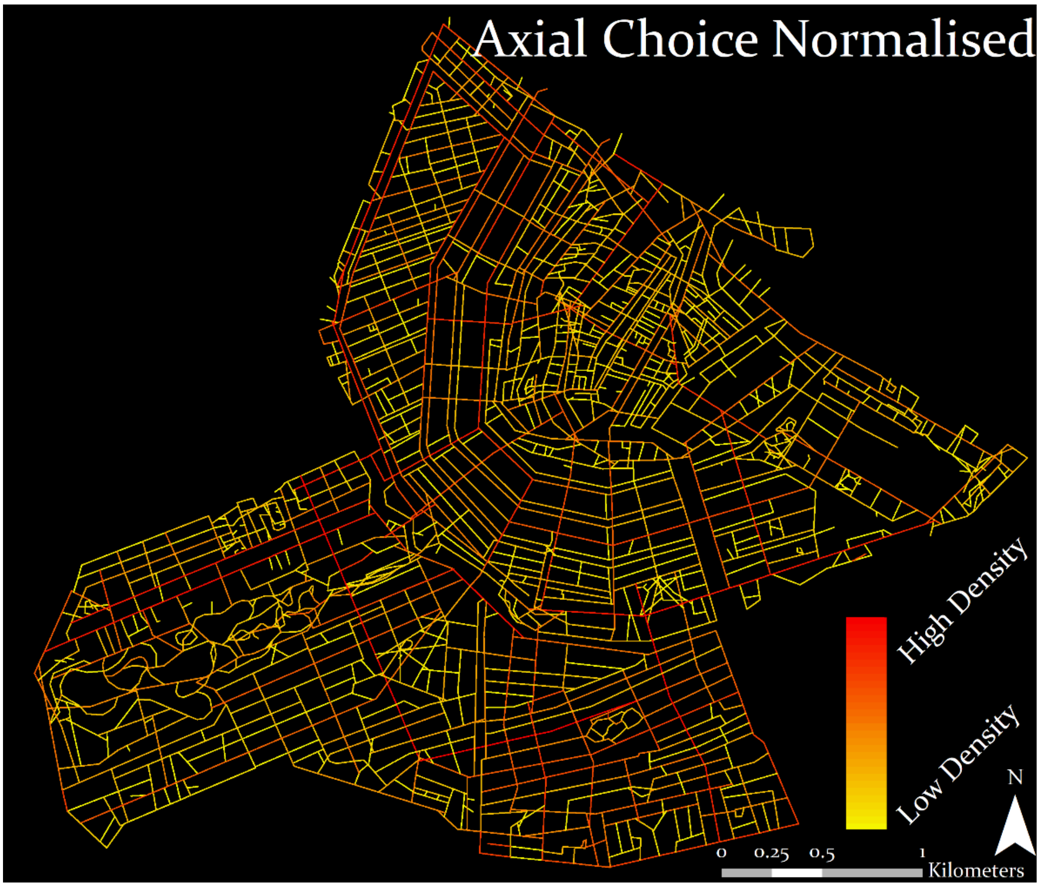
```
        float SegInt;
        float SegTD;
        float SegNach;
        float SegNain;
        int AxChoice;
        float AxChoiceNorm;
        float AxControl;
        float AxIntHH;
        float AxIntPV;
        float AxIntTK;
        float AxMD;
        float AxRA;
        float AxRRA;
        float AxRE;
        int AxTD;
        float AxRAInv;
        float AxRRAInv;
        int SegChoice_len;
        float SegInt_len;
        float SegTD_len;
        float SegNach_len;
        float SegNain_len;
        /** current value used in calculations */
        float CV <- AxIntHH;
        /** variables needed for calculations */
        float prob;
        float weight;
        /**variable to see how many times the road is walked on */
        int timesWalkedOn <- 0;

        rgb color <- #black;
        aspect geom {
                draw shape color: color;
        }

}

species buildings {
        string type;
        int maxPeople;
        int nrPeople;
        int Oid;
        int totalPeople;
        rgb color <- #grey;
        aspect geom {
                draw shape color: color;
        }
}

experiment Myexperiment type: gui {
        parameter "number of toerists" var: nb_toerists category: "toerists"
min: 1 max: 7000;
        output {
                display map type: opengl {
                        species buildings aspect: geom;
                        species roads aspect: geom;
                        species toerists aspect: base;
                        }
         }
}
```

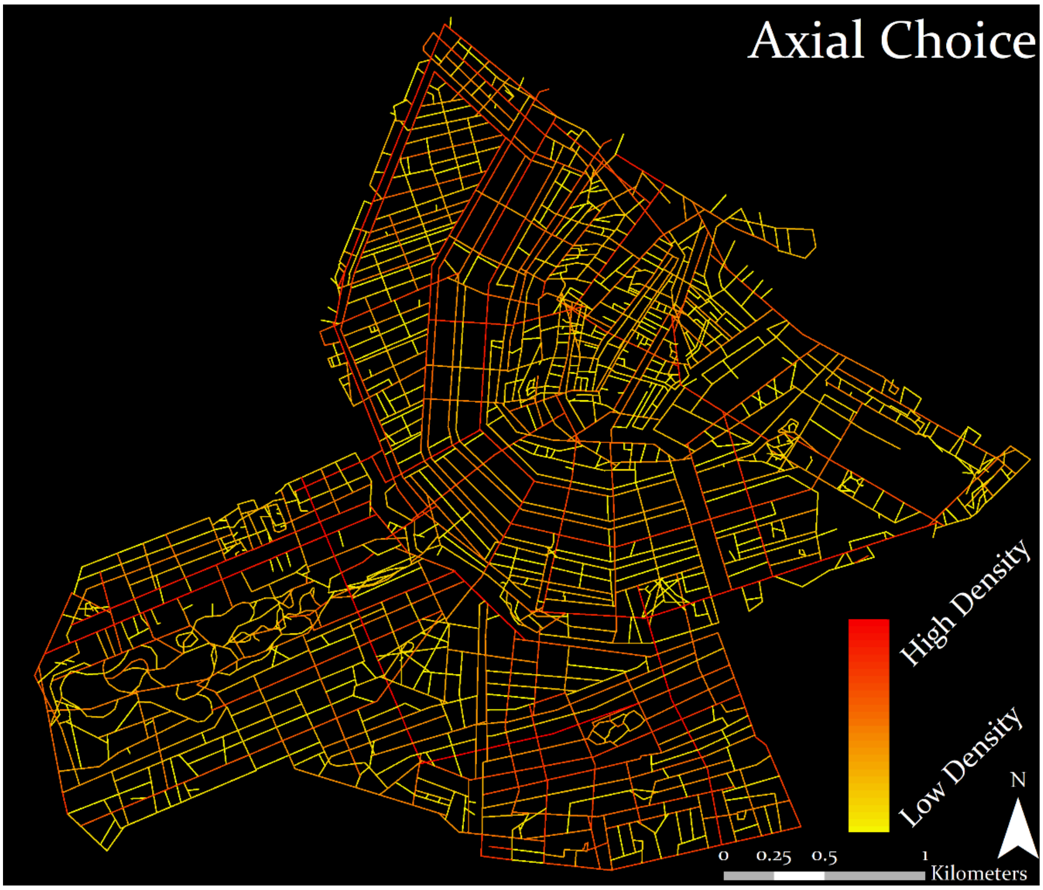
```
experiment 'Batch runs .csv one file' type: batch repeat: 1 keep_seed: true
until: (backHome = nb_toerists) {
      parameter "Batch run name:" var: runName category: "General
parameters";
      parameter "number of toerists" var: nb_toerists category: "toerists"
min: 1 max: 7000;

      reflex end_of_runs {
      ask simulations {
            ask roads {
                  save [Oid, timesWalkedOn] type: "csv" to: "../results/" +
"N_of_" + nb_toerists + "_value_" + runName + ".csv" rewrite: false;
                  }
            }
      }
   }
```

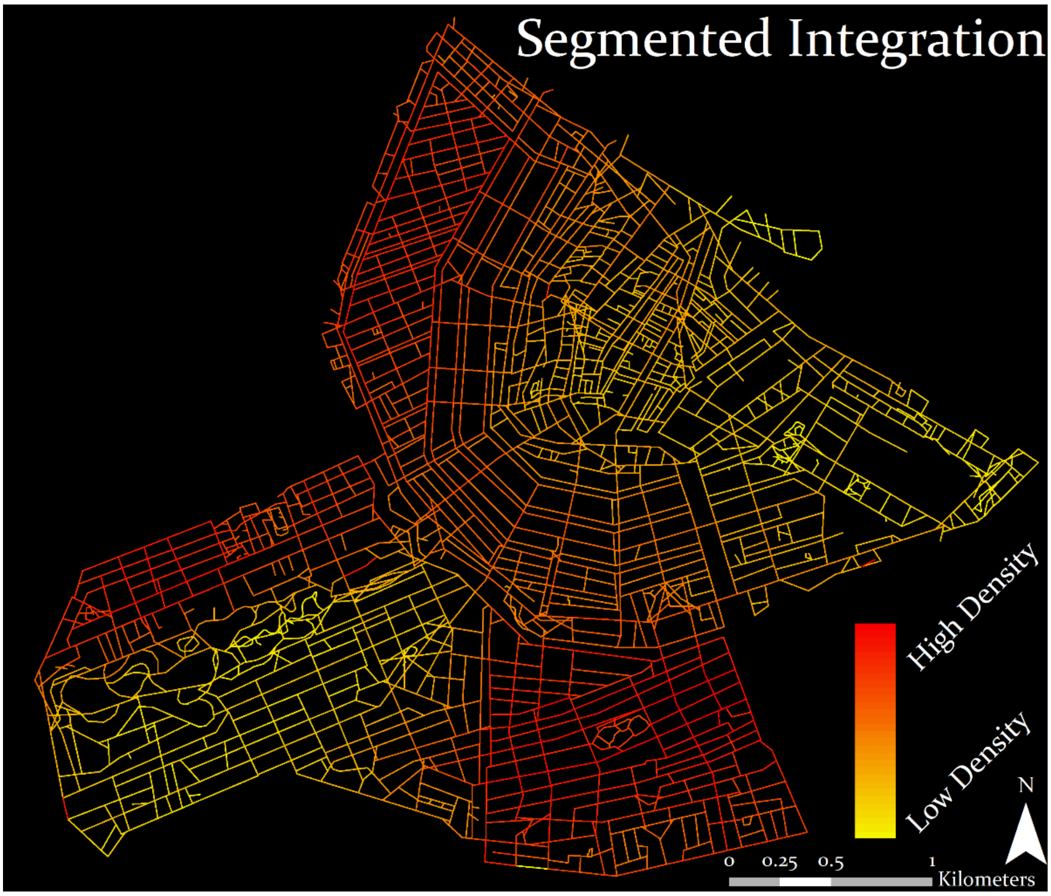
Appendix B. Movement pattern maps.

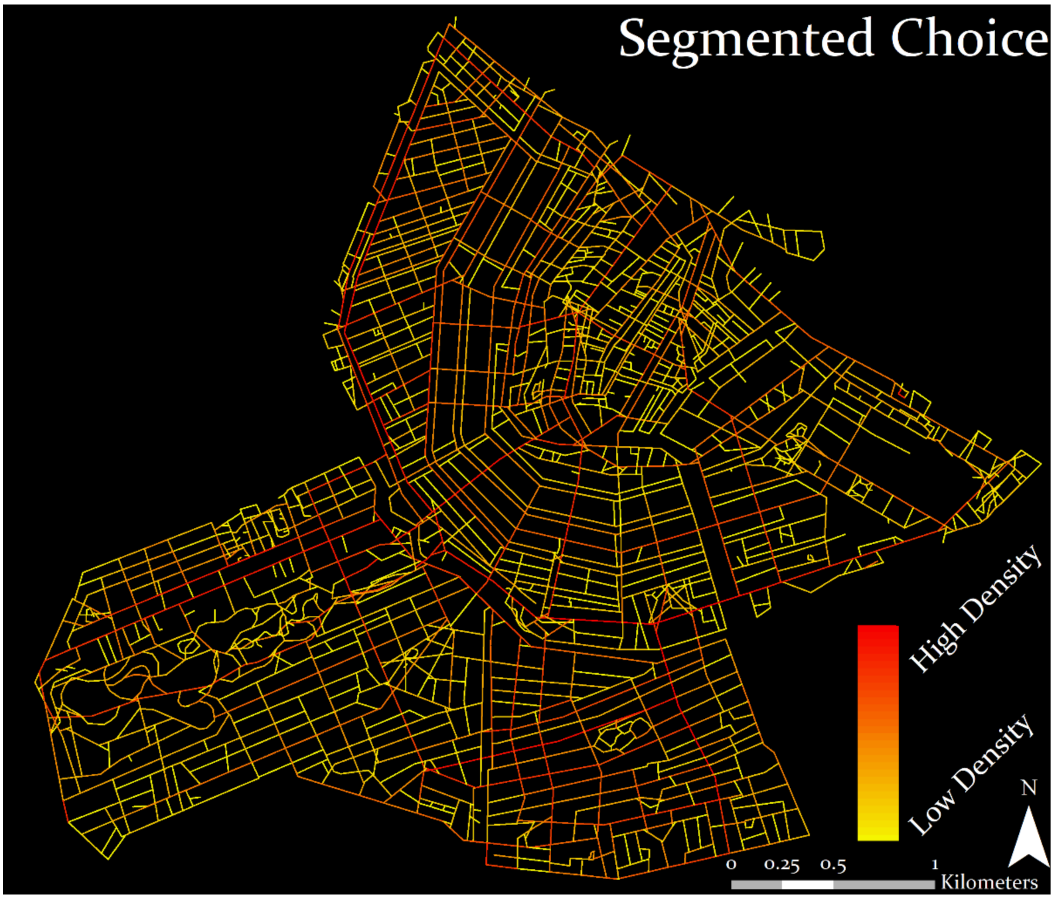


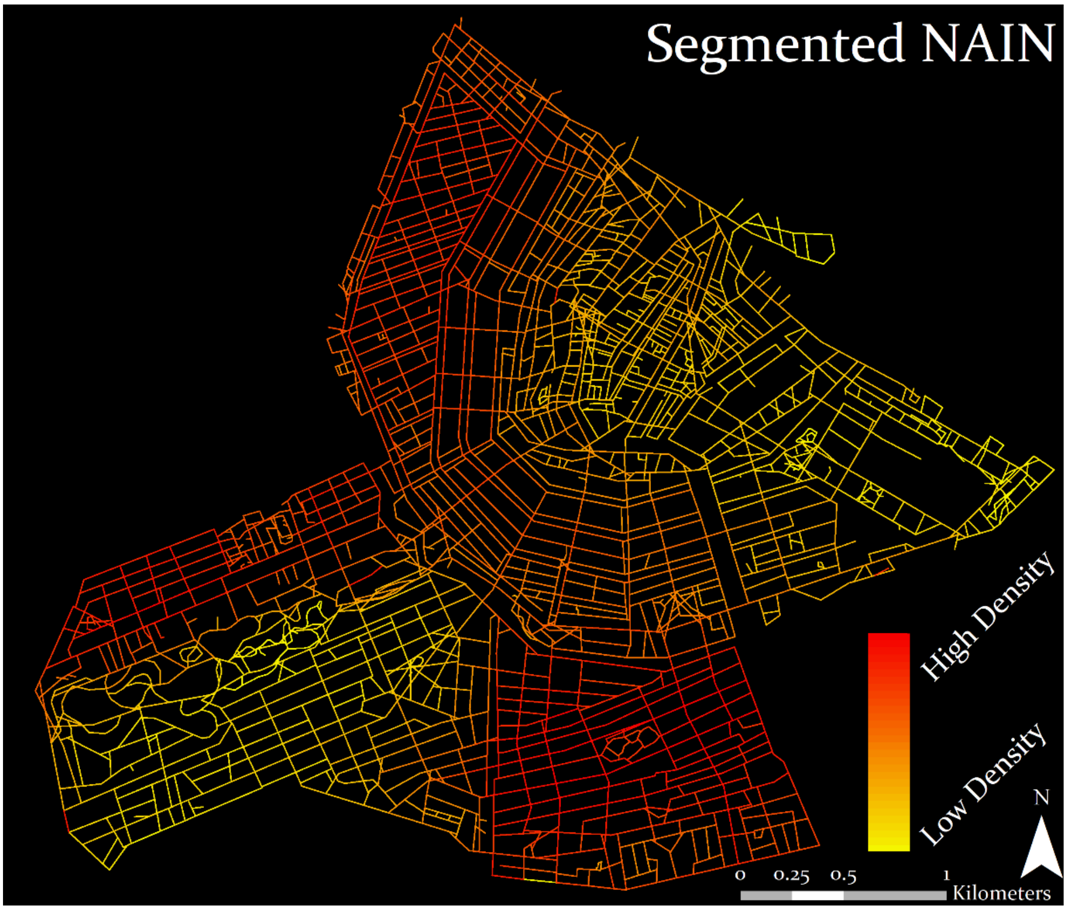


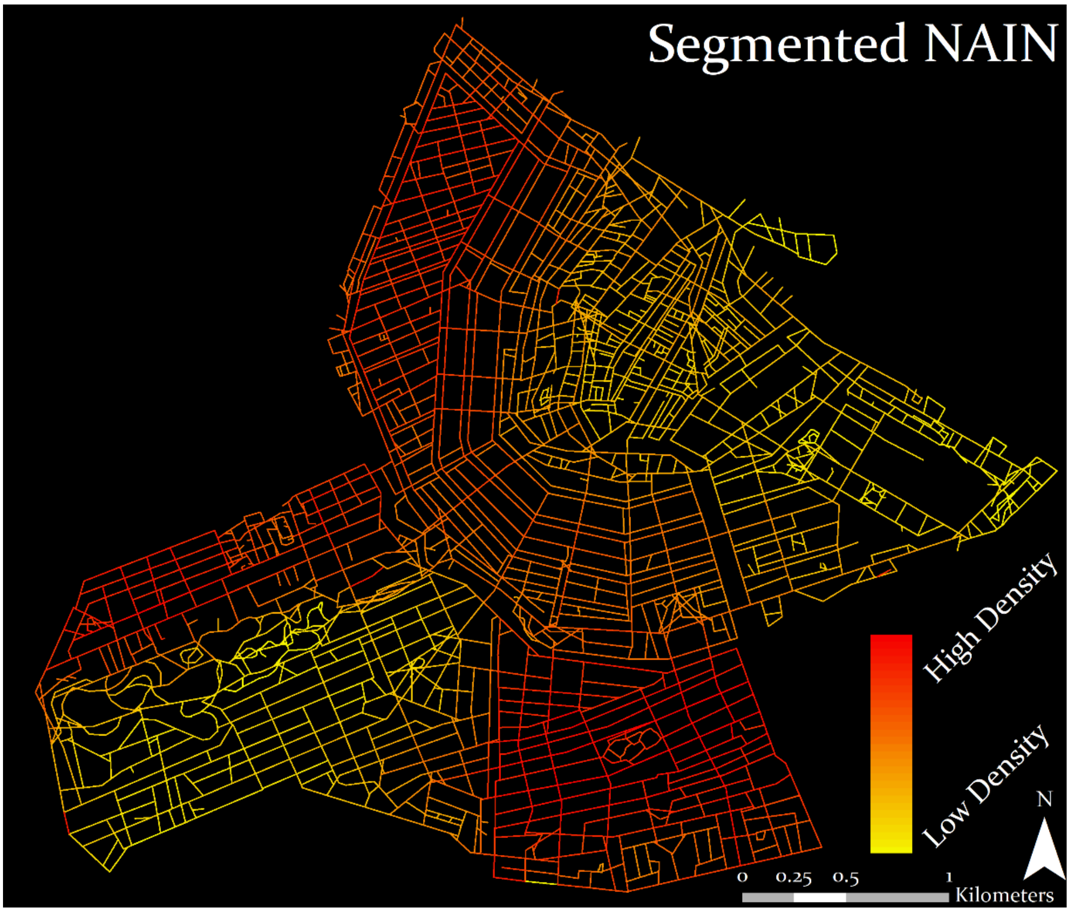

Axial Choice Normalised
High Density
Low Density
N
0 0.25 0.5 1
Kilometers

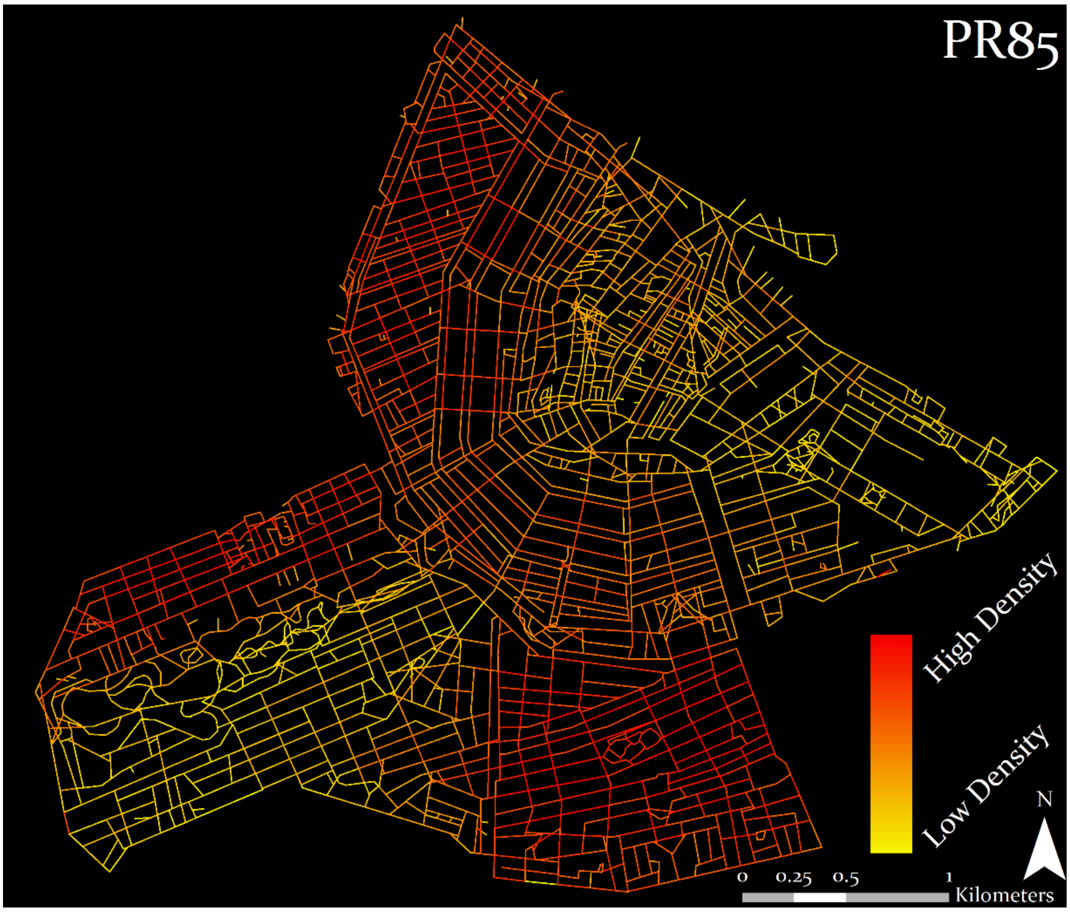

Axial Choice
High Density
Low Density
N
0 0.25 0.5 1
Kilometers

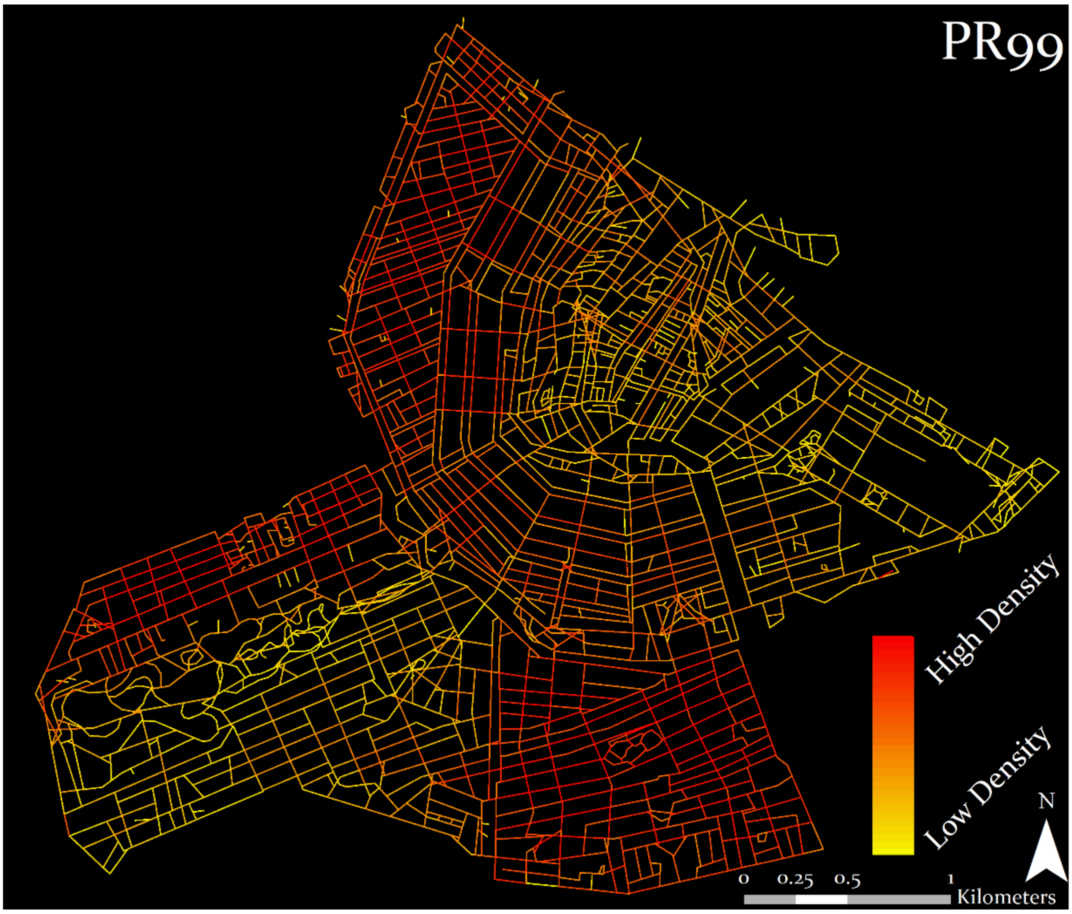

Segmented Integration
High Density
Low Density
N
0 0.25 0.5 1
Kilometers

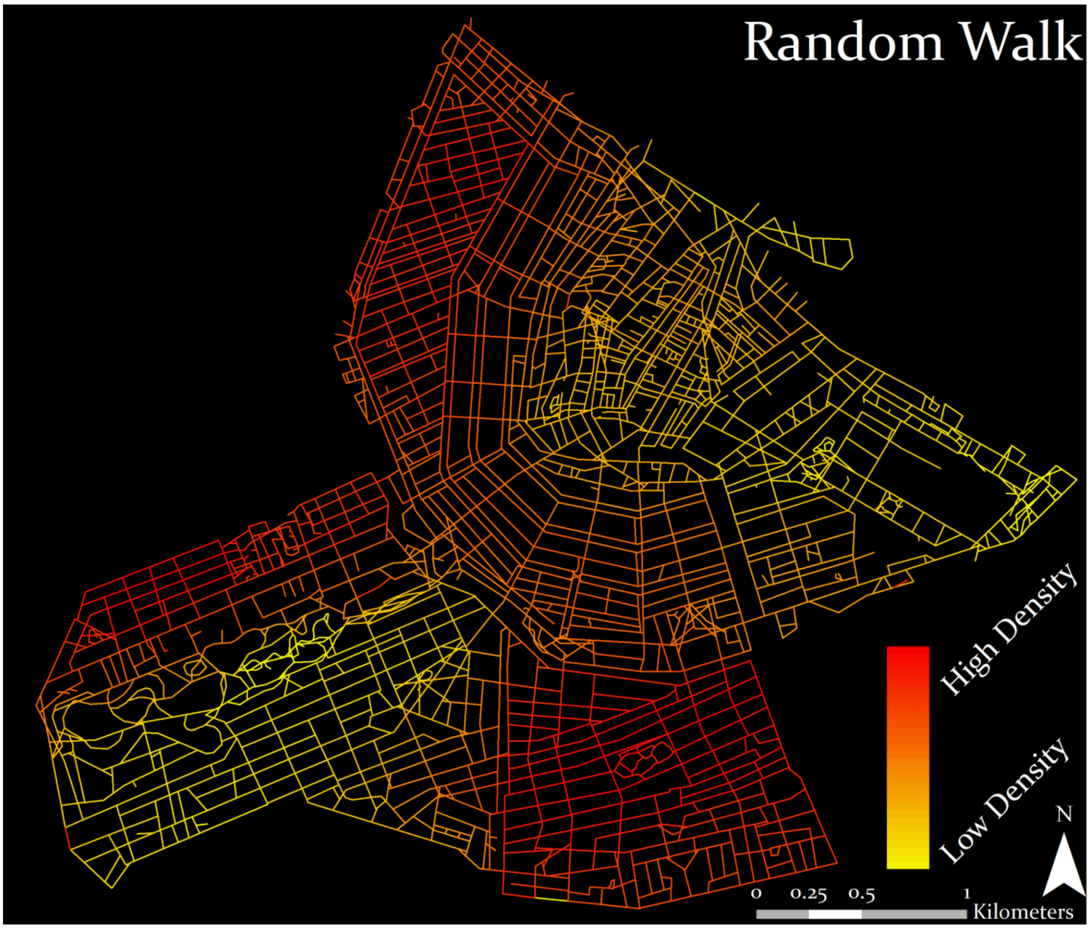

Segmented Choice
High Density
Low Density
N
0 0.25 0.5 1
Kilometers


Segmented NAIN
High Density
Low Density
N
0 0.25 0.5 1
Kilometers


Segmented NAIN
High Density
Low Density
N
0 0.25 0.5 1
Kilometers


PR85
High Density
Low Density
N
0 0.25 0.5 1
Kilometers


PR99
High Density
Low Density
N
0 0.25 0.5 1
Kilometers


Random Walk
High Density
Low Density
N
0 0.25 0.5 1
Kilometers