Homogeneous Merging (HoMer)

Flavien Loup - - Hadamard
Grenoble Informatics Lab (LIG),
Grenoble-Alps University
flavien.louphadamard@univ-grenoble-alpes.fr

Michael Belfrage
Department of Computer Science and Media Technology & Sustainable Digitalisation Research Centre, Malmö University
michael.belfrage@mau.se

Fanny Di Tursi
Grenoble Informatics Lab (LIG),
Grenoble-Alps University
fanny.di-tursi@univ-grenoble-alpes.fr

Fabian Lorig
Department of Computer Science and Media Technology & Sustainable Digitalisation Research Centre, Malmö University
fabian.lorig@mau.se

Carole Adam
Grenoble Informatics Lab (LIG),
Grenoble-Alps University
carole.adam@imag.fr

Paul Davidsson
Department of Computer Science and Media Technology & Sustainable Digitalisation Research Centre, Malmö University
paul.davidsson@mau.se

Julie Dugdale
Grenoble Informatics Lab (LIG),
Grenoble-Alps University
julie.dugdale@imag.fr

Christopher Frantz
Department of Computer Science, Norwegian University of Science and Technology
christopher.frantz@ntnu.no

ABSTRACT

As many cities are increasingly threatened by natural disasters, the study of evacuation and population behaviour during those disasters is identified as a crucial point by many stakeholders and researchers. As the environment is very important in those studies, integrating Geographic Information Systems (GIS) data into an Agent Based Social Simulation (ABSS) marks a significant step forward in crisis modelling. However, it also introduces a key challenge: adapting highly detailed spatial data to reduce environmental complexity while preserving the essential information necessary to approximate higher-resolution environments. Drawing on two disaster evacuation scenarios using different spatial data structures, namely graph-based and grid-based representations, this paper proposes an original methodology to address this challenge. The proposed methodology, HoMer (Homogeneous Merging), provides a structured approach to reducing spatial resolution while retaining crucial information and striking a balance in computational scale for environments that afford agent-environment interactions. It is composed of a three-step framework: (1) attribute reduction, (2) merging of spatial entities, and (3) evaluation of the output. Our results demonstrate that HoMer makes it possible to identify an appropriate ABSS resolution for a given project while maintaining maximum data integrity.

Keywords

Agent Based Social Simulation, Environment modelling, Spatial entity, Methodology, GIS Data

INTRODUCTION

As many cities are increasingly threatened by natural disasters (Van Aalst 2006), the study of evacuation and population behaviour during those disasters is identified as a crucial point by many stakeholders. Indeed, evacuation is one of the crisis management tools used in case of a big event, such as a forest fire, for example (Kirschenbnum 1992 ; Lindell 2021). However, field-based simulations of population evacuation are difficult, time-consuming, and costly to organise. Computer simulations of evacuation, which incorporate social behaviours, serve as a valuable tool for analysing such situations. They allow multiple scenarios to be examined quickly and at low cost.

Social simulations have been studied since at least 1971, with Schelling's model of segregation (Schelling 1971). They aim at reproducing social phenomena (Epstein 1999) with high "behavioural fidelity" while neglecting the "environmental fidelity". However, in the case of natural disasters, it is important to study human behaviour in the context of the specific environment in which the evacuation occurred (Grazia De Paoli et al. 2020). Thus, the environment should be represented in sufficient detail when Agent-Based Social Simulation (ABSS) is used for disaster management.

Modelling environments of disaster situations, along with human behaviour, requires a lot of data. Since the adoption of the European Union's Open Data Directive (EU 2019/1024)[1], high-value datasets (HVD) such as population statistics, cadastral and geospatial information, and mobility data are now becoming increasingly available through open access. Agent-programming frameworks, such as GAMA, Repast (Team 2006), or AnyLogic (AnyLogic 2014) integrate geographical information systems (GIS) support. This means that fine spatio-temporal resolution and more complex agent-environment interactions can be modelled (Mancheva et al. 2019, Kaziyeva et al. 2021).

However, modelling such complex interactions requires a very high resolution, which means small-sized objects simulated in the model, which may not be computationally feasible with large-scale simulations, i.e., simulations with a large studied area. The trade-off between spatial scale and representational resolution is exemplified by the distinction between macroscopic and microscopic approaches. For example, in road traffic simulation, when focusing on a single road intersection, it is possible to simulate each vehicle individually, capturing detailed behavioural dynamics (Doniec et al. 2008). When the study area encompasses an entire city or country, such fine-grained resolution becomes computationally impractical. In these cases, the issue is addressed by adopting flow-based models that represent aggregated traffic dynamics rather than individual vehicles. Furthermore, ABSS is a highly diverse field characterised by strong multidisciplinarity. The complexity of GIS data and their processing can be a barrier to their use by non-geomatics experts. Most current computer social simulation environments are either too precise or too abstract to be used to represent a natural disaster and to be useful to stakeholders. Therefore, in this paper, we address the following questions:

- How can GIS data be leveraged in ABSS by modelling it in interactive environments?
- How can the balance be struck between high resolution and large scale in an ABSS for interactive environments?

This paper proposes a methodology designed to reduce environmental complexity while preserving the essential information required for ABSS. Our methodology deliberately relies on simple and intuitive concepts to ensure accessibility for researchers from diverse disciplines, including those without expertise in geomatics, graph theory, or computer science. Therefore, the methodology can be readily extended or refined through the integration of domain-specific techniques tailored to the environment under study. We use GIS data from two separate disaster management projects to demonstrate proof-of-concept for the method across different data types and scenarios. The first section reviews the related work and existing simulation scales and the limitations of these approaches. We then present the theoretical aspects of the methodology in the second section. Finally, in the last section, we introduce two different evacuation scenarios using GIS data to highlight crucial methodological considerations.

ENVIRONMENT REALISM AND RESOLUTION IN ABSS

Early works in social simulation, including Schelling's model of segregation (Schelling 1971), were characterised by the use of simple rule-based agent behaviours embedded in highly abstract environments. This approach, commonly referred to as "Keep It Simple, Stupid" (KISS, Edmonds and Moss 2004), remains both influential and productive across numerous domains. NetLogo, an agent-based simulation platform, is a useful tool to explore the KISS paradigm in ABSS (Tisue, Wilensky, et al. 2004) and is still largely used (Avila-Garzon et al. 2022, Grimm et al. 2025, Pires et al. 2026). However, since the early 2000s, there has been a growing shift towards developing

[1] *Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on Open Data and the Re-use of Public Sector Information* 2019

more realistic and data-driven models, exemplified by the “Keep It Descriptive, Stupid” (KIDS) school of thought (Edmonds and Moss 2004). This evolution has been accompanied by a growing interest in the realistic representation of environments within agent-based systems (Weyns, Van Dyke Parunak, et al. 2004; Weyns, Omicini, et al. 2007). While the environment is an ubiquitous concept in ABSS and, more broadly, in multi-agent systems (MAS), the term is often used ambiguously, encompassing a wide range of aspects, and has received limited attention as a research subject in its own right (Weyns and Michel 2015).

GIS data integration enables agent-environment interaction with a realistic environment, improving the validity of simulations' predictions and improving the acceptance of those predictions by decision-makers. (Schoenharl and Madey 2011). The use of GIS data foregrounds the potential of high-resolution environments as natural and built settings that shape the affordances and constraints of agents, enabling more realistic simulation models (Papasimeon 2010). While the KISS approach can be effective in cases where the environment is relatively simple or can be neglected altogether, this comes at the cost of foregoing such environmental richness. The advantage of this simplification is reduced time complexity from an algorithmic standpoint. Time complexity otherwise increases with the scale and spatio-temporal resolution of the environment, as well as with the complexity of agent decision-making and agent–environment interactions. The downside is that de-emphasising, or even entirely neglecting, the environment risks overlooking emergent dynamics arising from agent–environment interactions, which may significantly influence simulation outcomes. For instance, extensions of the BOIDS model incorporate features such as coral reefs and other obstacles, illustrating how agent–environment interactions can fundamentally alter collective dynamics rather than merely scaling individual behaviour (Rahmani et al. 2020). In earthquake simulations including more complex environments, such as falling debris capable of obstructing evacuation routes, along with more realistic agent behaviour, agents' arrival to safe areas was significantly delayed (Iskandar et al. 2024). Ultimately, increasing the resolution of environments facilitates a shift from static digital models toward dynamic and complex environments, and enhances the perception of realism for all stakeholders.

In the context of European Union's Open Data Directive*Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on Open Data and the Re-use of Public Sector Information* 2019 agent-programming frameworks integrating geographical information systems (GIS) support offer the opportunities to implement finer spatiotemporal resolution and improve the perceived realism of users. GAMA, an agent-based simulation platform used to develop data-driven simulation models, enables the direct use of GIS data such as GeoJSON, Shapefile, or SVG. GAMA has already been used many times on a wide number of projects where the environment is described at a fine resolution in order to allow environment-user interaction Taillandier et al. 2019, multimodal movement (network, continuous or discrete) (Drogoul et al. 2013;Yang and Xu 2024).

Having a fine resolution is particularly important in the case of natural disasters, where the environment strongly influences human behaviour and evacuation (changing environment, risk perception). Highly dynamic environments, such as disaster situations with evolving hazards, modifications to the infrastructures and resource availability, introduce additional constraints that affect the dynamics of models as re-evaluation of the perceived environment or dynamic path computation (Moradi et al. 2025; Clark 1997). Crisis management is thus a particularly relevant domain for investigating environmental complexity in simulation, and bringing more realism will enhance both results and their perceived validity by users.

To make GIS data more suitable for simulations, a number of simplification strategies can be employed. In relation to network data, one approach is to prune the graph by removing nodes with low network centrality (Porta et al. 2006). However, this may reduce network connectivity and compromise local structural integrity, potentially distorting movement patterns. Relatedly, another graph-based strategy is node suppression, where chains of nodes with two connections (*i.e.*, non-intersection nodes) are removed (Boeing 2017). Since this approach does not account for curvature, it tends to straighten roads, causing the network as a whole to contract. A different approach is to treat individual edges as lines and apply line simplification algorithms. For instance the Ramer–Douglas–Peucker (RDP) algorithm removes intermediate points between two endpoints based on a tolerance threshold (Douglas and Peucker 1973). As this approach is purely geometric and developed for line simplification, it does not inherently account for network topology and may remove nodes representing intersections, potentially resulting in a loss of network connectivity. Another approach that seeks to preserve network structure is stroke-based generalisation, which relies on geometric continuity like angular tolerance to amalgamate edges that are semantically or geometrically similar into longer “strokes” (Z. Li and Dong 2010).

To represent continuous environments, there are many solutions to reduce data for simulation. The most commonly used for 3D environments is mesh representation, for instance, in shortest path algorithms in video games (Pelechano and Fuentes 2016). Meshes allows many reduction algorithms such as vertex clustering, or edge collapse (S. Li et al. 2018). However, 3D representations can be very computationally expensive with only minor gains in perceived realism. It is possible to make a 2D projection of the environment toreduce computational expense. This can be done through shape and geometry (Solmaz and Turgut 2017), or by using grid, as in one of the most used fire simulators (Finney 2006) that aggregates all of the information needed to simulate a forest fire into cells. (Sharma et al. 2016) present a LiDAR data reduction algorithm using a grid representation. A similar method with a simpler algorithm has been used for our case study presented in the section of this paper : "Grenoble Bastille: Evacuation of a cultural heritage site in case of forest fire". All of the above solutions have their own benefits and drawbacks. This paper does not propose to compare those solutions but to present a workflow to apply them to your own context.

METHOD: HOMOGENEOUS MERGING

We propose to simplify GIS data through an attribute-driven merging procedure that seeks to reduce spatial data while retaining all relevant information to increase computational efficiency for environments affording agent-environment interactions. GIS data typically includes spatial geometries (nodes, edges, grid cells, etc.) with associated attributes (*e.g.*length, altitude, speed limit, street names). We refer to these units as *Spatial Entities* (SEs). GIS databases have become widely available following the European Open Data Directive. However, raw GIS datasets are often large, highly detailed, and internally inconsistent, making them difficult, or even impossible, to load and use without preprocessing. The central challenge is, therefore, to preserve relevant spatial information while reducing computational overhead during runtime as much as possible. We address this challenge by aggregating SEs into the largest spatial unit possible with the same attribute value, termed *Largest Homogeneous Spatial Entities* (LHSEs). For this reason, we refer to the procedure as *Homogeneous Merging* (HoMer). Accordingly, to balance information richness with simplicity, HoMer merges SEs that share either similar attributes or geometric properties into LHSEs. This implies that two distinct forms of homogeneity can be distinguished and merged within GIS data: attribute homogeneity and spatial homogeneity.

Data selection and pruning can be guided by attribute homogeneity, since much GIS data contains information irrelevant to any specific ABSS model. For example, street names in a road network are often unnecessary, creating unnecessarily small SEs that could be merged if removed. Once the essential attribute data has been identified and the rest has been discarded, it is possible to identify the LHSE. Applying aggregation procedures (e.g., mean, minimum, sum) to attributes allows additional SEs to be merged without incurring any significant loss of information. Merging can also be guided by spatial homogeneity, defined as the presence of simple geometric shapes, where SEs with similar – but not identical – attribute values (*e.g.* two roads with comparable speed limits of 80–90 km/h) are combined into a single SE. Alternatively, spatial homogeneity may be maintained by preserving a higher number of individual structuring elements (SEs), instead of combining multiple simple-shaped SEs into a single, larger structuring element with increased geometric complexity. Another strategy to achieve spatial homogeneity is through a predefined environment (*e.g.* a grid) where raw GIS data are integrated into those spatial units. After merging, both spatial and attribute integrity must be verified. Spatial accuracy is checked by comparing the resulting geometries with the raw GIS data, while attribute loss is quantified to ensure that simplification does not compromise essential data. In order to find the perfect scale for your simulation, it's possible to make multiple merges with different parameters and compare their evaluation.

The methodology can be summarised in the following steps:

- **Step 1**: Attribute reduction — retain only the attributes of interest, pruning others and, where possible, reducing the number of unique values.
- **Step 2**: Perform merging — Merge SEs with neighbours if their attributes are the same or similar. Multiple approaches can be used to merge SEs into LHSEs with different parameters.
- **Step 3**: Evaluation — Compare the result with the original data set to ensure that data losses are acceptable to the study case and check the gain in computation time.

The following section presents two disaster-management case studies employing different types of spatial data using HoMer. The first subsection of each case provides background and covers steps 1 and 2, while the second focuses on evaluation, addressing the final step by comparing results with the original data.

TWO CASE STUDIES IN DISASTER MANAGEMENT

Malmö Citylab: modelling a Civil Defense Shelter

Developed in collaboration with Malmö City's disaster management division, the model functions as a decision-support tool that enables policy actors to test and compare policies in a safe, simulated environment (Belfrage et al. 2024). The model examines civil evacuation under conditions in which shelter accessibility and capacity are unevenly distributed across the city and constrained by the underlying road network. Road networks are typically represented as graphs including nodes (vertices) and links (edges) and are denoted as $G = (V, E)$ (Marshall et al. 2018), which we will also adopt. Publicly available data, such as Malmö's road network, is often incomplete and may therefore contain disconnected subcomponents that prevent agents from traversing the network as intended. To address this, we added 323 edges to connect each of the disconnected subcomponents to the main component, thereby ensuring full network connectivity. However, this also increased the size of the network, yielding a graph with vertices $|V| = 99.414$ and edges $|E| = 107.565$. The goal of this model is to agentify the road network so that each road link becomes an individual road agent with its own speed limit, repeated vehicle–road interactions to communicate speed limits, and computed shortest-path routes. This means that there is a substantial computational overhead for a network of this size. Therefore, the motivation for applying HoMer is to reduce Malmö's large-scale network, enabling a coarse-level analysis before proceeding to more granular simulations at the city district level, which is the smallest spatial unit for which aggregate population statistics are available.

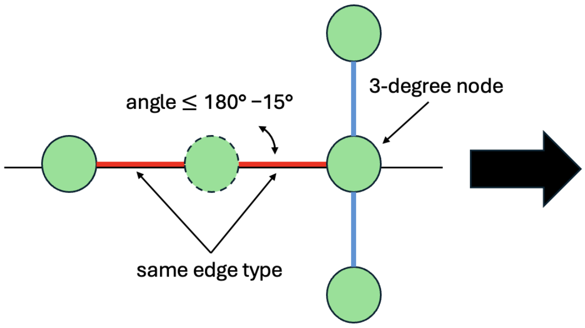

Figure 1. Logic of the curvature-aware 2-degree reduction

To reduce the size of Malmö's road network, various data reduction and optimisation techniques can be applied to reduce both the number of nodes and links, based on the speed limit of the roads, with the goal of retaining information and the structural integrity of the geography. Practically, this means that any straight road containing intermediate degree-2 vertices – where the degree of a node is defined as the number of links connecting it to other nodes (Barabási 2013) – along edges with the same speed limit can be merged without losing relevant information. This exemplifies the information breaking point – the Largest Homogeneous Spatial Entity (LHSE) – in the Malmö city road network, which can serve as a guiding principle for environmental modelling in ABSS. To this end, we tested two complementary approaches: (i) a node suppression (degree-2) algorithm targeting chains of nodes (*i.e.*, roads without intersections), and (ii) a stroke-based (curvature-aware) algorithm, which further targets degree-2 nodes where connected segments deviate by no more than 15° from a straight line (technically 180° - 15°). This logic is illustrated in Figure 1 where both algorithms focus on 2-degree nodes, since nodes with a higher degree (e.g., 3-degree) do not correspond to straight road segments but to intersections. The motivation for testing two different approaches is that even a 2-degree reduction may lead to a loss of geometric information, as it could cause the network to drift (roads change coordinates/locations) or contract by pulling nodes closer together, thereby making it structurally smaller.

Evaluation

Figure 2A (in red) shows the original road network of Malmö city in Sweden, G_A, with $|V(G_A)| = 99{,}414$ vertices and $|E(G_A)| = 107{,}565$ edges. The blue network, Figure 2B is the curvature-aware reduction, in which pairs of vertices with the same speed limit and with an angular deviation below 15° are merged, resulting in one vertex being removed (see Figure 1). This produces a moderate overall reduction with $|V(G_B)| = 70{,}329$ and $|E(G_B)| = 78{,}480$. The green network shown in Figure 2C was reduced using the more aggressive 2-degree reduction $|V(G_C)| = 18{,}342$, $|E(G_C)| = 26{,}036$, suppressing vertices on edges with the same speed limit without considering curvature, leading to a substantial decrease in geometric complexity. To support visual inspection of the geographical integrity post reduction, four areas of Malmö have been chosen: Västra Hamnen (V.H), the island in Limhamn (Ön), the Öresund Bridge (Ö.B) connecting Malmö to Denmark, and the docks in Bunkeflo (B.F). As seen around the Ö.B area, both G_B and G_C show some geometric degradation with respect to G_A. The loss of geometric information is considerably greater in the green (2-degree reduction) network, particularly when examining other areas of the network.

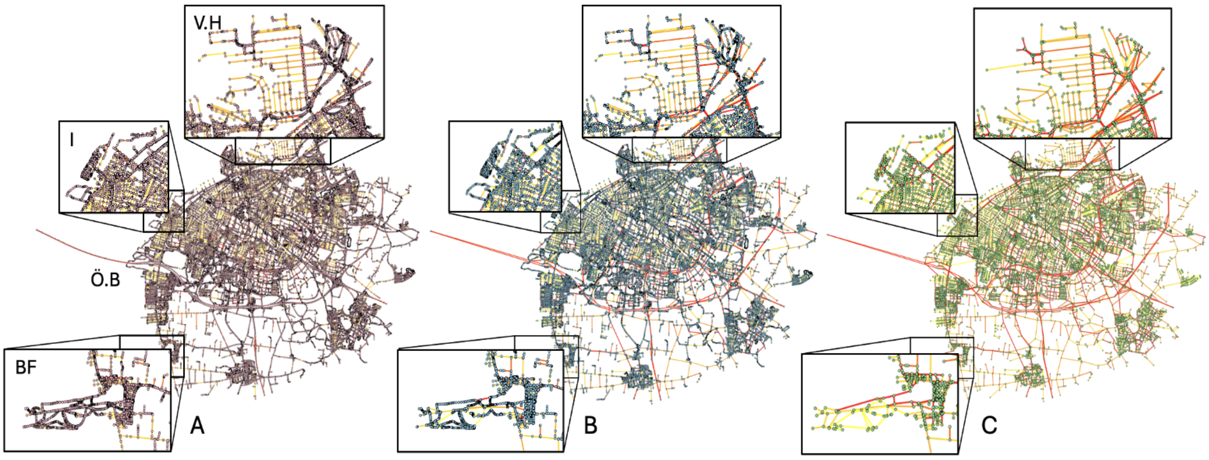

Figure 2. Three versions of Malmö City's road network. Network A (in red) represents the original network after connecting all previously disconnected components. Network B (in blue) is a curvature-aware 2-degree reduction using a 15° angular tolerance, while Network C (in green) is a 2-degree reduction merging edges of the same type without any angular consideration.

To benchmark the computational gains from reducing the road network, we ran three simulation experiments comparing the two reduced networks with the original network. We conducted a batch experiment with 20 iterations per condition. Each iteration included 20 randomly distributed agents, each with randomly assigned targets, and with the shortest paths to the targets being calculated using Dijkstra's algorithm. Given the simplicity of this experiment (comprising only two agent types, people and roads, and a limited set of functions), the interactions between model components and the environment are minimal. This makes the experiment well suited to test the effect of network reduction on runtime. After the experiments were completed, the average initialisation and computation times were calculated for each network to assess the impact of each reduction. Table 1 shows that the initialisation time varies negligibly across conditions, whereas the shortest-path computation time decreases markedly following network reduction. On average, computation on the original network takes 998.5 s (16.6 min), compared to 420.8 s (7.0 min) for the curvature-aware reduction and 73.7 s (1.2 min) for the two-degree reduction. Accordingly, there are substantial computational gains to be made from network reduction, with the 2-degree reduction providing the largest speed-up and the curvature-aware reduction offering a significant, though smaller, improvement in runtime.

Table 1. Time of shortest-path experiments reported in seconds.

Graphs	Original	Curvature reduction	2-degree reduction
Initialisation	1.768410	1.768413	1.768414
Shortest path computation	998.5	420.8	73.7

Next, we verified attribute consistency in the two reduced graphs by comparing the attribute data we had access to, namely speed limit, to that of the original graph. For this comparison, we summed the total length of all road segments sharing the same speed limit and then compared their proportional distribution across the network. Using the original network as ground truth, the numbers indicate that the curvature-aware reduction preserves both structural and attribute properties more faithfully than the 2-degree reduction. The total edge length is almost unchanged under the curvature-aware method (< 0.5%deviation), whereas the 2-degree reduction shortens the network by about 5.5%, indicating a structural contraction of the network. Speed limit distributions are nearly identical to the original for the curvature-aware graph, while the 2-degree reduction introduces small proportional shifts and a stronger reduction in variance, suggesting a decrease in attribute heterogeneity. More intuitively, when attribute heterogeneity decreases, extremes (very low or very high speeds) become relatively less represented.

Overall, the curvature-aware approach preserves geometric and attribute resolution more closely, whereas the 2-degree reduction induces small, yet systematic simplification effects relative to the overall size of the network.

Table 2. Attribute consistency comparison.

Graphs	Original	2-degree reduction	Curvature-aware reduction
Total edge length (km)	2244.19	2121.37	2234.05
0 Km/h	13.27%	12.57%	13.28%
30 Km/h	25.31%	24.70%	25.40%
50 Km/h	46.92%	48.05%	46.93%
70 Km/h	6.90%	7.07%	6.90%
90 Km/h	2.70%	2.47%	2.60%
110 Km/h	4.91%	5.14%	4.89%
Speed limit (avg)	37.39	35.86	33.67
Standard deviation	25.73	21.85	23.95

Beyond attribute consistency, we also considered geographical coherence, which seeks to quantify and assess the 'drift' introduced by network reductions. This aspect is particularly important in multidisciplinary projects or when the intended audience includes non-technical stakeholders. If the visual representation of the environment deviates too strongly from the real world, the intended message may be misinterpreted or not properly conveyed. With this in mind, we compared the two reduced networks to the ground truth, i.e., the original network, to ensure consistency in the spatial representation of the environment. By tracing each vertex of the original network to the nearest edge in each reduced graph, we quantified the drift between them in meters. To examine its spatial distribution, we classified the distances into five bands (0; 1–10; 11–50; 51–100; 5. > 100), which were then colour-coded. Accordingly, the closer the distances are to 0, the more closely the reduced graph resembles the original road network. Conversely, larger distances indicate greater deviations from the original network. The results of this analysis are shown in Figure 3.

Table 3. Distance comparison between the Malmö city's road network (points) and the reduced graphs (edges).

Graphs	Curvature-aware reduction		2-degree reduction	
Distance (in meters)	**Amount of points**	**Percentage**	**Amount of points**	**Percentage**
0	171940	79.92	89690	41.69
1-10	33010	15.34	87778	40.80
11-50	9206	4.28	32370	15.05
51-100	620	0.29	3762	1.75
> **100**	354	0.16	1530	0.71
Total	215130	100	215130	100

We calculated the distance from edges in the original network to equivalent edges in the reduced graphs 3. For the 2-degree reduction graph and the curvature-aware 2-degree reduction graph, the proportion of identical edges is 41.7% and 79.9%. When including small gaps of 1–10 meters between otherwise similar edges, the similarity increases to 80.5% for the standard 2-degree reduction and 95.3% for the curvature-aware reduction. Considering larger distances, the 2-degree reduction has 15.05% of edges within 11–50 m, 1.75% within 51–100 m, and 0.71% exceeding 100 m, while the curvature-aware reduction has only 4.28%, 0.29%, and 0.16% of edges in the same ranges, respectively. The maximum observed gaps are 583 m for the 2-degree reduction and 373 m for the curvature-aware reduction, although distances above 100 m represent less than 1% of edges in both cases. Echoing the result, from the visual inspection of Figure 2, the curvature-aware reduction preserves the original network's structural integrity better than the 2-degree reduction, with most edges either identical or only a few meters apart. Deviations above 100 m are rare in both cases, indicating that both approaches maintain overall network coherence, but the integrity of a few peripheral structures in the 2-degree reduction is compromised.

In conclusion, the results of the evaluation underscore a clear trade-off between computational gains and geographical resolution. Reducing the original network to the curvature-aware graph and the standard 2-degree graph leads to substantial reductions in runtime, particularly for the 2-degree network, which saves 924.8 seconds relative to the original (998.5 s to 73.7 s), compared to a 577.7-second reduction for the curvature-aware network (to 420.8 s). In terms of geographical coherence, the 2-degree reduction exhibits more drift from the original network, with only

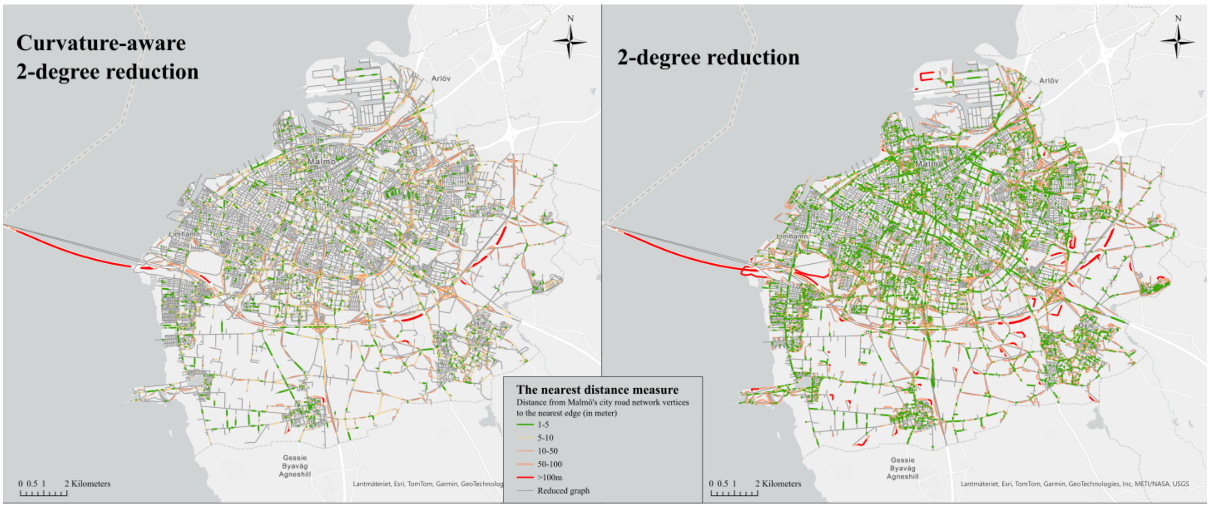

Figure 3. Differences between the reduced graphs and the Malmö city's road network

41.7% of edges remaining identical (compared to 79.9% in the curvature-aware graph) and more frequent mid-range deviations. However, these improvements come with differing structural consequences: the 2-degree reduction contracts total network length by approximately 5.5%, while demonstrating some loss of peripheral geographical structures, and reduces attribute heterogeneity, reflected in a lower standard deviation of speed limits (21.85 vs. 25.73 in the original). Overall, the curvature-aware approach achieves some computational gains while maintaining high geometric and attribute consistency, whereas the standard 2-degree reduction maximises efficiency at the cost of geographical resolution.

Grenoble Bastille: Evacuation of a cultural heritage site in case of forest fire

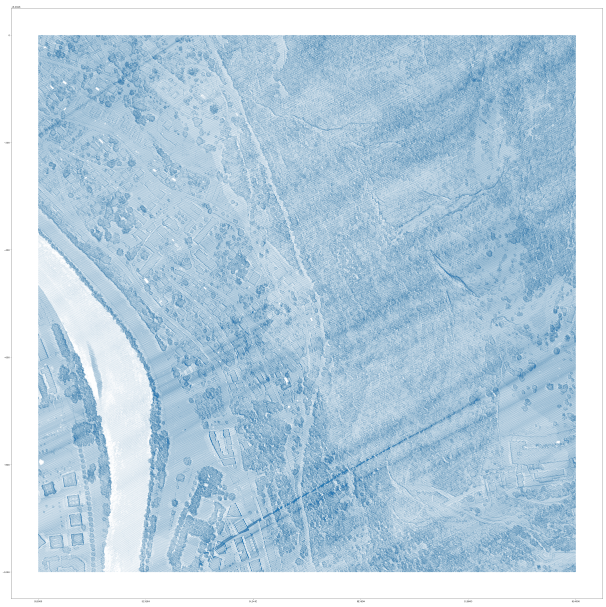

Figure 4. Bastille LiDAR Dataset

The Bastille project aims to study the evacuation of the local population and tourists in case of forest fire around the cultural site of the Grenoble Bastille. The goal is to develop an agent-based simulator where people are represented as agents moving around the Bastille environment. To model the environment, we first used the LiDAR (Light Detection and Ranging) dataset provided by the French National Geographic Institute (IGN 2022). The dataset consists of point clouds and provides environmental information such as elevation, vegetation type, roads, and buildings (IGN 2022, p.13). However, the raw data cannot be used directly, as the point clouds are discontinuous and agents cannot be simulated in such an environment. Furthermore, as illustrated in Fig. 4, LiDAR points are extremely numerous and unevenly distributed throughout the environment. Consequently, the gaps need to be filled,

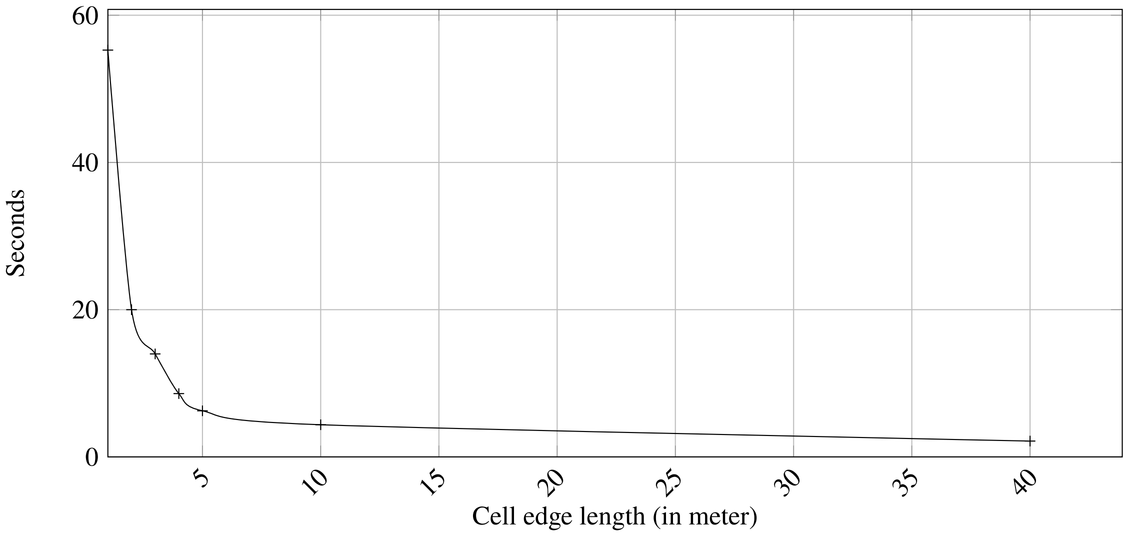

Figure 5. Initialization time depending on cell size

and the SEs should be evenly distributed in the area to represent the whole surface. To do so, we decided to use HoMer to create a grid representation. Merging all LiDAR data points into cells allows the aggregation of their attributes within each cell. Increasing cell size allows the reduction of environmental complexity. Among LiDAR dataset attributes, we can reduce the environment to two of them: elevation and class (buildings, vegetation level, water, etc.). To merge LiDAR data points in a grid, a cell has the same class value as the most common aboveground points covered by it, and the elevation is given by the mean value of all the ground points covered by that cell.

Evaluation

We simulated the environment on multiple cell sizes to compare computational complexity and data loss. By reducing the number of cells, we reduced both computational complexity and data resolution. This section examines the benefits and drawbacks of grids at different scales, with cell sizes ranging from 1 to 5 meters.

We compared the data complexity of different cell sizes by running a series of tests on the GAMA platform, which will be used for the simulation. The tests consist of launching the simulation multiple times and checking the time it takes to initialise the environment. The results of these experiments are presented in Fig. 5. This curve is shaped like a reciprocal function, clearly decreasing from the initial values onwards. We compared classification distributions for both original LiDAR data points and resulting grids. The results of this experiment are presented in Fig.6. As we want to classify a cell from the aboveground elements composing it, the disappearance of many ground points is expected.

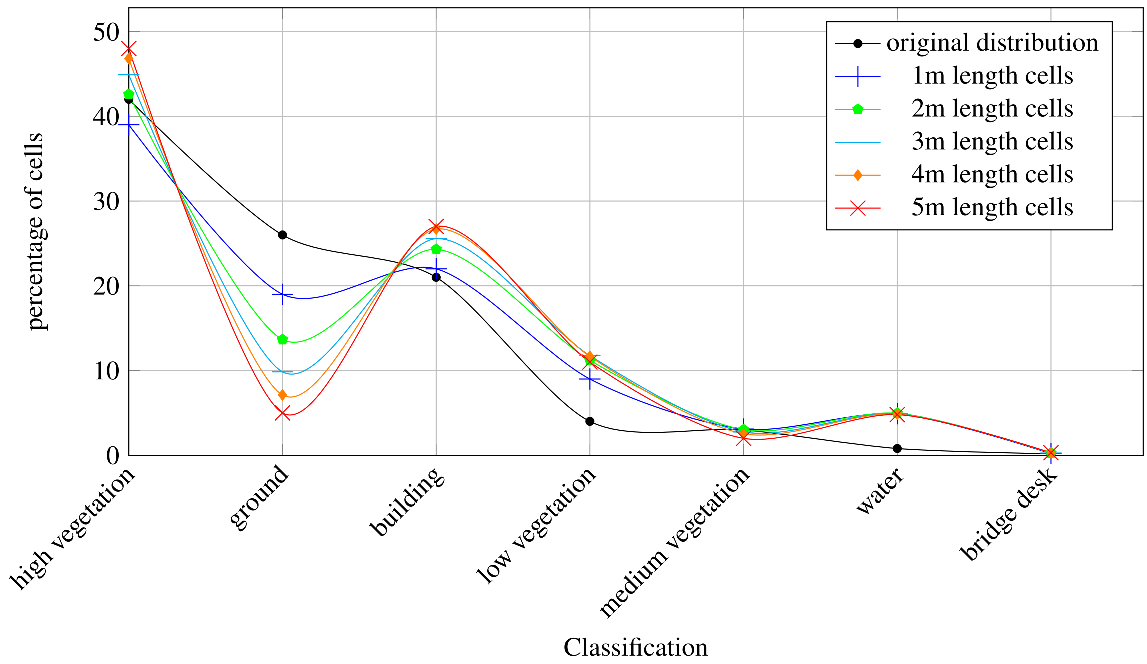

Figure 6. Classification comparison

The last step of our experiment is the evaluation of the spatial integrity of data. We choose to make our experiment only on the building entities because it is the most relevant case among the other classifications. We compared our results with previously validated data from the IGN topological database (BD Topo). This database includes a dataset (topoDS) providing a highly accurate representation of all buildings in France. It is already widely adopted and extensively used throughout the country. To achieve this, we calculated the Jaccard index. (Jaccard 1901), computed as follows:

$$J(A, B) = \frac{A \cap B}{A \cup B} \tag{1}$$

Where: (A) is the LiDAR DS; and (B) is the topoDS.

This allows us to compare the differences in loss and gain between the topoDS and LiDAR datasets. The LiDAR grid was created according to the following sizes: 1, 2, 3, 4, and 5 m. Instead of comparing datasets at the building resolution, we decided to do so at the islet resolution because the latter is used for an agent-based simulation, enabling us to apply a larger scale. We used a set theory formula, the Jaccard index

The result is a number between 0 (completely different shapes) and 1 (perfectly similar shapes in the two datasets). This index enables us to quickly compare each LiDAR grid size and assess the robustness of the dataset used.

The curves in figure 7 show the repartition of the buildings' islet over Jaccard's index. A building islet is the result of a merging of adjacent and grouped buildings. As expected, there are no, or almost no, buildings above 0.9 on the Jaccard index, as the LiDAR grid delimitation never fits topoDS buildings' shapes exactly. We can also notice that the curve shifts to the left when cell grid lengths rise, meaning that smaller grids are more accurate. We also computed the mean Jaccard index across all LiDAR datasets and weighted it by surface area. The resulting weighted mean is substantially higher, indicating that HoMer methodology performs more effectively for large surface areas. This finding further suggests that the most significant inaccuracies are primarily associated with small buildings, which account for only a limited proportion of the total surface area.

Here we have a visual representation of shape distortion in Fig. 8 between the building islet and the original building from the topoDS. In the zoomed sector of the Bastille, three scenarios are represented following the LiDAR grids of 1, 2, and 5 m in length. The darker the colour of the building islet, the higher the score of Jaccard's index is. So, it means that its shape is closely similar to the original buildings. Then, the scenarios with a LiDAR grid of 1 m and 2 m length are closer to the reality and are fully satisfactory for our study case.

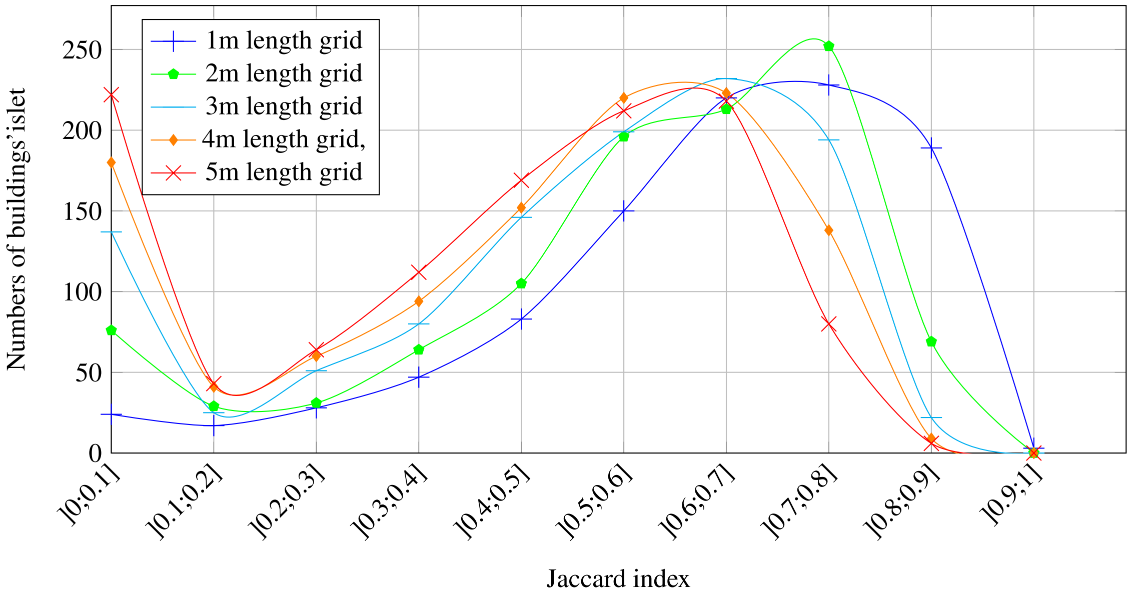

Figure 7. Jaccard index comparison

Table 4. Jaccard mean values

cell size	1m	2m	3m	4m	5m
Jaccard mean	0.63	0.55	0.49	0.44	0.41
weighted mean	0.77	0.71	0.65	0.58	0.51

We manually examined many of the extreme errors and found that they are largely due to differences in data-collecting methods between LiDAR and TopoDS rather than the merging procedure. For instance, some buildings in the LiDAR dataset are hidden by forest trees, leading to cells being falsely classified as high vegetation. Regarding all of the experiments, we conclude that it is necessary to increase cell size over 1 m to reduce computation time; even reducing resolution from 1 m to 2 m has a great impact on this aspect Fig.5. As it shows in Fig.6, attribute experiments do not show a shifting difference in attribute distribution on the map. The reduction of ground is induced by the merging method. Geographical studies show that the shape of buildings deteriorates at every cell size increase Fig.7. Regarding all those observations, the 2 m length grid is the best compromise between complexity and accuracy for our case study.

CONCLUSION

In this paper, we have proposed the original method of HoMer to adapt geographic data for ABSS. It is composed of a simple three-point framework: 1) attribute reduction, 2) merging in LHSE, and 3) evaluation. The method is purposefully flexible, guided by the LHSE principle, to strike a balance between resolution and scale in environments that afford agent-environment interactions. On the one hand, the strength of the method, as demonstrated by the two disaster management cases using different types of data structures, including graph and grid-based representations, is that it is sufficiently flexible to be applied across diverse environment types. On the other hand, a potential drawback of this flexibility is that, although identifying the LHSE can guide reduction without losing relevant information, multiple approaches often remain viable once the LHSE has been identified. However, by repeating the last two steps with different merging methods, it is possible to find out the ABSS resolution acceptable for a project. This acceptability can be defined by the balance between running time and dataset resolution. Due to the flexibility of the LHSE principle, we expect that it can be further generalised to other cases and types of environments. We will have to conduct this methodology on further projects in order to verify if it is applicable and useful in other contexts. In addition to those upcoming experiments, we want to test other graph reduction algorithms for Malmö to check if it is possible to find a better representation. Another important direction for future work is to conduct simulations using a range of simplification methods in order to examine their effects on perceived realism as well as on decision-making outcomes. Finally, in the Bastille case, we did not do an experiment on the shortest path algorithm. As we have already seen in the Malmö case, the shortest path algorithm is more sensitive to environmental complexity than initialisation. A similar behaviour can reasonably be expected in the context of the

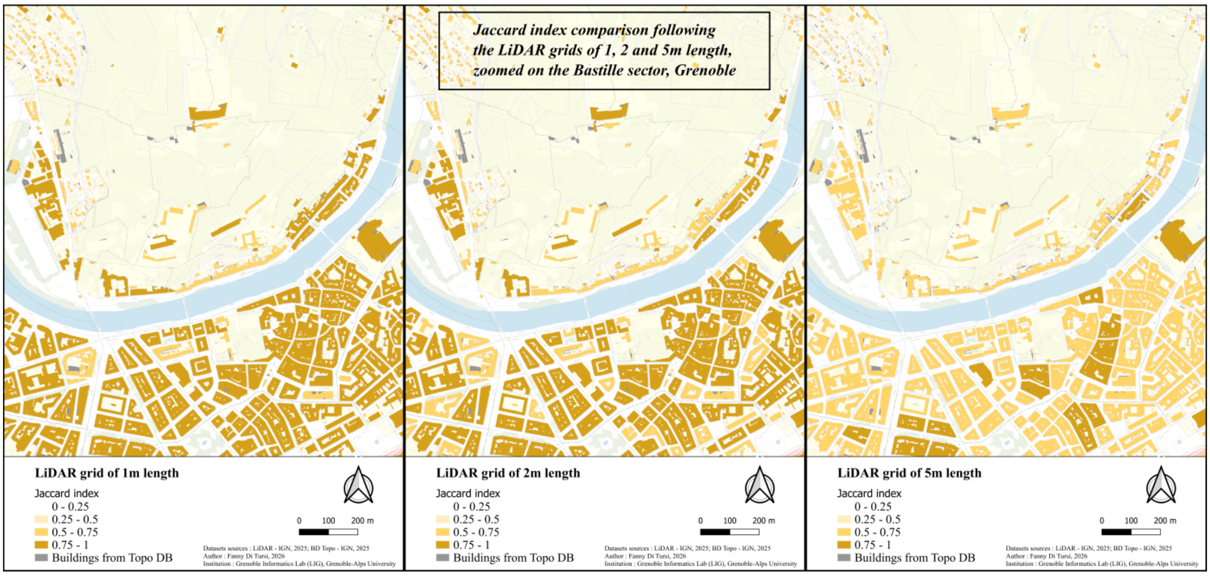

Figure 8. Jaccard Index representing LiDAR grids of 1, 2 and 5m length

Bastille experiment. However, this aspect requires further investigation in order to conduct a dedicated experiment and accurately assess the respective gains and losses associated with each dataset.

ACKNOWLEDGEMENTS

This work was partly supported by the Wallenberg AI, Autonomous Systems and Software Program – Humanities and Society (WASP-HS) funded by the Mariane and Marcus Wallenberg Foundation and the Marcus and Amalia Wallenberg Foundation. It is also partly supported by (PCI2024-153441) funded by MICIU/AEI/10.13039/501100011033 and cofunded by the European Union. (ANR-23-JCCH-0005)

REFERENCES

AnyLogic (Oct. 2014). *AnyLogic 7.1 Released*. https://www.anylogic.com/blog/anylogic-7-1-released/.

Avila-Garzon, C., Balaguera, M., and Tabares-Morales, V. (2022). "An Agent-Based Social Simulation for Citizenship Competences and Conflict Resolution Styles". In: *International Journal on Semantic Web and Information Systems (IJSWIS)* 18.1, pp. 1–23.

Barabási, A.-L. (2013). "Network science". In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 371.1987.

Belfrage, M., Lorig, F., and Davidsson, P. (2024). "Simulating change: A systematic literature review of agent-based models for policy-making". In: *2024 Annual Modeling and Simulation Conference (ANNSIM)*. IEEE, pp. 1–13.

Boeing, G. (2017). "OSMnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks". In: *Computers, environment and urban systems* 65, pp. 126–139.

Clark, A. (1997). *Being there*. mit Press Cambridge, MA.

Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on Open Data and the Re-use of Public Sector Information (2019). Official Journal of the European Union, L 172, 26 June 2019, pp. 56–83.

Doniec, A., Mandiau, R., Piechowiak, S., and Espié, S. (2008). "A behavioral multi-agent model for road traffic simulation". In: *Engineering Applications of Artificial Intelligence* 21.8, pp. 1443–1454.

Douglas, D. H. and Peucker, T. K. (1973). "Algorithms for the reduction of the number of points required to represent a digitized line or its caricature". In: *Cartographica: the international journal for geographic information and geovisualization* 10.2, pp. 112–122.

Drogoul, A., Amouroux, E., Caillou, P., Gaudou, B., Grignard, A., Marilleau, N., Taillandier, P., Vavasseur, M., Vo, D.-A., and Zucker, J.-D. (2013). "GAMA: multi-level and complex environment for agent-based models and simulations (demonstration)". en. In: *International conference on Autonomous agents and multi-agent systems*. United States, pp. 1361–1362.

Edmonds, B. and Moss, S. (2004). "From KISS to KIDS–an 'anti-simplistic'modelling approach". In: *International workshop on multi-agent systems and agent-based simulation*. Springer, pp. 130–144.

Epstein, J. M. (1999). "Agent-based computational models and generative social science". In: *Complexity* 4.5, pp. 41–60.

Finney, M. A. (2006). "An overview of FlamMap fire modeling capabilities". In: *In: Andrews, Patricia L.; Butler, Bret W., comps. 2006. Fuels Management-How to Measure Success: Conference Proceedings. 28-30 March 2006; Portland, OR. Proceedings RMRS-P-41. Fort Collins, CO: US Department of Agriculture, Forest Service, Rocky Mountain Research Station. p. 213-220*. Vol. 41.

Grazia De Paoli, R., Di Miceli, E., and Giuliani, F. (2020). "Disasters and Cultural Heritage: planning for prevention, emergency management and risk reduction". In: *IOP Conference Series: Materials Science and Engineering* 949, p. 012084.

Grimm, V., Berger, U., Calabrese, J. M., Cortés-Avizanda, A., Ferrer, J., Franz, M., Groeneveld, J., Hartig, F., Jakoby, O., Jovani, R., et al. (2025). "Using the ODD protocol and NetLogo to replicate agent-based models". In: *Ecological Modelling* 501, p. 110967.

IGN (2022). *LiDAR HD Version 1.0 - Descriptif de contenu*. fr. https://geoservices.ign.fr/sites/default/files/2025-03/DC_LiDAR_HD_1-0.pdf.

Iskandar, R., Dugdale, J., Beck, E., and Cornou, C. (2024). "Agent-based simulation of seismic crisis including human behavior: application to the city of Beirut, Lebanon". In: *Simulation* 100.4, pp. 357–377.

Jaccard, P. (1901). "Etude de la distribution florale dans une portion des Alpes et du Jura". In: *Bulletin de la Société Vaudoise des Sciences Naturelles* 37, pp. 547–579.

Kaziyeva, D., Loidl, M., and Wallentin, G. (2021). "Simulating Spatio-Temporal Patterns of Bicycle Flows with an Agent-Based Model". In: *ISPRS International Journal of Geo-Information* 10.2, p. 88.

Kirschenbnum, A. (1992). "Warning and evacuation during a mass disaster: a multivariate decision making model". In: *International Journal of Mass Emergencies & Disasters* 10.1, pp. 91–114.

Li, S., Marsaglia, N., Garth, C., Woodring, J., Clyne, J., and Childs, H. (2018). "Data reduction techniques for simulation, visualization and data analysis". In: *Computer graphics forum*. Vol. 37. 6. Wiley Online Library, pp. 422–447.

Li, Z. and Dong, W. (2010). "A stroke-based method for automated generation of schematic network maps". In: *International Journal of Geographical Information Science* 24.11, pp. 1631–1647.

Lindell, M. K. (2021). "Emergency management: evacuations". In: *Encyclopedia of Security and Emergency Management*. Springer, pp. 272–286.

Mancheva, L., Adam, C., and Dugdale, J. (2019). "Multi-agent geospatial simulation of human interactions and behaviour in bushfires". In: *ISCRAM*. Valencia, Spain.

Marshall, S., Gil, J., Kropf, K., Tomko, M., and Figueiredo, L. (2018). "Street network studies: from networks to models and their representations". In: *Networks and Spatial Economics* 18.3, pp. 735–749.

Moradi, H., Iskandar, R., Rodriguez, S., Singh, D., Dugdale, J., Tzempelikos, D., and Sfetsos, A. (2025). "A Framework for Constructing Agent-based Models for Evaluating Emergency Evacuation Procedures; A Case Study Applied to Egaleo, Greece". In: *58th Hawaii International Conference on System Sciences*.

Papasimeon, M. (2010). *Modelling agent-environment interaction in multi-agent simulations with affordances*. Tech. rep.

Pelechano, N. and Fuentes, C. (2016). "Hierarchical path-finding for Navigation Meshes (HNA∗)". In: *Computers & Graphics* 59, pp. 68–78.

Pires, R., Torres, P., Valente, N. A., Solteiro Pires, E. J., Reis, A., Oliveira, P. B. d. M., and Barroso, J. (2026). "Agent-Based Simulation of Forest Fire Spread with NetLogo". en. In: *HCI International 2025 – Late Breaking Papers*. Ed. by M. Antona and C. Stephanidis. Cham: Springer Nature Switzerland, pp. 212–224.

Porta, S., Crucitti, P., and Latora, V. (2006). "The network analysis of urban streets: a primal approach". In: *Environment and Planning B: planning and design* 33.5, pp. 705–725.

Rahmani, P., Peruani, F., and Romanczuk, P. (2020). "Flocking in complex environments—attention trade-offs in collective information processing". In: *PLoS computational biology* 16.4, e1007697.

Schelling, T. C. (1971). "Dynamic models of segregation". In: *Journal of mathematical sociology* 1.2, pp. 143–186.

Schoenharl, T. and Madey, G. (2011). "Design and implementation of an agent-based simulation for emergency response and crisis management". In: *Journal of Algorithms & Computational Technology* 5.4, pp. 601–622.

Sharma, R., Xu, Z., Sugumaran, R., and Oliveira, S. (2016). "Parallel landscape driven data reduction & spatial interpolation algorithm for big LiDAR data". In: *ISPRS international journal of geo-information* 5.6, p. 97.

Solmaz, G. and Turgut, D. (2017). "Modeling pedestrian mobility in disaster areas". In: *Pervasive and Mobile Computing* 40, pp. 104–122.

Taillandier, P., Grignard, A., Marilleau, N., Philippon, D., Huynh, Q.-N., Gaudou, B., and Drogoul, A. (2019). "Participatory Modeling and Simulation with the GAMA Platform". In: *Journal of Artificial Societies and Social Simulation* 22.2, p. 3.

Team, R. D. (Oct. 2006). *Repast Simphony Release Announcement*. https://www.gisagents.org/2006/10/repast-simphony.html.

Tisue, S., Wilensky, U., et al. (2004). "Netlogo: A simple environment for modeling complexity". In: *International conference on complex systems*. Vol. 21. Boston, MA, pp. 16–21.

Van Aalst, M. K. (2006). "The impacts of climate change on the risk of natural disasters". In: *Disasters* 30.1, pp. 5–18.

Weyns, D. and Michel, F. (2015). "Agent environments for multi-agent systems–a research roadmap". In: *Agent Environments for Multi-Agent Systems IV: 4th International Workshop, E4MAS 2014-10 Years Later, Paris, France, May 6, 2014, Revised Selected and Invited Papers*. Springer, pp. 3–21.

Weyns, D., Omicini, A., and Odell, J. (2007). "Environment as a first class abstraction in multiagent systems". In: *Autonomous agents and multi-agent systems* 14.1, pp. 5–30.

Weyns, D., Van Dyke Parunak, H., Michel, F., Holvoet, T., and Ferber, J. (2004). "Environments for multiagent systems state-of-the-art and research challenges". In: *International Workshop on Environments for Multi-Agent Systems*. Springer, pp. 1–47.

Yang, L. and Xu, J. (2024). "Utilizing Agent-Based Modeling on the GAMA Platform to Enhance Computational Urban Design". In: *GAMA Days 2024*. Ed. by A. Grignard. Online, France.