Computers, Environment and Urban Systems 86 (2021) 101573

Contents lists available at ScienceDirect

Computers, Environment and Urban Systems

journal homepage: www.elsevier.com/locate/ceus

# Modelling the effect of landmarks on pedestrian dynamics in urban environments

Gabriele Filomena *, Judith A. Verstegen

*Institute for Geoinformatics, University of Münster, Germany*

ARTICLE INFO

*Keywords:*
Agent-based modelling
Pedestrian navigation
Cognitive maps
Landmarks

ABSTRACT

Landmarks have been identified as relevant and prominent urban elements, explicitly involved in human navigation processes. Despite the understanding accumulated around their functions, landmarks have not been included in simulation models of pedestrian movement in urban environments. In this paper, we describe an Agent-Based Model (ABM) for pedestrian movement simulation that incorporates the role of on-route and distant landmarks in agents' route choice behaviour. Route choice models with and without landmarks were compared by using four scenarios: road distance minimisation, least cumulative angular change, road distance minimisation and landmarks, least cumulative angular change and landmarks. The city centre of London was used as a case study and a set of GPS trajectories was employed to evaluate the model. The introduction of landmarks led to more heterogeneous patterns that diverge from the minimisation models. Landmark-based navigation brought about high pedestrian volumes along the river (up to 13% of agents) and the boundaries of the parks (around 8% of the agents). Moreover, the model evaluation showed that the results of the landmark-based scenarios were not significantly different from the GPS trajectories in terms of cumulative landmarkness, whereas the other scenarios were. This implies that our proposed landmark-based route choice approach was better able to reproduce human navigation. At the street-segment level, the pedestrian volumes emerging from the scenarios were comparable to the trajectories' volumes in most of the case study area; yet, under- and over-estimation were observed along the banks of the rivers and across green areas (up to +7%, −11% of volumes) in the landmark-based scenarios, and along major roads (up to +11% of volumes) in the least cumulative angular change scenario. While our model could be expanded in relation to the agents' cognitive representation of the environment, e.g. by considering other relevant urban elements and accounting for individual spatial knowledge differences, the inclusion of landmarks in route choice models results in more plausible agents that make use of relevant urban information.

## 1. Introduction

Human movement patterns in cities emerge from the complex interactions between pedestrians and their external environment. In the last fifty years, street network analysis has emerged as one of the most widely used tools to study urban dynamics, through different lenses and for different purposes (Marshall, Gil, Kropf, Tomko, & Figueiredo, 2018). A street network analysis approach to human movement in urban spaces has originated from the idea that cities hide a mathematical order, as a result of discontinuous growth patterns and bottom-up processes (Blanchard & Volchenkov, 2009). At the same time, Agent-Based Modelling (ABM) has been described as a promising tool for mimicking self-organisation processes in pedestrian behaviour (Helbing, Molnár, Farkas, & Bolay, 2001) and unveiling the mechanisms behind movement patterns (Jiang & Jia, 2011). Therein, agents are modelled so to perceive the external environment and produce a series of actions that, in turn, reshape the environment itself. ABM allows investigating the distribution of road users across the street network, public transport services and urban spaces, under different conditions.

Yet, in existing simulation approaches, the agents' cognitive architecture and behavioural rules are usually built upon simplistic assumptions regarding route choice behaviour. 'The path selection problem has been ignored or assumed to be the result of minimizing procedures such as selecting the shortest path, the quickest path or the least costly path' (Golledge, 1995, p. 207). Shortest path algorithms based on road network distance (e.g. Torrens, 2014) or, more recently,

* Corresponding author.
*E-mail addresses:* gabriele.filomena@uni-muenster.de (G. Filomena), j.a.verstegen@uni-muenster.de (J.A. Verstegen).

https://doi.org/10.1016/j.compenvurbsys.2020.101573
Received 22 February 2020; Received in revised form 9 October 2020; Accepted 24 November 2020
Available online 13 January 2021

least cumulative angular change (e.g. Omer & Kaplan, 2017), have been employed to symbolise the cognitive functioning of urban actors in simulation models.

Nevertheless, wayfinding is a far more complex behaviour that involves perceptual and cognitive mechanisms (Allen, 1999; Gibson, 1979; Heft, 1996). To face complex and uncertain environments as cities, people tend to impose regularities, codify and organise external stimuli and spatial information. Categorisation, simplification, distortions and heuristic processes give shape to a hierarchical 'organized network of place information' (Allen & Golledge, 2007, p. 85), the so-called 'cognitive map'. According to Lynch (1960), a pioneer of mental imaging research, people decompose the city in bi-dimensional elements - landmarks, nodes, routes, edges and districts - that are connoted by high imageability. Such elements facilitate the acquisition and organisation of spatial knowledge, and support wayfinding tasks (Golledge, 1992; Lynch, 1960).

The role of landmarks is considered to be remarkable. Landmarks constitute fixed environmental features that are known and remembered for their distinctiveness within a specific environment (Evans, Smith, & Pezdek, 1982; Lynch, 1960). They are inherent to navigation and orientation: they allow an individual to place herself in the space, define goals and navigational steps (Sorrows & Hirtle, 1999). They 'function as a kind of environmental index' (Chown et al., 1995, p. 11), whose uniqueness is based on object recognition and categorisation processes. They are considered the foundation stones of cognitive maps (Chown et al., 1995; Couclelis, Golledge, Gale, & Tobler, 1987). Their relevance in human spatial behaviour (Appleyard, 1969; Presson & Montello, 1988; Sadalla, Burroughs, & Staplin, 1980) and the idea of including them in automatic navigation systems (Duckham, Winter, & Robinson, 2010; Löwen, Krukar, & Schwering, 2019; Richter, Tomko, & Winter, 2008) have prompted several attempts to formalise the concept of landmark and landmark salience. In GIScience and spatial cognition, researchers have advanced routing algorithms where the benefit of landmarks at decision points or along links are included in the formulation of shortest, more informative or simplest paths (e.g. Caduff & Timpf, 2005; Duckham et al., 2010; Duckham & Kulik, 2003; Elias & Sester, 2006). However, the enthusiasm towards landmarks has not been reflected in the ABM domain. Lynch's contribution suggests rethinking spatial knowledge formalisations in computational representations of the human-urban environment system (Filomena, Verstegen, & Manley, 2019), by incorporating information about landmarks and meaningful environmental elements.

The aim of this work is to devise and evaluate an agent-based model of pedestrian movement in urban areas that allows for the intertwined relationship between human cognition and an environment enriched with relevant geographic information. We advance a modelling approach to pedestrian route choice behaviour that combines landmark- and minimisation-based navigation. For this purpose, salient urban features - landmarks - are explicitly incorporated in the simulation environment. Agents are endowed with a simplified, rough cognitive map by means of a graph representation of the city (Gillner & Mallot, 1998; Hölscher, Tenbrink, & Wiener, 2011). We claim that such an explicit inclusion of landmarks into simulation models of pedestrian movement in urban spaces may allow reproducing human wayfinding strategies more faithfully. Furthermore, it might support the development of new theories about the reasons behind movement patterns in street networks. In this direction, we seek to answer the following research questions: a) To what extent does the incorporation of landmarks in pedestrian agents' route choice behaviour lead to different macro-level patterns as compared to pure minimisation models, i.e. shortest road-distance or least cumulative angular change paths? b) To what degree does the inclusion of a landmark-based route choice model in an ABM for pedestrian simulation contribute to generating more realistic distributions of pedestrians across an urban space?

The paper is structured as follows: literature on route choice behaviour is briefly reported as concerns the role of landmarks in wayfinding. Taking from previous research, a formalisation of *local* (on-route marks at decision points) and *global* (distant) landmarks is proposed. Thereafter, a computational approach to identify landmarks is described and employed to incorporate local and global landmarks' functions into a graph representation of the city. A landmark-based piloting approach for capturing pedestrian route choice behaviour is advanced and implemented within an ABM. Different ABM scenarios, wherein different route choice models are employed, are conceived to compare the resulting movement flows of agents for the city centre of London, UK. An evaluation of the model is also undertaken by contrasting those outputs with the volumes generated by a set of GPS trajectories.

## 2. Background: Wayfinding and spatial knowledge

Wayfinding is a complex behaviour that ensues from navigation and orientation processes (Gluck, 1991). Wayfinding in urban environments consists of processes and actions aimed at reaching distant destinations or locations. 'It involves search and decision making and its two essential components are environmental cognition and route choices' (Stern & Portugali, 1999, p. 99). Wayfinding triggers a set of processes directed at locating the destination and its direction, cognising distances and angles, orientating oneself by means of local and distant landmarks, and, finally, at embedding 'the route to be taken in some larger reference frame' (Golledge, 1999, p. 7). Planning routes in complex environments entails the use of indirect sources, as maps or navigational devices, or reliance on cognitive representations of space (Garling, Book, & Lindberg, 1984; Golledge, 1999).

### 2.1. Spatial knowledge structure and mechanisms

The relationship between wayfinding, route choice and cognitive mapping has been widely influenced by the theory of the acquisition and organisation of spatial knowledge advanced by Siegel and White (1975). The authors identify three structures - landmark knowledge, route knowledge, survey knowledge - that take shape in a sequential process. *Landmark knowledge* derives by 'unique patterns of perceptual events at a specific location' (Siegel & White, 1975, p. 23) and contains information regarding the features of a place and path-maintaining marks. *Route knowledge* works as a 'recognition-in-context' memory that makes explicit use of information regarding the order of landmarks and the consequent actions to undertake in proximity of each of them. *Survey knowledge* integrates the route and landmark representations generating complex and configurational relationships, besides involving the use of metric information.

Such a paradigm has been criticised for the assumptions regarding the sequential nature of these processes. Ishikawa and Montello (Ishikawa & Montello, 2006; Montello, 1998) have claimed that approximative quantitative information is acquired by individuals since the beginning of their experience with the environment. In this view, landmark and route knowledge would be acquired concurrently and potentially enriched with knowledge of distances and directions.

### 2.2. Landmark-based piloting and path-planning approaches

Despite the criticism concerning the sequential hypothesis theory (Siegel & White, 1975), the existence of different mechanisms designated to a) identify and recognise places, b) ascertain spatial associations between locations, c) make use of allocentric frames of references, has found support in psychobiology and neuropsychology (see Epstein, Patai, Julian, & Spiers, 2017; Grieves & Jeffery, 2017; Hafting, Fyhn, Molden, Moser, & Moser, 2005; Muller, Stead, & Pach, 1996; Nadel, 1990). As a result of a tangle of variables related to the characteristics and the complexity of the environment, as well as individual differences, people would exploit different strategies to reach their destinations: oriented search, path integration, landmark-based piloting, and referring to a cognitive map (Allen, 1999; Wiener, Büchner, & Hölscher, 2009). *Oriented search* is a strategy adopted for reaching the destination when the survey knowledge is poorly defined; few orientating elements are used to readjust one's direction (similar to the landmark, egocentric-navigation system described by Nadel (1990)), and an idiosyncratic, instinctual navigation approach is embraced by the explorer. *Path integration* consists in using cues generated by one's movement, as velocity and acceleration, so as to determine the current location in relation to the destination (Gallistel, 1990).

*Landmark-based piloting* (Allen, 1999; Epstein & Vass, 2014; Gallistel, 1990), or route-based navigation, refers to a more elaborate strategy whereby the position of known, local and distant landmarks is used to formulate a path. Heath, Cornell, and Alberts (1997) argue that points of reference are memorised when learning a route and used at a later time to indicate turning points. Pushing forward this idea, Lovelace, Hegarty, and Montello (1999) report that people recur to two types of landmarks: landmarks at choice-points and off-route landmarks. The first may be considered local landmarks or *route-marks*, while the second ones global or external points of reference (Richter & Klippel, 2005). They suggest what a walker should do at a specific junction, they help to follow the correct path across the street network and identify the next anchor-point (Michon & Denis, 2001). Thus, local landmarks are used to create a sequence of decision points and to chunk long and complex routes in sub-sections, while distant landmarks are associated with directions and orientation. Such an approach is employed when finding a route towards a novel or known destination and it requires a minimal knowledge of possible salient landmarks (Allen, 1999).

In line with Siegel and White's theory, some scholars further differentiate this type of strategy from a form of wayfinding based on cognitive maps. In this case, instead of employing an 'undifferentiated' and incomplete representation of space, a more integrated spatial schema, containing regional divisions and hierarchical relationships between features, is thought to support wayfinding decisions (Allen, 1999; Allen & Golledge, 2007; Kitchin, 1996). The difference between the piloting and cognitive maps navigation approaches is thin and may lay on the degree of accuracy of the space representation of an individual, although depending on similar cognitive mechanisms. Epstein and Vass (2014) indeed report how landmark-based piloting entails an interplay of different functional brain areas in relation to a) landmark identification, b) self-localisation, orientation and direction determination, c) 'encoding and retrieving' relationships between locations, within an allocentric frame of reference. Whereas a simple landmark strategy can be preferred over a survey-like knowledge in some circumstances (Foo, Warren, Duchon, & Tarr, 2005), metric information still plays a role in route choice decisions in urban environments (e.g. Golledge, 1995; Rodríguez et al., 2015). 'A combination of geometry and landmarks defines a given space and (…) such information has a privileged status in the creation of cognitive maps' (Nadel, 2013, p. 79). As landmarks shape and structure a topological survey-knowledge, where connectivity and sequences are encoded (Kuipers, Tecuci, & Stankiewicz, 2003), they prompt distance representations to possible next, known locations, as well as turning decisions. Thereby, on-route local landmarks may activate a 'locally minimising distance' (LMD) heuristic, which is based on straight-line computations and makes use of short-term local representations of space (Gärling, Säisä, Book, & Lindberg, 1986; Hayes-Roth & Hayes-Roth, 1979).

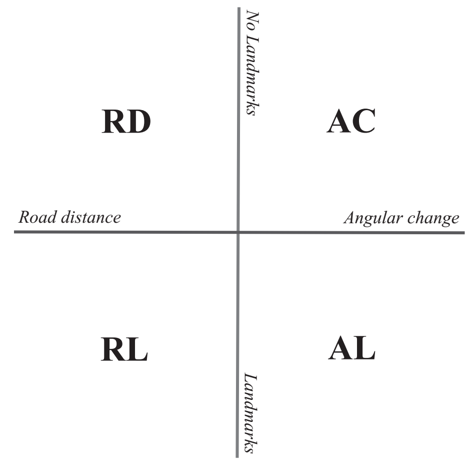



**Fig. 1.** Characteristics of the scenarios analysed with the ABM, with respect to minimisation approaches and incorporation of landmarks.

## 3. Methodology

### 3.1. Conceptual model and origin-destination matrix definition

A landmark-based route choice model is here proposed and incorporated in an ABM for pedestrian movement simulation to capture how people move between an origin and a destination. In a nutshell, agents, who represent pedestrians, are designed to complete trips between pairs of origins and destinations (OD) in the simulation environment. The

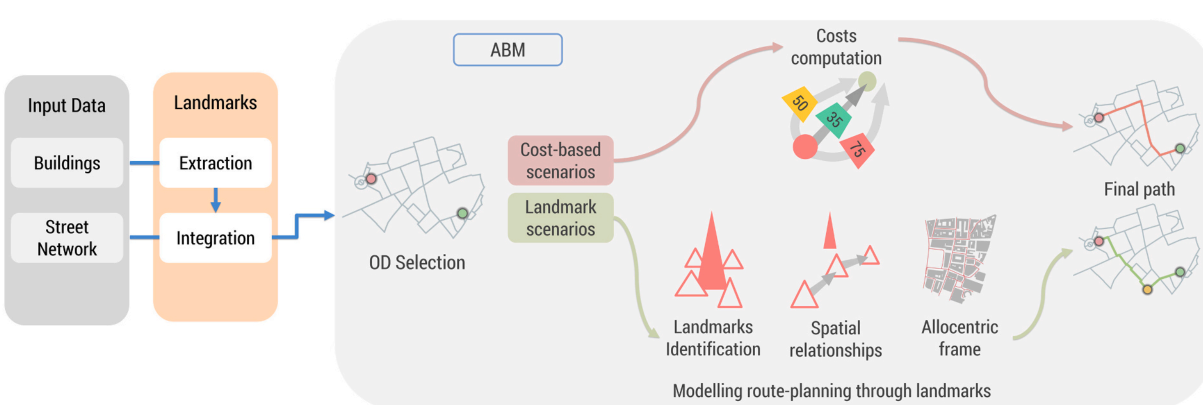



**Fig. 2.** A summary of the methodology. Prior to the simulation, input data are processed: landmarks are extracted and integrated into the street network. Within the model, a fixed set of OD pairs is defined per each run, on the basis of the length of the GPS trajectories used as observational data. Thereafter, four scenarios are implemented: the cost-based scenarios (RD, AC scenarios) minimise road costs only; the landmark-based scenarios (RL, AL scenarios), encompass three interdependent steps whereby landmarks and locations are identified, relationships are established and situated in an allocentric frame of reference.

environment is a graph representation of the case study area's street-network, enriched with landmark information (Section 3.2).

Four scenarios are devised to compare the patterns emerging from different route choice models employed by the agents (Fig. 1):

1. Road-distance minimisation, *RD scenario*;
2. Least cumulative angular change, *AC scenario*;
3. Landmark-based piloting route choice model, combined with a local, road-distance minimisation heuristic, *RL scenario*;
4. Landmark-based piloting route choice model, combined with a local, least cumulative angular change heuristic, *AL scenario*;

The first two scenarios are defined *cost-based scenarios* (Fig. 2) and are typically used in existing simulation models (e.g. Omer & Kaplan, 2017; O'Sullivan, Thurstain-Goodwin, & Schelhorn, 2001; Torrens, 2012); minimisation heuristics are here adopted at the global level. In the RD scenario, agents minimise road distance and take therefore the shortest route; in the AC scenario, they take the route with the least cumulative angular change between the origin and the destination. The Dijkstra's shortest path algorithm (Dijkstra, 1959) is employed to compute the routes: edges in a primal graph are weighted with road distance (RD scenario) while edges in a dual graph, representing street intersections, are weighted with the angle of deflection between the two corresponding intersecting road segments (AC scenario, see Hillier & Iida, 2005). In the other two scenarios, defined *landmark-based scenarios*, comprising the route choice models that we propose, agents, when planning their routes, employ landmarks along with the minimisation heuristics modelled in the first two scenarios. In these scenarios, the heuristics are used at the local level, between decision points (Section 3.3). The model was implemented in GeoMASON, a multi-agent simulation environment written in the Java programming language (Sullivan, Coletti, & Luke, 2010). Along with the case study input data, the model is available on a GitHub repository (Filomena, 2020a).

In all the scenarios, the perceived costs of distances and turn angles are modelled as stochastic variables to account for people's rough perception of distances and turn angles (Montello, 1997; Tversky, 1992). The perceived cost $i_e$ of a street segment (edge) $e$ is defined as:

$$i_e = cost_e + Z_f{}^*\sigma^*cost_e \quad \text{with } Z_f \sim N(0,1) \tag{1}$$

where $cost_e$ is the actual cost of the segment; $Z_f$ is a standard normal distribution and $\sigma$ is the standard deviation (equal to 0.1). Thus, the deviation from the actual cost is relative, i.e. the perception error is typically larger for longer segments or more abrupt turns.

In order to make the results comparable across the four scenarios, a single set of OD pairs is generated prior to starting one run of the simulation:

1. An origin node is selected from the whole set of nodes of the street network.
2. Afterwards, a destination node is picked at a distance equal to the length (±50 m) of a randomly chosen trajectory; this is a GPS trajectory belonging to the observational dataset employed for the model evaluation (see Section 3.5).
3. The pair is added to the set of OD pairs, which is in turn input into each scenario.

The four ABM scenarios are run 50 times each as Monte Carlo simulations, to account for the randomness resulting from the OD pairs selection. The model counts every time a pedestrian crosses a street segment. Per each scenario, the median count value of each segment across the 50 executions is calculated to obtain the agents' volume per segment (*meso-level* of analysis). In such a way, the patterns emerging from the scenarios are interpretable and meaningful beyond the specificity of a single run and its set of OD pairs. Finally, *macro-level* statistics are computed as regards the average length of the routes, the number of segments employed by the agents and the landmarkness accumulated along the routes. The Pearson product-moment correlation coefficient is computed between scenarios on the agents' volumes per street segment resulting from the model.

### 3.2. Landmarks extraction and integration into the street network

A methodology to computationally assess the cognitive salience of the city elements, built upon the work of Sorrows and Hirtle (1999) and Raubal and Winter (2002), is presented by Filomena, Verstegen, and Manley (2019). From that work, we borrow the method to compute landmark scores, here denoted as *landmarkness*, for each building in a case study area. Landmarkness is the combination of visual, structural, pragmatic and cultural components. Weights are used to manipulate not only the relevance of a specific component, but also the importance of different indexes within these components (see Table 1).

The global landmarkness of a building $b$ is defined as:

$$g_b = \frac{v_b - min(v_U)}{max(v_U) - min(v_U)}{}^*\alpha + \frac{s_b - min(s_U)}{max(s_U) - min(s_U)}{}^*\beta + \frac{c_b - min(c_U)}{max(c_U) - min(c_U)}{}^*\gamma + \frac{p_b - min(p_U)}{max(p_U) - min(p_U)}{}^*\delta \tag{2}$$

where $v_b$, $s_b$, $c_b$ and $p_b$ are the visual, structural, cultural and pragmatic landmark scores of a building $b$, and $\alpha$, $\beta$, $\gamma$ and $\delta$ are the weights of these components. Thus, $g_b$ is obtained by rescaling each component's value in relation to its highest and lowest values (e.g. $max(c_U)$, $min(c_U)$), within the entire set of buildings in a case study area (set $U$). Likewise, we compute the local landmarkness of a building, within a smaller number of adjacent buildings, that are inside the bounds of a buffer of $f$ meters around $b$ (set $T$), and using different weights as:

$$l_b = \frac{v_b - min(v_T)}{max(v_U) - min(v_T)}{}^*\kappa + \frac{s_b - min(s_T)}{max(s_U) - min(s_T)}{}^*\eta + \frac{c_b - min(c_T)}{max(c_U) - min(c_T)}{}^*\omega + \frac{p_b - min(p_T)}{max(p_U) - min(p_T)}{}^*\zeta \tag{3}$$

Finally, both $g_b$ and $l_b$ are rescaled across the whole set of buildings $U$ in order to obtain values ranging from 0 to 1.

In line with formalisations proposed in spatial cognition research (Lovelace et al., 1661; Michon & Denis, 2001; Richter & Klippel, 2005) (see Section 2.2), we here define as *distant landmark* a building visible from several locations (junctions in the street network) and endowed with high city-level landmarkness $g_b$. Such an element supports individuals' orientation processes: walkers use distant landmarks to understand in which part of the city they are, or simply to organise their spatial knowledge. Conversely, the local nature of a landmark indicates to the explorer what actions have to be undertaken along the route, to reach the chosen destination. A building is in this case an *on-route mark* with high local landmarkness $l_b$, that elicits spatial decisions. Here we consider as possible on-route marks local landmarks at decision points.

The next step consists in incorporating the distribution of local and global landmarks[1] across the urban space into the graph representation of the urban environment. For this purpose, each node $n$ (street junction) is assigned with three sub-sets of buildings:

- $B_n$: the set of local landmarks at the junction; all the buildings within a certain distance $q$ from the junction.
- $V_n$: the set of visible, global landmarks (with $g_b > y$, $y$ defined by the researcher, to exclude meaningless buildings) from the node $n$, located at least $r$ meters away from it. Visible buildings from a location are identified by constructing 3d sightlines.

[1] Local and global landmarks are considered respectively on-route marks and distant landmarks when the agent employs them throughout a certain trip.

**Table 1**
Landmark extraction: weights of the indexes and components when computing global $g$ and local $l$ landmarkness (see Filomena, Verstegen, & Manley, 2019).

| Component | Index | Index weight | Weight $g$ | Weight $l$ |
|---|---|---|---|---|
| Visual ($v$) | 3d Visibility | 0.50 | $\alpha = 0.50$ | $\kappa = 0.25$ |
| | Façade area | 0.30 | | |
| | Height | 0.20 | | |
| Structural ($s$) | Area | 0.30 | $\beta = 0.30$ | $\eta = 0.35$ |
| | 2d advance visibility | 0.30 | | |
| | Neighbours (150 m buffer) | 0.20 | | |
| | Road distance | 0.20 | | |
| Cultural ($c$) | Historical importance | 1.0 | $\gamma = 0.10$ | $\omega = 0.10$ |
| Pragmatic ($p$) | Land-Use (200 m buffer) | 1.0 | $\delta = 0.10$ | $\zeta = 0.30$ |

- $A_n$: global landmarks (with $g_b > y$) within a buffer of $x$ meters around $n$; they guide the agent when the node $n$ is selected as a destination.

### 3.3. Modelling route choice behaviour through landmarks

The stages of the route choice models here proposed are loosely built upon the spatial knowledge model proposed by Siegel and White (1975), in light of the findings of Ishikawa and Montello (2006) concerning the employment of metric and direction information at all the decisional stages (see Section 2.1). In our conceptualisation, landmark navigation (landmark knowledge), relationships between locations (route knowledge), the use of an allocentric frame of reference, distance and direction information (survey knowledge), interact across the whole route choice process, as discussed in Section 2.2. The information stored in the graph representation of the urban environment is used by the agents in the landmark-based scenarios when formulating a path throughout the environment. We speculate that the distance to be walked and the legibility of the environment surrounding the agent determine the *wayfinding complexity* (Allen & Golledge, 2007; Lynch, 1960), and, in turn, to what extent the agent has to resort to on-route marks and distant landmarks.

In short, the route choice model with landmarks consists of one or more steps, depending on the complexity of the environment between the origin and the destination (Fig. 3). The phases are described in the next two sub-sections but refer to the appendix for the detailed implementation.

### 3.4. Selection of intermediate decision points with on-route marks

In the model, when formulating a route, the agent is able to represent the external environment and take spatial decisions only within a certain distance between the origin (or a successive intermediate location) and the destination. Such a *search-space* is identified on the basis of the perceived wayfinding complexity between the current location[2] of the agent and the destination, by using the set of buildings $U$ (Fig. 3a). If the wayfinding complexity is higher than a certain threshold $u$ defined by the researcher, the agent needs on-route marks to find its way towards the destination and to better represent distances and turns. Thus, it considers a set of known junctions (nodes) that are vivid in its representation of the environment (*cognitive map*); it evaluates them in relation to local landmarkness, using the set $B_n$, and chooses the best one as an intermediate decision point (Fig. 3b). The selection of an intermediate decision point (a node $n$), within the search-space, is based on a measure of attractiveness $a_n$ computed as:

$$a_n = k_n{}^*\rho + \frac{e_{c,d} - e_{n,d}}{e_{c,d}}{}^*\tau \tag{4}$$

[2] With *current location* we indicate either the origin or an already selected intermediate point from where the agent assesses the complexity of the environment; except for the origin, the agent is not actually placed there.

where $k_n$ represents a score of local salience of the landmarks at the node $n$, given the set $B_n$ (see Section 3.2), $e_{c,d}$ is the Euclidean distance between the current location of the agent and the destination $d$, and $e_{n,d}$ the Euclidean distance between the candidate node $n$ and $d$. Hence, the last factor represents the gain in distance towards the destination that the node would grant to the agent. $\rho$ and $\tau$ are the weights assigned to the two attractiveness components. The agent reassesses the wayfinding complexity of the environment (Fig. 3c) from the chosen intermediate decision point. If the complexity between the new location and the destination is still too high to mentally conceive the space around the destination (i.e. $>u$) the agent looks for a further decision point within the new cognisable space, and so forth. In-between decision points, agents make use of local minimisation heuristics, in combination with global landmarkness, by using the sets $V_n$ and $A_d$ ($d$ stands for the destination node, see the next subsection). Finally, a global path is formulated by joining the different chunks of the route.

### 3.5. Connecting the sequence of decision points and formulating the final path by employing distant landmarks

If the complexity of the environment is low, a simple route is computed without intermediate decision points by exclusively employing a minimisation heuristic (the sequence contains two nodes only: the origin and the destination). In this case, road distance (*RL scenario*) or turn angle (*AL scenario*) costs are used to compute a route by means of the Dijkstra's shortest path algorithm between the OD pair (Dijkstra, 1959). On the contrary, when the agent has identified a sequence of intermediate decision points (i.e. the full sequence of nodes including origin and destination consists of more than two nodes), the cost minimisation heuristic is combined with a maximisation of the distant landmarks' visibility so to formulate routes in-between intermediate decision points (nodes), again by means of the Dijkstra's shortest path algorithm (Dijkstra, 1959).

To account for the orientation functions played by distant landmarks, a score $h_n$, based on the visibility of distant landmarks, is assigned to each node $n$ when the agent is formulating its path. We assume that pedestrians use distant landmarks only when they are far away from the destination. Whilst the agent is travelling, for each possible traversed node that is more distant than $r$ meter from the destination (parameter defined by the researcher), the score $h_n$ is computed as:

$$h_n = max\left(g_b{}^*\frac{e_{n,d}}{e_{b,d}}\right) \quad with \;\; b \in (V_n \cap A_d) \tag{5}$$

where $V_n$ is the set of visible buildings from the node $n$ and $A_d$ the set of orienting landmarks around the destination node $d$ (see Section 3.2); $h_n$ is the best value amongst the resulting products of the global score $g_b$ of each building $b$, belonging to the sets $V_n$ and $A_d$, and a distance factor. The latter is determined by the location of the node $n$ in relation to the destination and the distance between the landmark $b$ and the destination, with $e_{n,d} > r$. In this way, the impact of a reference point is relative

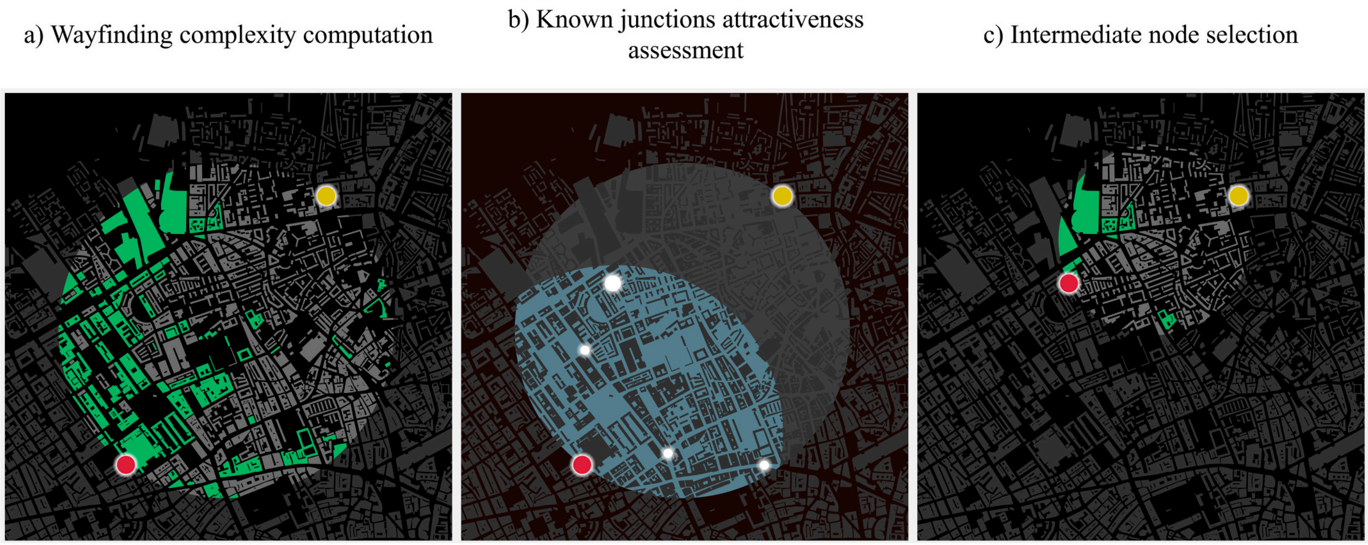



**Fig. 3.** Illustration of the landmark-based piloting route choice model. In a) the agent computes the wayfinding complexity of the area between a location (red node) and the destination (yellow node), based on the environmental legibility (local landmarks in the area) and distance separating the agent from the destination; b) known junctions are identified and assessed in relation to the local salience of their adjacent buildings and the possible gain in distance towards the destination. The most advantageous junction is selected as an intermediate decision point (the largest node in the figure); c) from the chosen intermediate decision point (red node) phase a) is repeated. (For interpretation of the references to colour in this figure caption, the reader is referred to the web version of this article.)

to its location and the possible position of the agent. If the distance factor is >1, the value is fixed at 1. When there are no orienting landmarks around the destination, $h_n$ is set to 0 and indicates poor orientation towards the destination.

Therefore, when formulating a path between two nodes in the sequence, the road cost of an edge *e* in the graph representation is combined with the inverse of its *to-node*'s *h* score as follows:

$$weight_e = (1 - h_n)^* \rho^* i_e \tag{6}$$

where $\rho$ is the weight used to regulate the contribution of $h_n$, $i_e$ represents a road cost, i.e. road distance or turn angles. In this manner, for example, a certain route may be chosen over a shorter one (or characterised by a lower cumulative change of direction) because of the positive influence of distant, orienting landmarks visible along the route's junctions.

### 3.6. The case study: London city centre

The area defined by the boundary of the Congestion Charge Zone, London (UK) was used as case study area (see Fig. 4). London's transport authority has initiated a 'Walking action plan' to design healthy urban spaces with the goal of reaching a percentage 80% of trips by foot by 2041 (Transport for London, 2018). This vision aims at reducing reliance on cars and entails an explicit integration of its transport network with walking behaviour. Such policies, the city's extended transport network, morphological diversity in the urban layout and socio-cultural complexity make London a challenging and appealing case study area for pedestrian simulation models.

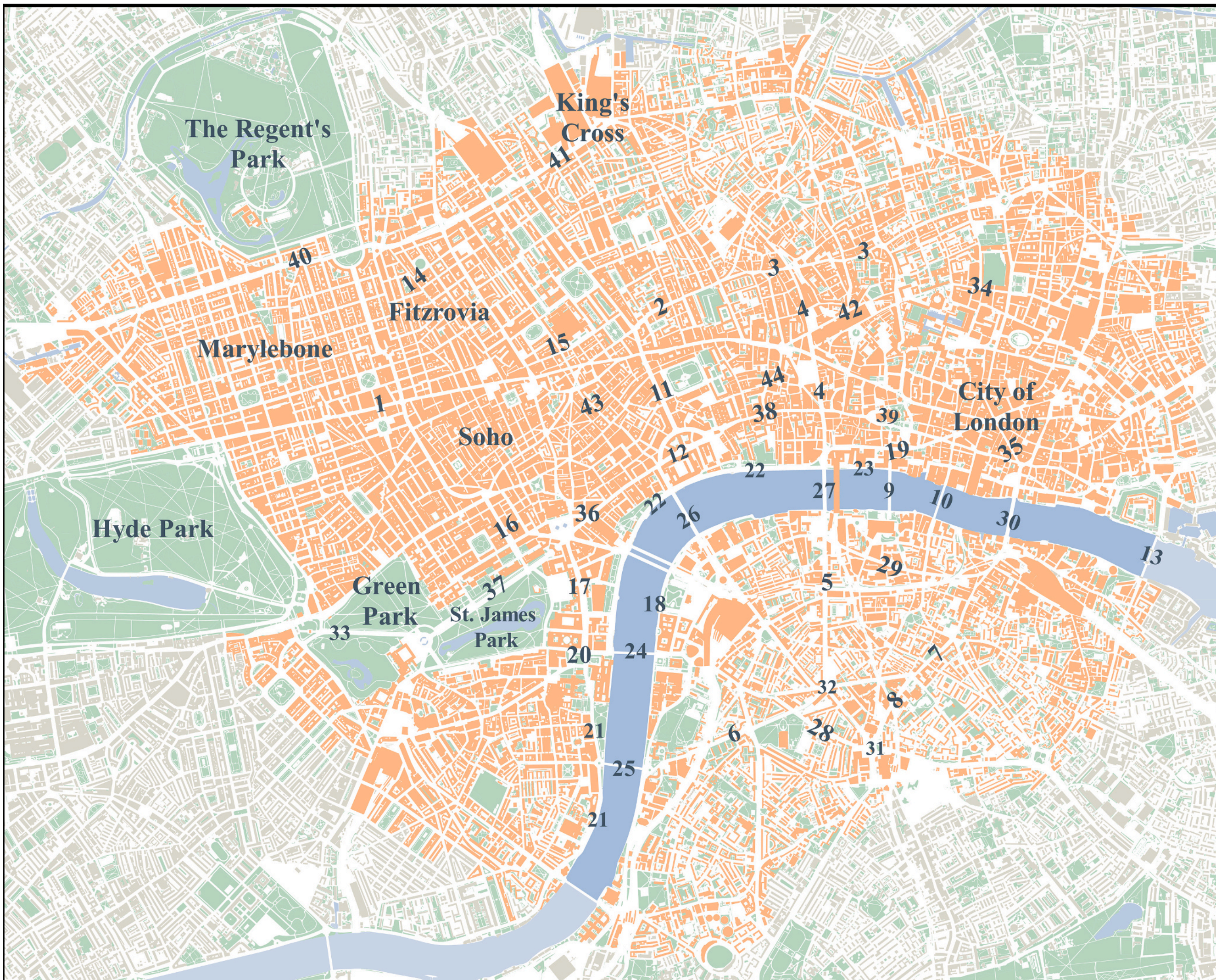
The values of the parameters mentioned in the previous sections were defined for this case study area on the basis of the buildings distribution, the extent of the case study area and subjective judgements (Table 2). The number of trips in each simulation run was set to 255, as the number of collected GPS trajectories (see Section 3.5), in order to facilitate the comparison between the scenarios' and trajectories' volumes. To better illustrate the functioning of the landmark-based piloting route choice model, additionally to the four scenarios, a set of agents was designed to navigate only employing on-route marks and filling in gaps between intermediate decision points by minimising road distance; another set of agents would instead navigate maximising the visibility of distant landmarks through the route. In the text we refer to these two further types of agents as *LL* and *GL* respectively.

### 3.7. Model evaluation with observational data

We evaluated the performance of the model under the four different scenarios with observational data. A set of GPS trajectories was obtained from *gpsies.com*. Only trajectories categorised as *walking-one way* or *hiking-one way* were kept. Additionally, a subjective examination of each individual trajectory was undertaken: incomplete trajectories and trajectories with a limited number of way-points were disregarded. After this selection process, 255 valid tracks were obtained. On the basis of the proximity to nodes and edges of the street network graph, the trajectories' way-points were snapped to the street network. A Python Jupyter Notebook documenting the analysis is available on a Github repository, along with the necessary input files (Filomena, 2020b).

The scenarios were evaluated as follows:

- For all the routes resulting from the scenarios and the trajectories, cumulative measures of local and global landmarkness were computed. Cumulative local landmarkness over a route was calculated by summing the highest local score amongst the buildings adjacent to each of the nodes traversed. Cumulative global landmarkness was computed as the sum of each traversed node's *h* (distant landmarks visibility score) as described in Section 3.3, Eq. 5. Such indexes were compared with the landmarkness values of the trajectories, to explore the ability of the different route choice models of the ABM to make use of landmarks alike the individuals of the GPS trajectories. The Games-Howell post-hoc test (Games & Howell, 1976), was employed to assess whether the differences were statistically significant (with a 0.05 *p*-value). The Games-Howell post-hoc test is a non-parametric alternative for the *t*-test that can be applied to compare multiple groups with unequal variances.
- For each street segment, a count of all the crossing trajectories was used to determine the observed pedestrian volumes. The distributions of agents across the street segments, resulting from each scenario, were compared with these observed volumes. The Root Mean Square Error (RMSE) was computed over the entire street network to indicate the ability of each scenario to generate realistic results at the macro-level. Furthermore, the difference between modelled and




**Roads**

| | | | |
|---|---|---|---|
| Blackfriars Rd | 5 | Marylebone Road | 40 |
| Bloomsbury St | 15 | Millbank | 21 |
| Borough High St | 7 | Newington Causeway | 8 |
| Charterhouse Street | 42 | Oxford St | 1 |
| Chiswell St | 34 | Paul's Walk | 23 |
| Clerkenwell Rd | 3 | Queen Victoria St | 19 |
| Cleveland St | 14 | Regent St Saint James's | 16 |
| Constitution Hill | 33 | Southwark St | 29 |
| Endell Street | 43 | St Paul's | 39 |
| Euston Road | 41 | St. George's Rd | 28 |
| Farringdon St | 4 | The Mall | 37 |
| Fetter Lane | 44 | The Queen's Walk | 18 |
| Fleet Street | 38 | The Strand | 12 |
| King William St | 35 | Theobalds Rd | 2 |
| Kingsway | 11 | Victoria Embankment | 22 |
| Lambeth Rd | 6 | Whitehall | 17 |

**Squares and bridges**

| | |
|---|---|
| Blackfriars Bridge | 27 |
| Charing Cross | 36 |
| Elephant and Castle | 31 |
| Lambeth Bridge | 25 |
| London Bridge | 30 |
| Millennium Bridge | 9 |
| Parliament Square | 20 |
| Southwark Bridge | 10 |
| St George's Circus | 32 |
| Tower Bridge | 13 |
| Waterloo Bridge | 26 |
| Westminster Bridge | 24 |

**Fig. 4.** case study area: The city centre of London (UK), delimited by the Congestion Charge Zone. The labelled locations are mentioned in the results and discussion section.

**Table 2**
Values assigned to the parameters (weights for computing landmarkness are defined in Table 1).

| Parameter | Description | Value |
|---|---|---|
| $f$ | Distance from a building for computing its local salience | 1500 m |
| $q$ | Max. distance between a building and a node for extracting buildings at a junction | 80 m |
| $r$ | Min. distance between a node $n$ and landmarks, when computing 3d visibility from $n$. | 300 m |
| $y$ | Threshold for identifying global and local landmarks | 0.30 |
| $u$ | Wayfinding complexity threshold above which a route between a location and a destination would require intermediate decision points | 0.05 |
| $x$ | Distance around a node for identifying its orienting landmarks | 2000 m |
| $\rho$ | Weight of $k_n$, local salience of landmarks at the node $n$, in Eq. 4 | 0.60 |
| $\tau$ | Weight of the distance factor in Eq. 4 | 0.40 |
| $\rho$ | Weight of $h_n$, the distant landmarks visibility score at a node $n$, in Eq. 6 | 0.80 |

observed volumes - expressed in the number of standard deviations (computed over the 50 Monte Carlo runs) away from the volumes derived from the GPS trajectories - was computed to examine where the model may have over-, under- or correctly estimated the observed movement flows. This approach was chosen to take into account the variation between runs stemming from the random selection of OD pairs.

## 4. Results and discussion

### *4.1. Comparison amongst scenarios: Emerging pedestrian agents' volumes across the street network*

The landmark-based scenarios bring about a number of changes at the global level, compared to the cost-based scenarios, that could be considered more realistic, as outlined in the following (see Fig. 5 for macro-level patterns). The average length of the RL agents' routes is 559 m longer than the RD agents' (+16%), whereas the average length of the AL agents' routes is 440 m longer than the AC agents' (+11%) (Table 3). The correlation matrix in Table 4 suggests that the introduction of landmarks led to more divergent patterns between the RD and RL scenarios (correlation of 0.62) than between the AC and AL scenarios (correlation of 0.83). This is also evoked by Fig. 5, as the patterns emerging from the AL scenario retrace the skeleton of the main paths depicted by the agents in the AC scenario; conversely, the patterns of the RD and RL scenarios look more diverse from each other. Throughout this section, examples are given to elaborate on these observations and to speculate why they come about.

In the RD scenario, the inner ring emerges as a path used to move across the city centre. This is particularly evident in the north. In the central part of the case study area, other major roads stand out: sections of the A40 road (i.e. Oxford Street, Theobalds Road and Clerkenwell Road), A201 (i.e. Blackfriars Road and Farringdon Street), the A3 (i.e. Borough High Street, Newington Causeway) and Lambeth Road represent the main lines of movement in this scenario. These roads were crossed by between 14 and 22 agents, 5% to 8% of all the agents. Furthermore, agents made an even use of most of the bridges in the central area, apart from Southwark Bridge and the pedestrian Millennium Bridge, which were neglected and rarely traversed. This is the scenario with the lowest number of segments traversed at least by 5 agents (605) and the second highest as regards the number of segments that were crossed at least once (5053). This indicates that agents solely designed to walk as little as possible do use a wide range of roads.

In the RL scenario, novel paths are depicted by the agents' routes. Parts of the inner ring and some major roads were still used regularly, but Oxford Street and Theobalds Road, often walked in the RD scenario, were rarely crossed (2 to 5 times). Agents in the RL scenario made use of a similar number of roads to RD (5421 segments were crossed at least once) but recurring detours and deviations made certain roads emerge more prominently (e.g. Cleveland Street, Endell Street, Fetter Lane, Charterhouse Street and Bloomsbury Street up to 10 agents). In the east, Queen Victoria Street (up to 19 agents, 7% of the agents) stands out along with several roads around St. Paul's Cathedral.

However, the most prominent differences between the RD and RL scenarios can be ascribed to the agents' behaviour along the north bank of the Thames[3], and in the proximity of St James's Park and Green Park (peaks of 15 agents, 6%). Up to 34 agents (13%) walked around the Big Ben and Westminster (Parliament Square). See route 2 (Fig. 6) for an example of how Westminster, as an on-route mark, brings the agent to deviate from the shortest path and take a different bridge across the Thames, towards the south. Victoria Embankment is featured by similar figures (peaks of 32 agents, 12.5%) west of St. Paul's; from there the flow spreads towards different origins and destinations. In route 3 (Fig. 6), the agent prefers to walk along the river and cross it only after having passed an on-route mark. Streets adjacent to parks and following the north bank of the river are indeed featured by high visibility of global orienting landmarks (see bottom-right Fig. 5). Furthermore, several prominent landmarks in Parliament Square, Marylebone, between Fitzrovia and Soho in the west, and in the City of London, in the east, had an influence on reshaping the routes, by means of intermediate decision points (see bottom-left panel in Fig. 5).

The south riverbank was rarely traversed in the RL scenario. An exception is represented by the path between Blackfriars and London Bridge, walked by up to 14 agents (5%). Low volumes in the westward sections of Waterloo Bridge could be due to the fact that the agents preferred to cross the river in the western part of the city (as in route 2, Fig. 6) and to move along Lambeth Road (9 agents), or through the streets that branch off Westminster Bridge. Moreover, the visibility of several global landmarks placed along or in the proximity of the south bank itself may raise the attractiveness of the north riverbank. The latter was significantly preferred by agents in the RL and AL scenarios.

In the AC and AL scenarios major roads played a central role. In the AC scenario, the A40, between the west and the east, and the A201, between the north and the south, were crossed by up to 35 agents per segment (Farringdon Street, 13% of the agents). The north section of the ring was often walked between King's Cross and Regent Park, and the north bank of the river was traversed by different routes in almost all its extension (peaks of 26 agents, 10% of the agents). Agents wandered between Green Park, the Big Ben (Parliament Square) and then along the river till London Bridge. Due to its role in linking Farringdon Street and Blackfriars Road (A201), Blackfriars Bridge (29 agents, 11%) was privileged over other bridges which are not placed along the main lines of movement. Furthermore, the south bank was traversed mainly in its western section, till the Waterloo Bridge by up to 14 agents (5%). A high number of agents also walked along Southwark Street (18 agents, 7%) and other primary roads in the south, depicting a clear skeleton of paths that would allow people to move minimising cumulative angular change. Such route choice model brought agents to make use of a very small number of segments. Indeed, this scenario presents the lowest number of street segments (3015) crossed at least once.

The AL scenario follows a similar pattern but a wider set of segments and routes were exploited by the agents: a higher number of segments was crossed at least once (3686 vs 3015 in the AC scenario). The major roads emerging from the AC scenario are still featured by high volumes of AL agents and the north bank of The Thames was even traversed more frequently (32 agents along Victoria Embankment, 12%). Nevertheless, the number of agents in Theobalds Road and Clerkenwell Road are

[3] We refer in the text to the north bank as the bank on the north side of the Thames, e.g. Millbank, Victoria Embankment.

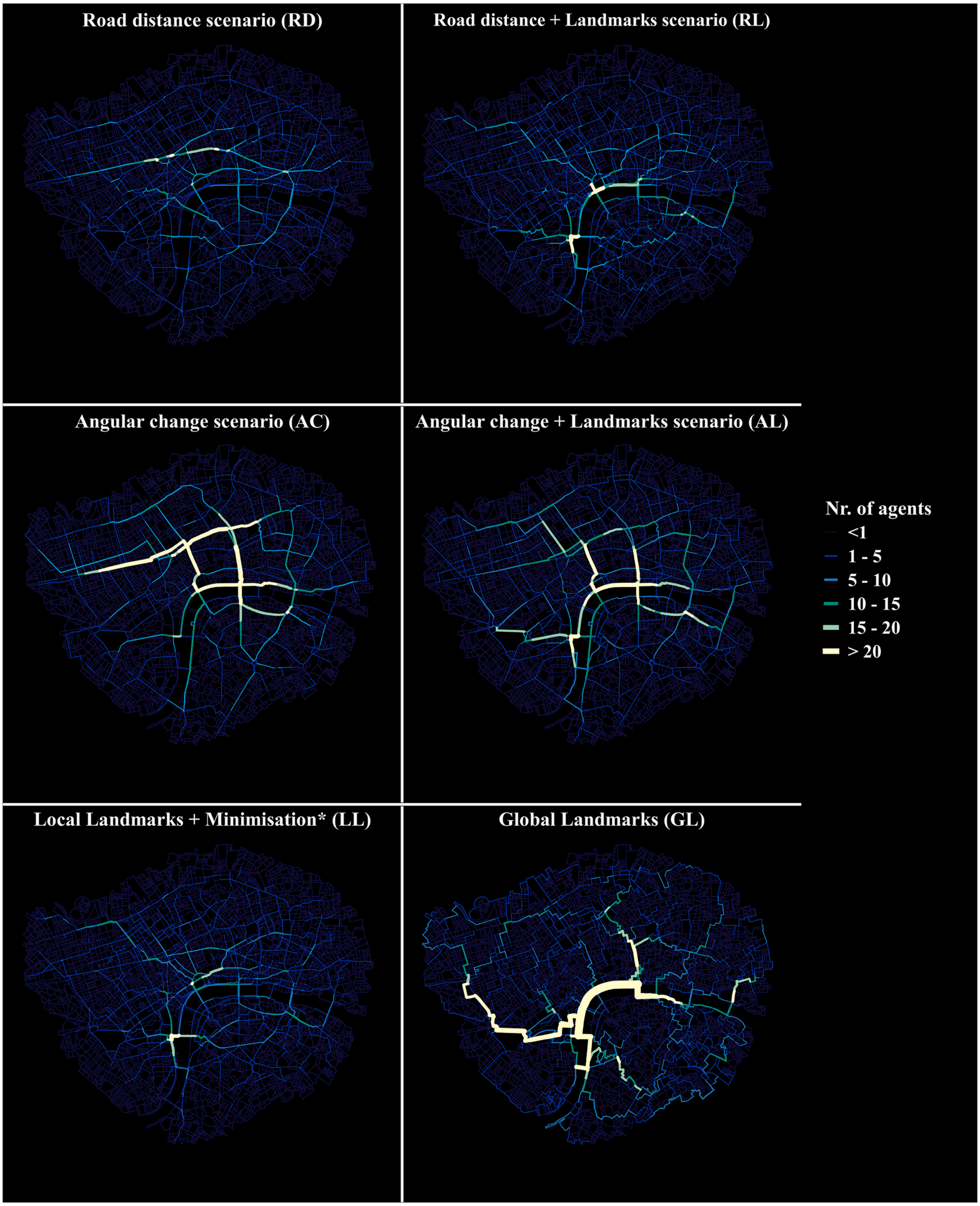



**Fig. 5.** Movement flows of agents on street segments resulting from the four ABM scenarios and the two sub-components of the landmark-based piloting route choice model (median across the runs). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

halved. In the east, Chiswell Street and other secondary roads around the Barbican Centre, a relevant landmark (used as an on-route mark in route 1, Fig. 6), attracted the agents that in the AC scenario would employ Clerkenwell Road. Conversely to the RL scenario, minor roads in The City were not traversed much by the AL agents, as they are not convenient from an angular perspective, especially around St. Paul's Cathedral.

Moreover, agents in the AL scenario, compared to the AC scenario, resorted less to Farringdon Street and its continuations (around 24 times in the AL scenario, up to 33 in the AC scenario); instead, they more evenly distributed in the east, while walking as much as in the AC scenario along Kingsway (up to 26 agents in both scenarios). More than in the RL scenario, the agents in the AL scenario wandered around the parks in the west (Hyde Park, St. James's Park, Buckingham Palace Park, Green Park) and along the north riverbank, even in the east (e.g. Paul's Walk, up to 32, 12%). Finally, similarly to the comparison between the

**Table 3**
Statistics of the ABM scenarios. Percentage values indicate the percentage of street segments over the entire network.

| | RD | RL | AC | AL |
|---|---|---|---|---|
| Mean routes' length | 3350.95 m | 3910.83 m | 3953.68 m | 4393.14 m |
| Mean nr. agents per street | 0.95 | 1.07 | 1.02 | 1.09 |
| Stdv. nr. agents per street | 1.91 | 2.05 | 3.39 | 2.87 |
| Max nr. agents per street | 21.5 | 34.5 | 35.0 | 32.0 |
| Nr. segments volumes $\geq 1$ | 5053 (39.0%) | 5421 (41.9%) | 3015 (23.3%) | 3686 (28.5%) |
| Nr. segments volmes $\geq 5$ | 605 (4.6%) | 713 (5.5%) | 770 (5.9%) | 920 (7.1%) |

**Table 4**
Correlation Matrix of the ABM scenarios' resulting volumes (Pearson product-moment correlation coefficients). GL and LL are two sub-components of the landmark-based piloting route choice model.

| | RD | RL | AC | AL | LL | GL |
|---|---|---|---|---|---|---|
| RD | – | 0.62 | 0.55 | 0.57 | 0.77 | 0.12 |
| RL | 0.62 | – | 0.41 | 0.64 | 0.81 | 0.34 |
| AC | 0.55 | 0.41 | – | 0.83 | 0.44 | 0.22 |
| AL | 0.57 | 0.64 | 0.83 | – | 0.66 | 0.35 |
| LL | 0.77 | 0.81 | 0.44 | 0.66 | – | 0.20 |
| GL | 0.12 | 0.34 | 0.22 | 0.35 | 0.20 | – |

RC and RL scenarios, a notable difference between the AC and the AL scenarios lies in the role of Parliament Square and the adjacent streets to and from the river, the parks and Lambeth Bridge.

On the whole, as compared to the divergence between the patterns emerging from the AC and AL scenarios, the contrast between the patterns emerging from the RD and RL scenarios is more tangible. We put forward two possible explanations. The first one relates to the nature of the cost that is minimised. Routes resulting from distance minimisation may be more dependent on the OD pair than the angular change approach. Adjacent pairs of origins and destinations may generate routes, perhaps similar in terms of direction or shape, that do not or barely overlap. On the contrary, angular change is more affected by the street morphology and more prone to produce converging routes and flows (as the low number of traversed segments suggests) (Dalton, 2001; Omer & Kaplan, 2019); even routes between pair of locations far away from each other share a considerable number of street segments, since certain roads, or systems of roads, are always appealing when minimising change of direction.

The second explanation regards the impact of the street morphology on the visibility of distant landmarks. The distant landmarks visibility maximisation component (GL in Table 4) has a low correlation with the RD scenario (0.12), but a moderate one with the AC scenario (0.35). Looking at the emerging patterns, one can see a certain overlap of the main lines of movement (high volumes) in the GL and AC scenarios, in particular along the north bank of the river and the boundaries of the parks in the west. It appears that roads that are advantageous in terms of angular change are also attractive in terms of visibility of distant landmarks. To a certain extent the correlation between these two scenarios may unveil that, at least in this case study area, cumulative angular change minimisation already conduces agents through junctions that allow keeping an eye on orienting landmarks.

Nevertheless, the AL scenario is featured by a wider range of traversed street segments. The Thames' south bank presents its highest volume of agents in this scenario. Yet, they still appear very low and unbalanced when considering the ABM flows along the north bank, as well as the large number of tourists and urban explorers who usually wander there. Finally, compared to the RL scenario, the AL scenario also produces patterns that are easier to interpret. The RL scenario generates flow patterns that are quite dispersed, and that are less insightful, especially in the south, in indicating concentrations of agents.

### 4.2. Model evaluation with observational data

At the macro-level, the global landmarkness accumulated along the routes of the AL scenario (median 7.53) is the most similar to the global landmarkness accumulated along the GPS trajectories (7.77) (see Fig. 7). The global landmarkness of the AL scenario's and trajectories' routes are indeed not significantly different when performing a Games-Howell post-hoc test (see appendix for the detailed results of the significance test, Tables B1 and B2). At the same time, the global landmarkness accumulated along the routes of the RL scenario is not significantly different from the trajectories (median: 8.91). As regards the local landmarkness, there are no significant differences between the RL and AL scenarios; local landmarkness values in the landmark-based scenarios are significantly lower than the trajectories' local landmarkness, but also significantly higher than the cost-based scenarios.

Although the trajectories' routes are longer than the routes of the ABM scenarios, the fact that the landmark-based scenarios, cumulative global landmarkness is not significantly different from the one resulting from the trajectories is noteworthy. In particular, one could argue that the observational sample is somehow composed of pedestrians who kept in sight orienting, distant landmarks as much as the agents of the AL and RL scenarios. Hence, the landmark-based route choice models seem to incorporate the "need of landmarkness" inherent to the observed routes.

Furthermore, while the RD and the RL scenarios display an evident overlap with the GPS trajectories for what concerns the number of segments crossed up to 7 agents (Fig. 8), the AC scenario presents a number of segments much lower than the observed trajectories, for the same frequencies of agents per segment. In particular, the RD and RL scenarios and the GPS trajectories exhibit a very similar number of street segments which were crossed up to one time (2490, 2477 and 2092 segments respectively, Fig. 8). Conversely, the AL scenario features the most similar distribution to the trajectories for frequencies between 7 and 12 agents. Generally speaking, the RD scenario, and to a certain extent the RL scenario, brings about distributions considerably divergent from the ones emerging from the GPS trajectories, especially for highly used segments.

At the meso-level, the routes of the 255 GPS trajectories generated a pattern that is at first glance quite different from the outcomes of the ABM scenarios (Fig. 9). The pedestrians of the GPS trajectories moved mainly along the south bank of the river (up to 33 individuals, 13%), making use of Tower Bridge (31), Millennium Bridge (24) and Westminster Bridge (21). Whitehall was crossed by up to 24 pedestrians between Charing Cross and Parliament Square. Remarkably, a large number of pedestrians moved along or through the parks (e.g. Constitution Hill, The Mall, from 15 to 26 trajectories), between The Strand and St. Paul's (Fleet Street, 15 pedestrians, 6%), through Oxford Street (peaks of 15 agents), across Soho and the south part of The City. The north bank was traversed only in the centre of the case study area (Victoria Embankment, 14 trajectories). Extremely low volumes characterise both the north (apart from Marylebone Road, in the vicinity of The Regent's Park) and the south, crossed by just a bunch of trajectories.

The RMSE indicates that the RD and the RL scenarios produce the lowest error in estimating the number of agents per street segment (RMSE 2.8). On the other hand, the AC and AL scenarios exhibit higher RMSE; however, in this case, the error is reduced by the incorporation of

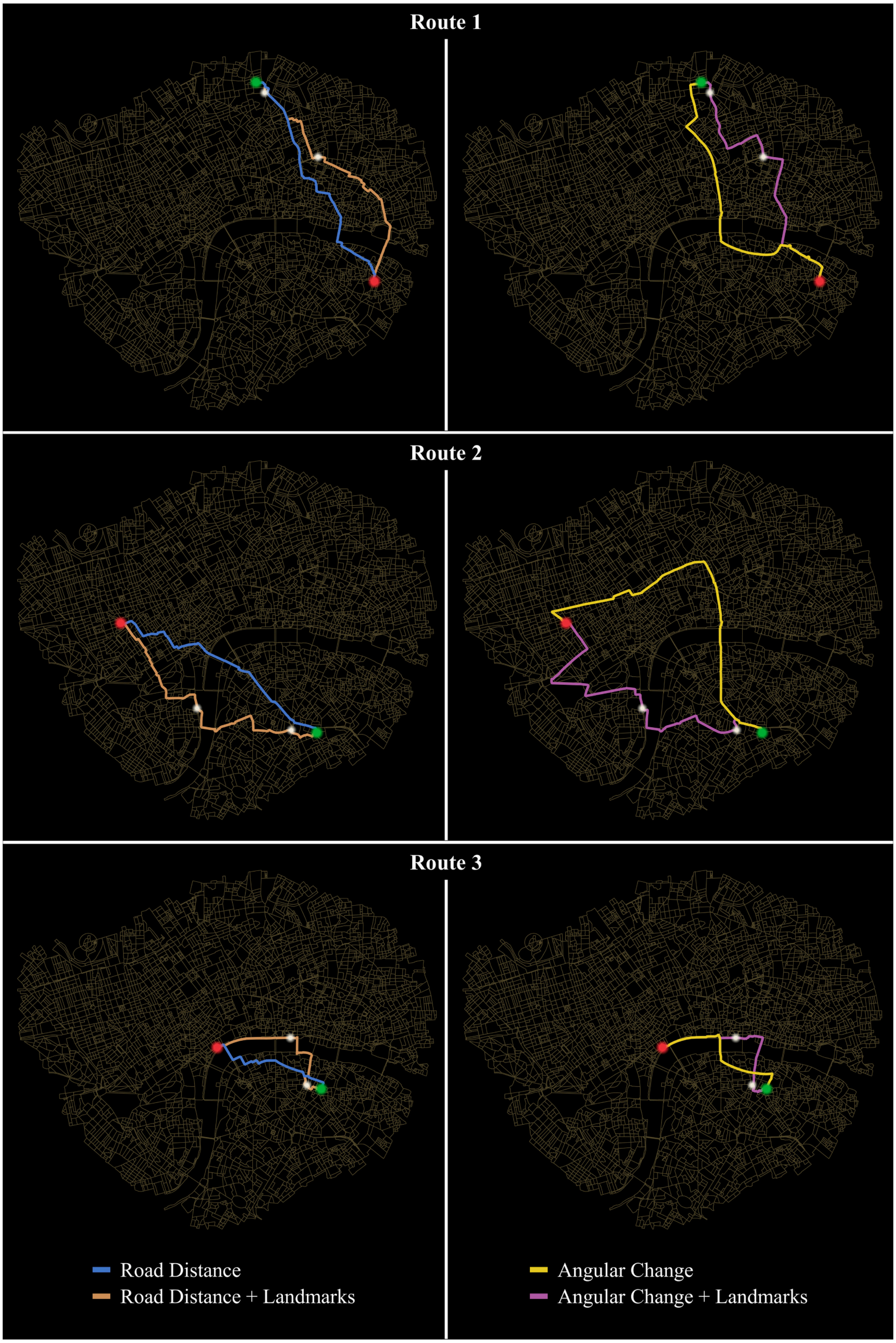



**Fig. 6.** Examples of individual routes generated in the four ABM scenarios: RD vs RL and AC vs AL. Red locations represent the origins of the routes, green locations the destinations, white locations intermediate decision points chosen for their local landmarkness. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

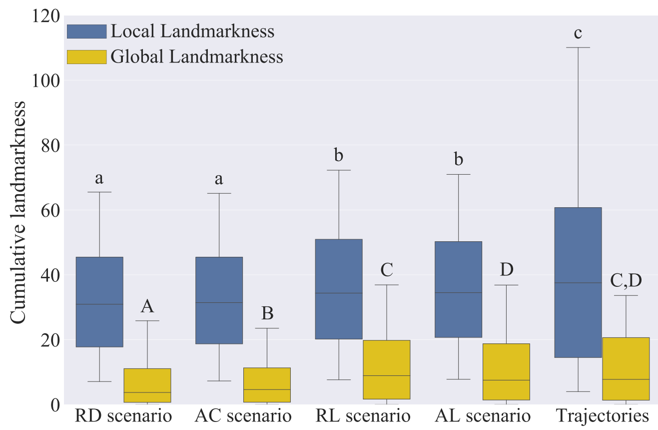



**Fig. 7.** Cumulative local and global landmarkness along the routes of the ABM scenarios and the trajectories. Different capital (global landmarkness) and lowercase (local landmarkness) letters indicate statistically significant differences between scenarios (Games-Howell post-hoc test). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

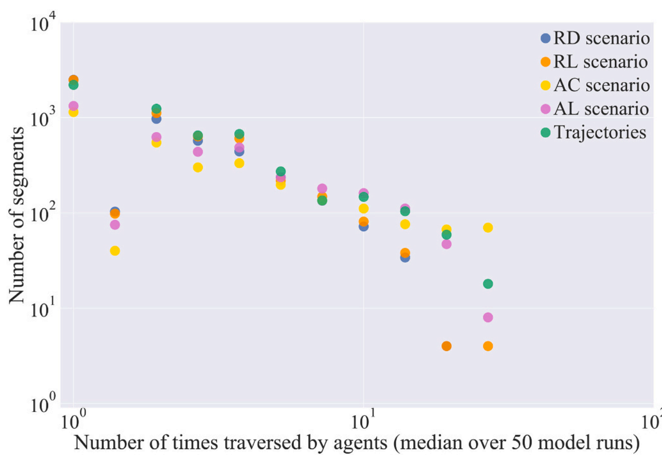



**Fig. 8.** Frequency distribution of pedestrian volumes across street segments in the four scenarios. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

landmarks from 3.8 to 3.2 (Fig. 10). At the street segment level, the error - expressed in number of standard deviations that the modelled median volumes are away from the GPS trajectories volumes - is low or equal to 0 in most of the areas (street segments coloured in dark grey, Fig. 10). The landmark-based scenarios led a high number of agents towards the river, as in the GPS trajectories. Yet, the north bank is featured by overestimation and the south by underestimation of volumes. The error is indeed remarkably high, in a negative direction, along street segments in the south bank of the river and inside the parks.

More in detail, the AC scenario overestimated the number of agents along major roads such as Theobalds Road, Clerkenwell Road, Farringdon Street and Blackfriars Road (error between 4 and 6 standard deviations) and nearby King's Cross, Euston Road (3 to 4 standard deviations); this can be partly observed also in the AL scenario (e.g. Kingsway, up to 5 standard deviations) and both in the AC and AL scenarios along Southwark Street (5 and 4 standard deviations). Therefore, in these two scenarios, the role of Blackfriars Bridge resulted excessively highlighted (6 standard deviations in the AC scenario, 4 in the AL scenario). These are all multi-lane roads, highly congested and not particularly attractive, in reality, for pedestrians. Nevertheless, although the importance of Tower Bridge (AC: −11, AL: −12 standard deviations) and Millennium Bridge (AC, AL: −16 standard deviations) was not captured, the AL scenario brought about volumes alike the trajectories along Constitution Hill (16 agents against 18 trajectories, 0 standard deviations) and partly along the south bank, (The Queen's Walk, 14 agents, 22 trajectories, 0 to 2 standard deviations). The RD and AC scenarios generated patterns comparable to the GPS trajectories along Oxford Street, one of the few major roads actually crossed by the pedestrians. What stands out from this analysis is, however, the overestimated volumes along the north bank of the Thames, in the AL, RL and AC scenarios. These scenarios, to some extent, "misplaced" the agents walking along the river.

Overall, some insights are of interest in relation to pedestrian route choice behaviour. The least cumulative angular change route choice model led agents, in absence of other constraints or rules, to widely use major roads. Although this mechanism has proven to be a valuable predictor for vehicular traffic analysis (cars, bus and taxi flows) (e.g. Manley, Orr, & Cheng, 2015), it is probably less determinant for pedestrians, at least for large case study areas (see route 2, Fig. 6, which appears implausible for a pedestrian, but realistic for car drivers). Several studies have indeed reported promising results in small urban areas or towns (e.g. Omer & Kaplan, 2017). We have shown here that these conclusions are less valid for large cities like London, and that the incorporation of landmarks helps to partly overcome the tendency of agents minimising cumulative angular change to rely on major roads. In addition, parks and rivers partly emerged as movement attractors in the landmark-based scenarios.

However, on the one hand, none of the scenarios here evaluated fully generated volumes alike the trajectories within the parks; on the other hand, the introduction of landmarks did not contribute much in identifying the south riverbank as more prone to be traversed, despite the fact that several touristic attractions and landmarks are located along this side of the Thames. Additionally, whereas the streetscape of the south bank is specifically designed for pedestrians, the north bank's footways are narrower, often interrupted by traffic lights and major thoroughfare. Regardless the distribution of agents between the south and the north bank, we can claim that the modelled movement flows at least make the river emerge as a movement attractor, consistently with the trajectories' volumes.

### 4.3. Limitations and future work

Our approach presents several limitations as pertains a) the formalisation of route choice behaviour; b) the characteristics of the agents; c) the observational dataset employed to evaluate the outcomes of the model. To begin with, several components, which have been proven to intervene in human wayfinding and route choice processes, were neglected in our work. Barriers (or edges in the Lynchian definition) – sometimes formalised as particular kinds of landmarks (Tomko & Winter, 2013) – such as rivers, parks, or obstructing structures, may lead people to restructure their path, avoid specific junctions, or walk along a natural delimiting element. Their role partially emerged across the model in the landmark-based scenarios (the boundaries of the parks and the north-bank of the Thames), but the fact that we did not explicitly include them in our route choice models becomes evident when considering the ABM mispredictions around such dividing urban elements (the Thames itself, the parks in the west and in the north, etc.). Therefore, this work can be expanded either by explicitly modelling the barrier effect on spatial behaviour in pedestrian simulation models, or extending the concept of landmarks to natural elements; in our model only buildings can indeed constitute landmarks.

Moreover, in other works, regions (or districts) have been incorporated in pedestrian and vehicular traffic simulation models (see Filomena, Manley, & Verstegen, 2019; Manley et al., 2015), to account for, findings on the influence of regions on route choice behaviour (Wiener & Mallot, 2003) and the hierarchical organisation of spatial knowledge (e.g. Hirtle & Jonides, 1985; Stevens & Coupe, 1978). There, the

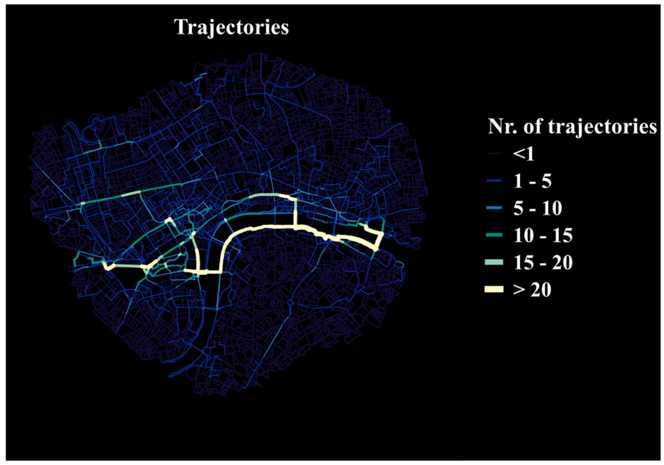



**Fig. 9.** Movement flows of pedestrians on street segments resulting from 255 GPS trajectories (GPSies.com). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

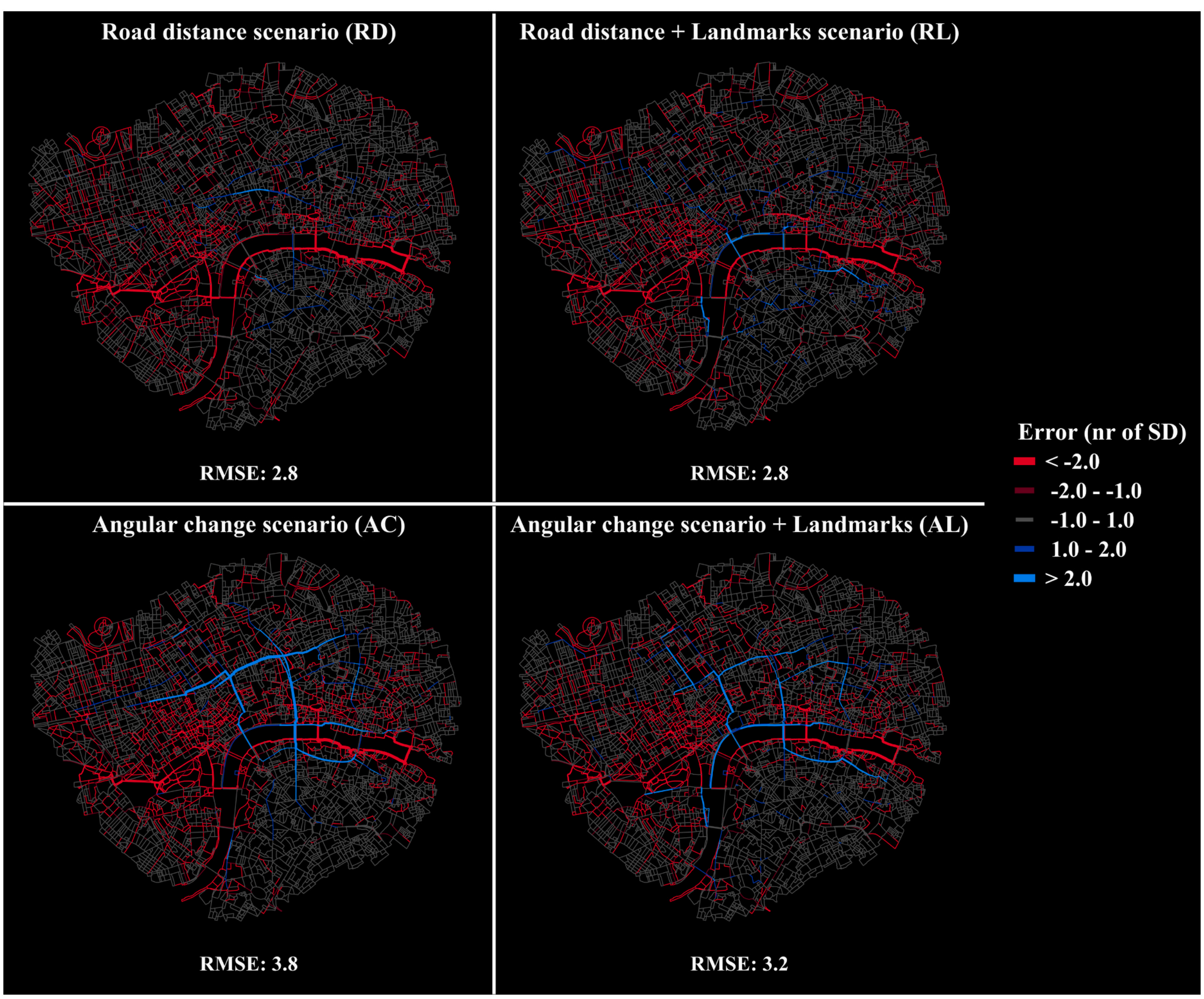



**Fig. 10.** Difference from the volumes of the GPS trajectories and the four ABM scenarios, expressed in standard deviation per each street segment. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

perceived division of an urban area in regions is used by the agents to formulate a rough plan before making use of other street attributes. A further direction could explore the combination of a region-based navigation approach and the landmark-based model here described.

Secondly, in our model, the agents are assumed to all have the same knowledge of the urban environment. We did not incorporate individual differences either in terms of agents' goals or in terms of spatial experience. However, pedestrians and urban explorers make use of different elements on the basis of their own experience and their knowledge. Certain roads or certain buildings are known only to few individuals and may be therefore avoided or not considered by others when formulating a route. The introduction of spatial knowledge differences has been explored in an ABM for traffic simulation by employing a hierarchisation of the street network and a locally defined agent's knowledge (Manley & Cheng, 2018). Furthermore, attributes regarding street layout, sidewalks' width, presence of green elements or areas, and street lighting, aspects which boil down to the so-called walkability of a street (see Forsyth, 2015; Lo, 2009), were not represented in our model. Yet these aspects do influence pedestrians' choices and, in turn, the roads that are more likely to be walked. Whereas their explicit representation in pedestrian simulation is not trivial, certain road attributes could be used to infer the suitability of certain street segments for pedestrians over other road users. Yet, it should also be considered that a model is by definition a simplification of reality: abstraction should be balanced with realism. As such, it is not desirable to include all possible aspects of human behaviour into modelling frameworks in great detail.

Finally, the observational set of trajectories presents limitations both in terms of quantity – just 255 routes over a large urban area – and, even more importantly, in terms of what type of movement they represent. The average length of the trajectories, 5107.14 m, may suggest that some of the routes, even though categorised within pedestrian descriptors or tags, constitute paths driven by leisure or outdoor activities. The composition of the sample is likely to represent a very small portion of walkers, namely people who are willing to track their routes and share them on GPS trajectories platforms. In addition, the distribution of the trajectories across the case study area is uneven. For example, the trajectories in the south of the case study area do not seem to be representative of pedestrian movement, as only two of them traverse this region. A more complete observational dataset of movement flows, in terms of spatial coverage as well as type of individual (e.g. not only GPS trajectories enthusiasts), would allow a more precise evaluation of the model. Volumes collected via multiple footfall sensors or by manually counting pedestrians across a large number of streets (e.g. Omer & Kaplan, 2017) might support calibration and validation processes.

## 5. Conclusion

The aim of this work was to devise an agent-based model (ABM) of pedestrian movement in urban areas enriched with salient information about the distribution of landmarks. A landmark-based piloting route choice model, built on existing theories and empirical evidence in spatial navigation research (Epstein & Vass, 2014; Lynch, 1960; Siegel & White, 1975), was developed so as to represent complex cognitive processes in human route choice behaviour. The function of on-route marks (local landmarks) and distant landmarks (global landmarks) were modelled through a graph representation of the urban environment. Four scenarios were conceived within the ABM to represent classes of agents endowed with different route choice models: a) road distance minimisation b) road distance minimisation and landmarks, c) least cumulative angular change, d) least cumulative angular change and landmarks.

We sought to answer the following research questions a) To what extent does the incorporation of landmarks in pedestrian agents' route choice behaviour lead to different macro-level patterns as compared to pure minimisation models, i.e. shortest road-distance or least cumulative angular change paths? b) To what degree does the inclusion of a landmark-based route choice model in an ABM for pedestrian simulation contribute to generating more realistic distributions of pedestrians across an urban space? A set of GPS trajectories was used to evaluate the outcomes of the simulation. The least cumulative angular change scenario, widely used in existing simulation models, led agents to make use of too few channels of movement, disregarding minor roads; it caused, therefore, overestimation along main thoroughfares. The introduction of landmarks in the agents' route choice behaviour drove them to resort more extensively to minor roads and novel paths. The landmark-based scenarios brought about higher volumes along the river and parks and more heterogeneity in the distribution of the agents across the case study area. Moreover, on the one hand, the landmark-based piloting approach generated routes whose cumulative global landmarkness was the most similar to the observed routes'; on the other hand, although the landmark-based scenarios showed a marked improvement, none of the ABM scenarios resulted in routes alike the observed ones, in terms of local landmarkness. At the street level, the landmark-based scenarios underestimated the volumes along the south-bank of the river but did identify the river as a movement attractor.

The incorporation of a landmark-based route choice model in pedestrian simulation offers insights regarding movement and concentration of pedestrians at relevant decision points. This knowledge may encourage modellers, as well as policy makers, to reconsider pedestrian navigation and the intertwined interaction between navigation, cognitive representations of space, and urban elements. Finally, the landmark-based route choice approach could be further enhanced to allow for individual differences in spatial knowledge and demographic characteristics.

## Declaration of Competing Interest

None.

## Appendix A. Identification of decision points in the landmark-based piloting approach

The following steps describe in detail the selection of decision points within the ABM:

1. *Origin and destination identification*

The agent is assigned with a pair of origin and destination nodes. The agent is placed at the origin.

2. *Wayfinding complexity estimation*

The wayfinding complexity $w$ of the space defined by the circle $K$, formed around the diameter between the current location $c$ and the destination $d$, is computed (see Fig. 3a) as:

$$w = \left(\frac{e_{c,d}}{max_e}*0.5 + \frac{X_K - X_K(L)}{X_K}*0.5\right) \tag{A1}$$

where $e_{c,d}$ (meters) is the Euclidean distance between the current location and the destination $d$; $max_e$ (meters) represents the maximum walkable Euclidean distance within the city. $X_K$ is the number of buildings within $K$; $X_K(L)$ corresponds to the number of buildings in the same area that are also considered local landmarks[4], namely with a local salience index $l_b > y$, as defined above. The value of $w$ determines whether the agent needs landmarks guidance to continue towards the destination, or if it is able to move based on a minimisation heuristic towards the destination. When $w$ is higher than a threshold $u$, defined above, the agent moves to the next step and defines a cognisable search-space; otherwise, the agent completes the path through the decision points identified so far (if any), towards the destination, by employing a cost minimisation heuristic.

3. *Search-space definition*

The wayfinding complexity of the environment affects the search capability of the agent within the circle $K$. Given the location of the agent and the destination, the complexity of the route lets the agent cognise and take decisions only within a certain space, that can be considered a chunk of the entire route. Such a *search-space* is the portion of the circle $K$ that is within a *search-distance* ($z$, meters) from the current location (see Fig. 3b); the search-distance $z$ (meters) is computed as:

$$z = e_{c,d}*(1-w) \tag{A2}$$

where $(1 - w)$ is the inverse of the wayfinding complexity and indicates the *easiness* to come up with a defined route between the current location and the destination. The agent's cognitive map of this portion of space contains information about known street junctions and landmarks, and it is used to identify decision points in the next step.

4. *Identification of possible decision points within the search-space*

When looking for decision points, the agent does not examine all possible nodes within its search-space, but only nodes that represent known junctions (see Fig. 3b). Known junctions are here identified on the basis of a shortest path betweenness centrality measure computed on nodes in a primal graph representation of the street network (Filomena, Verstegen, & Manley, 2019; Porta, Crucitti, & Latora, 2006), assuming that high centrality values indicate the likelihood of a node to be cognitively represented. The set of known junctions $N$ is defined by selecting the nodes whose betweennes centrality value is above the 75*th* percentile, within $K_{c,d}$.

5. *Decision points assessment: Local salience of landmarks at the junction*

At this stage the agent evaluates the attractiveness of a possible intermediate node, amongst the junctions identified in the previous stage, as regards the functions associated with on-route marks (see Sections 2.2 and 3.2 for references and definition). For each known junction $n$, the local salience $l$ of its adjacent buildings $B_n$ is employed. Thus, to reward decision points featured by salient buildings, the local salience of a decision point (node $n$) is defined as:

$$k_n = max(l_b) \quad \text{with } n \in N, \;\; b \in B_n \tag{A3}$$

The score $k_n$ is the highest local salience score amongst the buildings adjacent to a known junction $n$.

6. *Final assessment: Local salience and distance gain*

For each decision point candidate $n$ in $N$, within the search-space, a measure of attractiveness is finally obtained as (see Eq. 4 in the main text):

$$a_n = k_n*\rho + \frac{e_{c,d} - e_{n,d}}{e_{c,d}}*\tau \tag{A4}$$

where $k_n$ represents the landmark local salience of the node $n$, on the basis of the set $B_n$, $e_{c,\ d}$ is the Euclidean distance between the current location of the agent and the destination $d$, and $e_{n,d}$ the Euclidean distance between the node $n$ and $d$; $\rho$ and $\tau$ are the weights assigned to the two components. As across the rest of this planning process, the distance gain is assessed based on a Euclidean distance heuristic that does not take into account possible barriers and detours caused by the street morphology.

7. *Intermediate decision point selection*

The node with the highest attractiveness value $a$ amongst the known junctions $N$ is picked and added to a sequence of intermediate decision points (nodes). The process continues from step 2 for this new decision point, where the agent computes the wayfinding complexity of the space between the

[4] In this equation, global landmarks could be used instead, for example when representing agents with low experience in the case study area. In fact, given the lower number of global landmarks across the urban environment, the increase in the perceived wayfinding complexity may induce the agent to identify more intermediate decision points.

new location and the destination.

## Appendix B. Cumulative landmarkness: significance test values

**Table B1**
Results of the significance test (Games-Howell post-hoc test) on cumulative local landmarkness computed on the ABM scenarios, and trajectories, routes. *Mean(A)* indicates the mean of the first term of comparison, *Mean(B)* the mean of the second term of comparison. *Diff* is the difference between the mean values of the compared scenarios. *T* represents the difference in units of standard error. *p-Value* indicates the significance level.

| | | Mean (A) | Mean (B) | Diff | *T* | *p*-Value |
|---|---|---|---|---|---|---|
| AC scenario | AL scenario | 33.009 | 36.415 | −3.406 | −14.484 | 0.001 |
| AC scenario | RD scenario | 33.009 | 32.684 | 0.325 | 1.443 | 0.499 |
| AC scenario | RL scenario | 33.009 | 36.460 | −3.451 | −14.543 | 0.001 |
| AC scenario | Trajectories | 33.009 | 43.684 | −10.675 | −4.664 | 0.001 |
| AL scenario | RD scenario | 36.415 | 32.684 | 3.731 | 15.729 | 0.001 |
| AL scenario | RL scenario | 36.415 | 36.460 | −0.046 | −0.183 | 0.900 |
| AL scenario | Trajectories | 36.415 | 43.684 | −7.269 | −3.174 | 0.013 |
| RD scenario | RL scenario | 32.684 | 36.460 | −3.776 | −15.778 | 0.001 |
| RD scenario | Trajectories | 32.684 | 43.684 | −11.000 | −4.806 | 0.001 |
| RL scenario | Trajectories | 36.460 | 43.684 | −7.224 | −3.154 | 0.014 |

**Table B2**
Results of the significance test (Games-Howell post-hoc test) on cumulative global landmarkness computed on the ABM scenarios, and trajectories, routes. *Mean(A)* indicates the mean of the first term of comparison, *Mean(B)* the mean of the second term of comparison. *Diff* is the difference between the mean values of the compared scenarios. *T* represents the difference in units of standard error. *p-V*alue indicates the significance level.

| | | Mean (A) | Mean (B) | Diff | *T* | *p*-Value |
|---|---|---|---|---|---|---|
| AC scenario | AL scenario | 7.734 | 11.617 | −3.883 | −28.044 | 0.001 |
| AC scenario | RD scenario | 7.734 | 7.286 | 0.448 | 3.799 | 0.001 |
| AC scenario | RL scenario | 7.734 | 12.402 | −4.668 | −33.460 | 0.001 |
| AC scenario | Trajectories | 7.734 | 11.955 | −4.222 | −5.126 | 0.001 |
| AL scenario | RD scenario | 11.617 | 7.286 | 4.331 | 31.920 | 0.001 |
| AL scenario | RL scenario | 11.617 | 12.402 | −0.785 | −5.070 | 0.001 |
| AL scenario | Trajectories | 11.617 | 11.955 | −0.338 | −0.410 | 0.900 |
| RD scenario | RL scenario | 7.286 | 12.402 | −5.116 | −37.411 | 0.001 |
| RD scenario | Trajectories | 7.286 | 11.955 | −4.670 | −5.673 | 0.001 |
| RL scenario | Trajectories | 12.402 | 11.955 | 0.447 | 0.540 | 0.900 |

## References

Allen, G. L. (1999). Spatial abilities, cognitive maps, and wayfinding - bases for individual differences in spatial cognition and behavior. In R. G. Golledge (Ed.), *Wayfinding behavior: Cognitive mapping and other spatial processes* (pp. 46–80). Baltimore, London: Johns Hopkins University Press.

Allen, G. L., & Golledge, R. G. (2007). Wayfinding in urban environments. In T. Gärling, & L. Steg (Eds.), *Threats from car traffic to the quality of urban life* (pp. 79–101). Amsterdam: Elsevier. https://doi.org/10.1108/9780080481449.

Appleyard, D. (1969). Why buildings are known. *Environment and Behavior, 1*(2), 131–156. https://doi.org/10.1177/001391656900100202.

Blanchard, P., & Volchenkov, D. (2009). *Mathematical analysis of urban spatial networks*. Berlin, Heidelberg: Springer.

Caduff, D., & Timpf, S. (2005). The landmark spider: Representing landmark knowledge for wayfinding tasks. In *AAAI Spring Symposium: Reasoning with mental and external diagrams: Computational modeling and spatial assistance* (pp. 30–35).

Chown, E., Kaplan, S., & Kortenkamp, D. (1995). Prototypes, location, and associative networks (PLAN): Towards a unified theory of cognitive mapping. *Cognitive Science, 19*(1), 1–51. https://doi.org/10.1207/s15516709cog1901_1.

Couclelis, H., Golledge, R. G., Gale, N., & Tobler, W. (1987). Exploring the anchor-point hypothesis of spatial cognition. *Journal of Environmental Psychology, 7*(2), 99–122. https://doi.org/10.1016/S0272-4944(87)80020-8.

Dalton, N. S. (2001). *Fractional configurational analysis and a solution to the Manhattan problem* (pp. 1–26). Atlanta, GA: Proceedings of the 3rd International Space Syntax Symposium.

Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische Mathematlk, 1*, 269–271.

Duckham, M., & Kulik, L. (2003). "Simplest" paths: Automated route selection for navigation. In W. Kuhn, M. Worboys, & S. Timpf (Eds.)*, Vol. 2825. Spatial information theory. Foundations of geographic information science. COSIT 2003. Lecture notes in computer science* (pp. 169–185). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-540-39923-0_12.

Duckham, M., Winter, S., & Robinson, M. (2010). Including landmarks in routing instructions. *Journal of Location Based Services, 4*(1), 28–52. https://doi.org/10.1080/17489721003785602.

Elias, B., & Sester, M. (2006). Incorporating landmarks with quality measures in routing procedures. In M. M. Raubal, H. J. Miller, A. U. Frank, & M. F. Goodchild (Eds.), *Geographic information science. GIScience 2006. Lecture notes in computer science, vol 4197* (pp. 65–80). Berlin, Heidelberg: Springer. https://doi.org/10.1007/11863939_5.

Epstein, R. A., Patai, E. Z., Julian, J. B., & Spiers, H. J. (2017). The cognitive map in humans: Spatial navigation and beyond. *Nature Neuroscience, 20*(11), 1504–1513. https://doi.org/10.1038/nn.4656.

Epstein, R. A., & Vass, L. K. (2014). Neural systems for landmark-based wayfinding in humans. *Philosophical Transactions of the Royal Society, B: Biological Sciences, 369* (1635). https://doi.org/10.1098/rstb.2012.0533.

Evans, G. W., Smith, C., & Pezdek, K. (1982). Cognitive maps and urban form. *Journal of the American Planning Association, 48*(2), 232–244. https://doi.org/10.1080/01944368208976543.

Filomena, G., Manley, E., & Verstegen, J. A. (2019). Route choice through regions by pedestrian agents. In S. Timpf, C. Schlieder, M. Kattenbeck, B. Ludwig, & K. Stewart (Eds.), *COSIT 2019, 14th international conference on spatial information theory* (pp. 1–5). Dagstuhl: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. https://doi.org/10.4230/LIPIcs.COSIT.2019.5.

Filomena, G., Verstegen, J. A., & Manley, E. (2019). A computational approach to "the image of the City". *Cities, 89*, 14–25. https://doi.org/10.1016/j.cities.2019.01.006.

Filomena, Gabriele (2020). PedSimCity-Modelling_Landmarks (Version v1.09). Zenodo. https://doi.org/10.5281/zenodo.4308591.

Filomena, Gabriele (2020). PedSimCity-Evaluation (Version v1.02). Zenodo. https://doi.org/10.5281/zenodo.4308442.

Foo, P., Warren, W. H., Duchon, A., & Tarr, M. J. (2005). Do humans integrate routes into a cognitive map? Map- versus landmark-based navigation of novel shortcuts. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 31*(2), 195–215. https://doi.org/10.1037/0278-7393.31.2.195.

Forsyth, A. (2015). What is a walkable place? The walkability debate in urban design. *Urban Design International, 20*(4), 274–292. https://doi.org/10.1057/udi.2015.22.

Gallistel, C. R. (1990). Representations in animal cognition: An introduction. *Cognition, 37*(1–2), 1–22. https://doi.org/10.1016/0010-0277(90)90016-D.

Games, P. A., & Howell, J. F. (1976). Pairwise multiple comparison procedures with unequal N's and/or variances: A Monte Carlo study. *Journal of Educational Statistics, 1*(2), 113. https://doi.org/10.2307/1164979.

Garling, T., Book, A., & Lindberg, E. (1984). Cognitive mapping of large-scale environments: The interrelationship of action plans, acquisition, and orientation. *Environment and Behavior, 16*(1), 3–34. https://doi.org/10.1177/0013916584161001.

Gärling, T., Säisä, J., Book, A., & Lindberg, E. (1986). The spatiotemporal sequencing of everyday activities in the large-scale environment. *Journal of Environmental Psychology, 6*(4), 261–280. https://doi.org/10.1016/S0272-4944(86)80001-9.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston, MA: Houghton Mifflin.

Gillner, S., & Mallot, H. A. (1998). Navigation and acquisition of spatial knowledge in a virtual maze. *Journal of Cognitive Neuroscience, 10*(4), 445–463. https://doi.org/10.1162/089892998562861.

Gluck, M. (1991). Making sense of human wayfinding: Review of cognitive and linguistic knowledge for personal navigation with a new research direction. In D. M. Mark, & A. U. Frank (Eds.), *NATO ASI series (Series D: Behavioural and social sciences): Vol. 63. Cognitive and linguistic aspects of geographic space* (pp. 117–135). Dordrecht: Springer. https://doi.org/10.1007/978-94-011-2606-9_9.

Golledge, R. G. (1992). Do people understand spatial concepts: The case of first-order primitives. In A. U. Frank, I. Campari, & U. Formentini (Eds.)*, Vol. 639. Theories and methods of Spatio-temporal reasoning in geographic space. Lecture notes in computer science* (pp. 1–21). Berlin, Heidelberg: Springer. https://doi.org/10.1007/3-540-55966-3_1.

Golledge, R. G. (1995). Path selection and route preference in human navigation: A progress report. In A. U. Frank, & W. Kuhn (Eds.)*, Vol. 1329. Spatial information theory a theoretical basis for GIS. COSIT 1997. Lecture notes in computer science* (pp. 207–222). Berlin: Springer. https://doi.org/10.1007/3-540-60392-1_14.

Golledge, R. G. (1999). Human wayfinding and cognitive maps. In R. G. Golledge (Ed.), *Wayfinding behavior: Cognitive mapping and other spatial processes* (pp. 5–45). Baltimore, London: Johns Hopkins University Press.

Grieves, R. M., & Jeffery, K. J. (2017). The representation of space in the brain. *Behavioural Processes, 135*, 113–131. https://doi.org/10.1016/j.beproc.2016.12.012.

Hafting, T., Fyhn, M., Molden, S., Moser, M. B., & Moser, E. I. (2005). Microstructure of a spatial map in the entorhinal cortex. *Nature, 436*(7052), 801–806. https://doi.org/10.1038/nature03721.

Haklay, M., O'Sullivan, D., Thurstain-Goodwin, M., & Schelhorn, T. (2001). "So go downtown": Simulating pedestrian movement in town Centres. *Environment and Planning. B, Planning & Design, 28*(3), 343–359. https://doi.org/10.1068/b2758t.

Hayes-Roth, B., & Hayes-Roth, F. (1979). A cognitive model of planning. *Cognitive Science, 3*(4), 275–310. https://doi.org/10.1016/S0364-0213(79)80010-5.

Heft, H. (1996). The ecological approach to navigation: A Gibsonian perspective. In J. Portugali (Ed.), *The construction of cognitive maps* (pp. 105–132). Dordrecht: Kluwer Academic.

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

Heth, D. C., Cornell, E. H., & Alberts, D. M. (1997). Differential use of landmarks by 8- and 12-year-old children during route reversal navigation. *Journal of Environmental Psychology, 17*(3), 199–213. https://doi.org/10.1006/jevp.1997.0057.

Hillier, Bill, & Iida, Shinichi (2005). Network effects and psychological effects: A theory of urban movement. In A. G. Cohn, & D. M. Mark (Eds.), *Spatial Information Theory. COSIT 2005. Lecture Notes in Computer Science, vol 3693* (pp. 553–564). Berlin, Heidelberg: Springer-Verlag. https://doi.org/10.1007/11556114_30.

Hirtle, S. C., & Jonides, J. (1985). Evidence of hierarchies in cognitive maps. *Memory & Cognition, 13*(3), 208–217. https://doi.org/10.3758/BF03197683.

Hölscher, C., Tenbrink, T., & Wiener, J. M. (2011). Would you follow your own route description? Cognitive strategies in urban route planning. *Cognition, 121*(2), 228–247. https://doi.org/10.1016/j.cognition.2011.06.005.

Ishikawa, T., & Montello, D. R. (2006). Spatial knowledge acquisition from direct experience in the environment: Individual differences in the development of metric knowledge and the integration of separately learned places. *Cognitive Psychology, 52*(2), 93–129. https://doi.org/10.1016/j.cogpsych.2005.08.003.

Jiang, B., & Jia, T. (2011). Agent-based simulation of human movement shaped by the underlying street structure. *International Journal of Geographical Information Science, 25*(1), 51–64. https://doi.org/10.1080/13658811003712864.

Kitchin, R. M. (1996). Increasing the integrity of cognitive mapping research: Appraising conceptual schemata of environment-behaviour interaction. *Progress in Human Geography, 20*(1), 56–84. https://doi.org/10.1177/030913259602000104.

Kuipers, B. J., Tecuci, D. G., & Stankiewicz, B. J. (2003). The skeleton in the cognitive map: A computational and empirical exploration. *Environment and Behavior, 35*(1), 81–106. https://doi.org/10.1177/0013916502238866.

Lo, R. H. (2009). Walkability: What is it? *Journal of Urbanism, 2*(2), 145–166. https://doi.org/10.1080/17549170903092867.

Lovelace, K. L., Hegarty, M., & Montello, D. R. (1999). Elements of good route directions in familiar and unfamiliar environments. In C. Freksa, & D. M. Mark (Eds.)*, Vol. 1661. Spatial information theory. Cognitive and computational foundations of geographic information science. COSIT 1999. Lecture notes in computer science* (pp. 65–82). Springer: Berlin, Heidelberg. https://doi.org/10.1007/3-540-48384-5_5.

Löwen, H., Krukar, J., & Schwering, A. (2019). Spatial learning with orientation maps: The influence of different environmental features on spatial knowledge acquisition. *ISPRS International Journal of Geo-Information, 8*(3), 149. https://doi.org/10.3390/ijgi8030149.

Lynch, K. (1960). *The image of the city*. Cambridge, MA: MIT Press.

Manley, E., & Cheng, T. (2018). Exploring the role of spatial cognition in predicting urban traffic flow through agent-based modelling. *Transportation Research Part A: Policy and Practice, 109*, 14–23. https://doi.org/10.1016/j.tra.2018.01.020.

Manley, E., Orr, S., & Cheng, T. (2015). A heuristic model of bounded route choice in urban areas. *Transportation Research Part C: Emerging Technologies, 56*, 195–209. https://doi.org/10.1016/j.trc.2015.03.020.

Marshall, S., Gil, J., Kropf, K., Tomko, M., & Figueiredo, L. (2018). Street network studies: From networks to models and their representations. *Networks and Spatial Economics, 18*(3), 735–749. https://doi.org/10.1007/s11067-018-9427-9.

Michon, P.-E., & Denis, M. (2001). When and why are visual landmarks used in giving directions? In D. R. Montello (Ed.)*, Vol. 2205. Spatial information theory. COSIT 2001. Lecture notes in computer science* (pp. 292–305). Berlin: Springer. https://doi.org/10.1007/3-540-45424-1_20.

Montello, D. R. (1997). The perception and cognition of environmental distance: Direct sources of information. In S. C. Hirtle, & A. U. Frank (Eds.)*, Vol. 1329. Spatial information theory a theoretical basis for GIS. COSIT 1997. Lecture notes in computer science* (pp. 297–311). Berlin: Springer. https://doi.org/10.1007/3-540-63623-4_57.

Montello, D. R. (1998). A new framework for understanding the acquisition of spatial knowledge in large-scale environments. In M. J. Egenhofer, & R. G. Golledge (Eds.), *Spatial and temporal reasoning in geographic information systems* (pp. 143–154). New York: Oxford University Press.

Muller, R. U., Stead, M., & Pach, J. (1996). The hippocampus as a cognitive graph. *The Journal of General Physiology, 107*(6), 663–694. https://doi.org/10.1085/jgp.107.6.663.

Nadel, L. (1990). Varieties of spatial cognition. *Annals of the New York Academy of Sciences, 608*(1), 613–636. https://doi.org/10.1111/j.1749-6632.1990.tb48912.x.

Nadel, L. (2013). Cognitive maps. In D. Waller, & L. Nadel (Eds.), *Handbook of spatial cognition* (pp. 155–171). Washington, DC: American Psychological Association.

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, 64*, 57–67. https://doi.org/10.1016/J.COMPENVURBSYS.2017.01.007.

Omer, I., & Kaplan, N. (2019). 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, 46*(6), 1182–1200. https://doi.org/10.1177/2399808318760571.

Porta, S., Crucitti, P., & Latora, V. (2006). The network analysis of urban streets: A dual approach. *Physica A, 369*(2), 853–866. https://doi.org/10.1016/j.physa.2005.12.063.

Presson, C. C., & Montello, D. R. (1988). Points of reference in spatial cognition: Stalking the elusive landmark. *British Journal of Developmental Psychology, 6*(4), 378–381. https://doi.org/10.1111/j.2044-835X.1988.tb01113.x.

Raubal, M., & Winter, S. (2002). Enriching wayfinding instructions with local landmarks. In M. J. Egenhofer, & D. M. Mark (Eds.)*, Vol. 2478. Geographic information science. GIScience 2002. Lecture notes in computer science* (pp. 243–259). Berlin: Springer. https://doi.org/10.1007/3-540-45799-2_17.

Richter, K.-F., & Klippel, A. (2005). A model for context-specific route directions. In C. Freksa, M. Knauff, B. Krieg-Brückner, B. Nebel, & T. Barkowsky (Eds.)*, Vol. 3343. Spatial cognition IV. Reasoning, action, interaction. Spatial cognition 2004. Lecture notes in computer science* (pp. 58–78). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-540-32255-9_4.

Richter, K.-F., Tomko, M., & Winter, S. (2008). A dialog-driven process of generating route directions. *Computers, Environment and Urban Systems, 32*(3), 233–245. https://doi.org/10.1016/J.COMPENVURBSYS.2008.02.002.

Rodríguez, D. A., Merlin, L., Prato, C. G., Conway, T. L., Cohen, D., Elder, J. P., … Veblen-Mortenson, S. (2015). Influence of the built environment on pedestrian route choices of adolescent girls. *Environment and Behavior, 47*(4), 359–394. https://doi.org/10.1177/0013916513520004.

Sadalla, E. K., Burroughs, W. J., & Staplin, L. J. (1980). Reference points in spatial cognition. *Journal of Experimental Psychology: Human Learning and Memory, 6*(5), 516–528. https://doi.org/10.1037/0278-7393.6.5.516.

Siegel, A. W., & White, S. H. (1975). The development of spatial representations of large-scale environments. *Advances in Child Development and Behavior, 10*, 9–55. https://doi.org/10.1016/S0065-2407(08)60007-5.

Sorrows, M. E., & Hirtle, S. C. (1999). The nature of landmarks for real and electronic spaces. In C. Freksa, & D. M. Mark (Eds.)*, Vol. 1661. Spatial information theory. Cognitive and computational foundations of geographic information science. COSIT 1999. Lecture notes in computer science* (pp. 37–50). Berlin, Heidelberg: Springer. https://doi.org/10.1007/3-540-48384-5_3.

Stern, E., & Portugali, J. (1999). Environmental cognition and decision making in urban navigation. In R. G. Golledge (Ed.), *Wayfinding behavior: Cognitive mapping and other spatial processes* (pp. 99–120). Baltimore, London: Johns Hopkins University Press.

Stevens, A., & Coupe, P. (1978). Distortions in judged spatial relations. *Cognitive Psychology, 10*(4), 422–437. https://doi.org/10.1016/0010-0285(78)90006-3.

Sullivan, Keith, Coletti, Mark, & Luke, Sean (2010). GeoMason: GeoSpatial Support for MASON. *Technical Report GMU-CS-TR-2010-16, George Mason University*. http://mars.gmu.edu/handle/1920/8739.

Tomko, Martin, & Winter, Stephan (2013). Describing the functional spatial structure of urban environments. *Computers, Environment and Urban Systems, 41*, 177–187. https://doi.org/10.1016/j.compenvurbsys.2013.05.002.

Torrens, P. M. (2012). Moving agent pedestrians through space and time. *Annals of the Association of American Geographers, 102*(1), 35–66. https://doi.org/10.1080/00045608.2011.595658.

Torrens, P. M. (2014). High-fidelity behaviours for model people on model streetscapes. *Annals of GIS, 20*(3), 139–157. https://doi.org/10.1080/19475683.2014.944933.

Transport for London. (2018). *Walking action plan*. London: Mayor of London. http://content.tfl.gov.uk/mts-walking-action-plan.pdf?intcmp=54543. (Accessed 1 February 2020).

Tversky, B. (1992). Distortions in cognitive maps. *Geoforum, 23*(2), 131–138. https://doi.org/10.1016/0016-7185(92)90011-R.

Wiener, J. M., Büchner, S. J., & Hölscher, C. (2009). Taxonomy of human wayfinding tasks: A knowledge-based approach. *Spatial Cognition and Computation, 9*(2), 152–165. https://doi.org/10.1080/13875860902906496.

Wiener, J. M., & Mallot, H. A. (2003). "Fine-to-coarse" route planning and navigation in regionalized environments. *Spatial Cognition and Computation, 3*(4), 331–358. https://doi.org/10.1207/s15427633scc0304_5.