doi: 10.36689/uhk/hed/2022-01-034 HED Hradec Economic Days

Smart Cities: GIS Data for Realistic Simulations

Milan KOŘÍNEK* and Kamila ŠTEKEROVÁ

University of Hradec Králové, Hradec Králové, Czech Republic; milan.korinek@uhk.cz; kamila.stekerova@uhk.cz

* Corresponding author: milan.korinek@uhk.cz

Abstract: More and more cities are incorporating smart elements for their better management, increased service quality and resource optimization. A digital twin is suitable for creating the concept of a smart city. Smart cities, digital twins and agent-based simulations are mutually intertwined research fields. It is an accurate simulation of a real entity, thanks to which it is possible to explore and predict various situations and respond better to them. The usage of GIS data is important part of the process of development of realistic spatial simulations. GIS data is the essential source of parameters of the environment surrounding agents. This paper summarizes the usability of GIS data in agent-based simulations in the development of digital twins of smart cities. Sample traffic model in NetLogo was designed to demonstrate this approach.

Keywords: GIS; OpenStreetMap; agent-based model; digital twin; smart city; simulation

JEL Classification: C63

1. Introduction

A smart city is where the combination of modern information technology and IoT results in better governance, quality of service, security, and reduced management and resource costs (Kumar et al., 2020). According (Nam & Pardo, 2011), there are numerous objectives of smart cities such as improving convenience, preserving energy, improving water quality and air quality, recovering from disasters or collecting data for better decision making. Four major contributes of smart cities are sustainability, urbanization, quality of life and smartness (Khan et al., 2022). Recent reviews show that the research of smart city models and methodologies grows significantly, see Myeong et al. (2022), Winkowska et al. (2019).

The digital twin is the most accurate virtual copy of a real entity. It helps to understand, explore and predict the behavior of real subjects. Digital twin simulations exist in parallel with real entities (Grieves & Vickers, 2017) in our case smart cities. Other areas where digital twins can be used are engineering, robotics or healthcare (Barricelli et al., 2019; Boje et al., 2020). Agent-based simulation is the methodology successfully applicable in creation of digital twins (Clemen et al., 2021; Mykoniatis & Harris, 2021).

The aim of the paper is to show how it is possible to use existing geographical data for agent-based simulations. The first part of the paper describes GIS data and how to obtain them for the purpose of the realistic traffic simulation in NetLogo. The extraction of parameters of the urban environment from GIS data is explained, maps from OpenStreetMap serve as a source of data for our sample model. It is also explained how to modify the data in relation to

subsequent simulation scenarios. The sample model of traffic in urban area is presented to demonstrate the approach.

2. Methodology

Agent-based simulations are composed of agents and environment. Agent-based models of cities were summarized by (Crooks et al., 2021). Important trends are the development of large-scale simulations (see e.g., Santana et al. (2018), Manzoor et al. (2021), Huang et al. (2022)) and real-time GIS applications (Li et al., 2020).

Realistic simulation requires a precise specification of the environment, especially if it asked to be even an exact copy of the original (digital twin in this case). If the simulation is about a representation of building interior, it is possible to use BIM data to create environment (Kořínek et al., 2021). On the contrary for geographical area, importing data from Geographic Information Systems (GIS) is the most efficient method.

Geographic Information Systems are used for storing and visualizing data related to a specific place on Earth at a specific time. The two main representation methods are raster and vector representation of spatial entities and their relations. Both types of representation can be applied for visualization of the space. Vectors are used to display the elements, raster image data is preferred to visualize the background of the scene (Decker, 2001).

A raster representation is focused on the detailed display of the area using images and photographs. The basic building blocks are pixels that are arranged in a 2D grid. The pixel is represented by the color as in the photo and can be extended by other parameters. The raster representation does not have to be used only to display a photograph of the area, but, for example, to display temperature, population density or elevation. The raster representation is used to represent continuous space. It is easier to use raster representation for computational operations, but at the same time it is not as accurate when the scale changes.

A vector representation focuses on specific elements from the area, such as buildings, roads, rivers, etc. Because the vector representation shows the elements as separate objects, it is easier to work with and edit other parameters that are hidden under these objects and do not need to be displayed. These components are represented by three basic geometric elements (points, polygons, curves) which can change their purpose depending on the level of detail required:

- The representation of **points** may vary depending on the scale of the display area or the way the data is used. The point can represent a city or a building or even the entrance of the building. See **Figure 1**, for example of the point outside the polygon area: here the point is the center of the object.
- **Polygons** are used to represent object, e.g., buildings. Furthermore, it is possible to use polygons to display areas, similarly to raster representation, but with a discrete space (i.e., having precise dimensions and boundaries of objects).
- **Curves** consist of connections of two or more points and are used to represent roads, rivers, or the boundaries of areas, but in this case, it is only an edge, not the space inside (polygons are used for this).

Some tools can also work with another formats like 3D representations of buildings or terrain. This can be especially useful for various visualizations.

There are numerous freely available sources of GIS data such as ArcGIS HUB (ArcGIS HUB, 2022), Natural Earth Data (Natural Earth, 2022), Google Maps (Google, 2022b) Google Earth (Google, 2022a) etc.

The source of GIS data for this article is OpenStreetMap (OpenStreetMap, 2022), which is an ODbL-licensed project focusing on creating and sharing geographic data around the world. It works on the principle of Wikipedia, where users contribute, edit and use all available data, including the access to the history of edits. When viewing maps on the web, it is possible to immediately see a preview of the GIS parameters added to objects (**Figure 2**).

GIS data contains a large amount of information of different type. When creating the environment for agent-based simulation, it is advisable to use one of the commercial software designed for data preparation, e.g. ArcGIS (Esri, 2022).

Figure 1. Example of point and polygon representation of the space

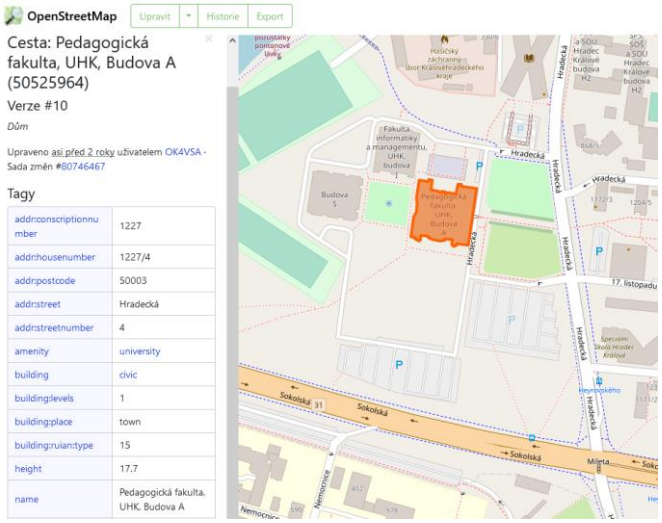

Figure 2. OpenStreetMap object with parameters. (https://www.openstreetmap.org)

Autodesk company offers AutoCad Map 3D (Autodesk, 2021) software which is suitable for 3D models or Building Information Model (BIM) data. It is also possible to use freely available software, e.g. Grass GIS (Grass GIS, 2022), gvSIG (gvSIG, 2009) or QGIS (QGIS, 2022). QGIS was created in 2002, but receives regular updates and is suitable for article sample model.

Finally, the agent-based platform is suitable for development of the smart city simulation. The overview of platforms such se NetLogo (NetLogo, 2016), AnyLogic (AnyLogic, 2022), Repast (Repast, 2021) and others (Morvaj et al., 2011; Lom & Pribyl, 2021).

3. Realistic Traffic Simulation

To demonstrate the applicability of GIS data in realistic agent-based simulation of smart cities, a sample model of urban traffic was designed in NetLogo. The model is composed of GIS-based environment and agents. The specification of the environment is as follows.

OpenStreetMap was used to get real world data (maps). Data is stored in OpenStreetMap format, which would be converted to one of the GIS formats (University of Delaware, 2022) (in article example it will be **.shp* format).

There is no need to download data from OpenStreetMap before, because QGIS, like many other software, already has built-in access to maps. Therefore, it is possible to open OpenStreetMap directly in the application, select the area and save it. This tool works similarly to graphical editors and data are divided into layers. After selecting and confirming the area, the previously mentioned geometric objects (polygons, lines and points) can be seen on the left side. The layers can be turned on and off as needed.

To obtain specific elements from OpenStreetMaps, it is recommended to use another plugin that uses query data editing. In case of traffic agent-based model, information about buildings and routes has to be obtained. In the simulation, it would be required to apply an exact representation of the routes where agents move. For our model, a simplified representation of the movement along the lines (more precisely, between the endpoints of the lines) is sufficient. This decision was made for optimization of the course of the simulation using desktop agent-based platform NetLogo. For this reason, node information needs to be attached to the data.

When working with OpenStreetMap objects, *QuickOSM* plugin is used to create layers of objects. The list of objects is written on the OpenStreetMap wiki page under the map features tab. Objects are divided here into keys (parent category) and values (sub-categories). For example, buildings and roads are required to be specified for the simulation.

The building layer will be created firstly. Buildings are marked with the keyword *building*. It is also possible to specify a sub-category (e.g. small houses or hotels correspondingly to the purpose of the simulation). In article example model, all objects from category building are required. Queries can be chained to select multiple categories or multiple subcategories. Individual selections can be combined using the logical operators *and, or*. Finally, it is also possible to select in which geographical area the objects are searched. After running the query, a new layer with the extracted data is created. **Figure 3** shows example of buildings extraction.

Further adjustments are needed for the roads. If roads were exported as they were extracted, point information would be missing. Roads would contain endpoints and some midpoints, but many would be missing. The simulation would be too simplified. QGIS includes many editing tools. One of the geometric tools is the breakpoint extraction tool. This creates a new layer with all points. This completes the editing of the roads (**Figure 4**). Other adjustments could follow, such as evenly adding more points along the way or simplifying. The path layer together with the newly created point layers can be exported as another file. The important thing is that extracted objects retain their parameters. This completes the preparation of data that can be used subsequently in simulation tools.

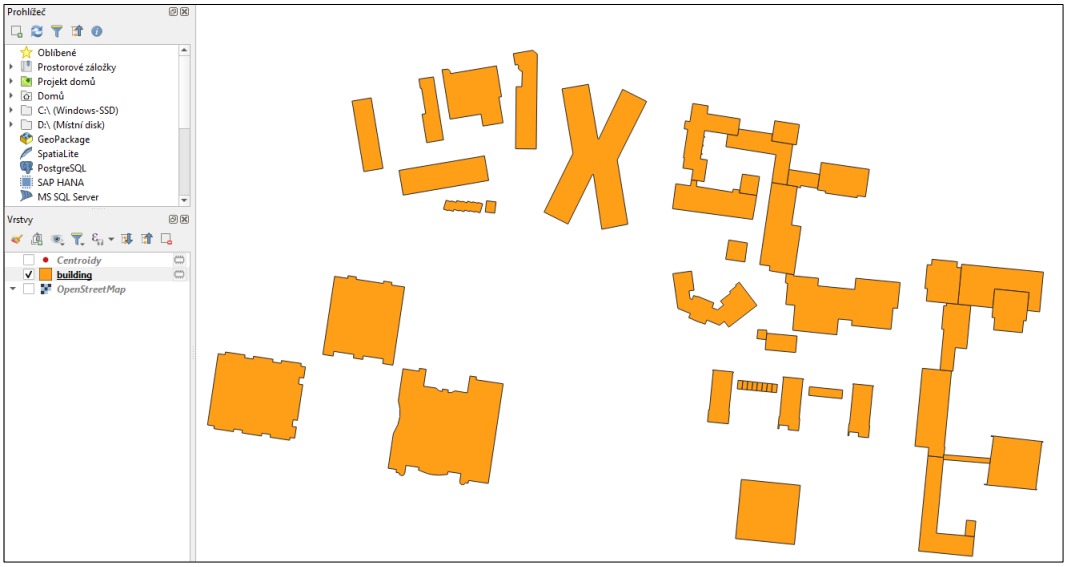

Figure 3. Extracted building layer

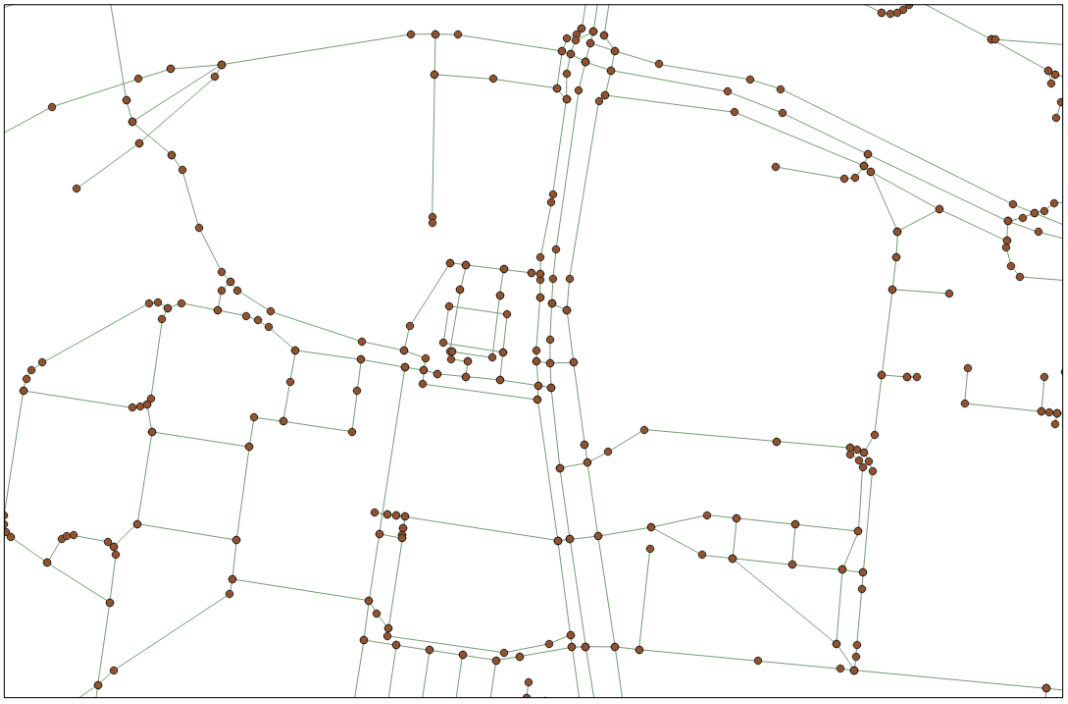

Figure 4. Extracted roads with generated breaking points

Now that the data are prepared and it is possible to start creating an agent simulation. To create a traffic model, a NetLogo was chosen. As mentioned earlier, agent simulation consists of two main parts. The environment was made from prepared GIS data. Lines and nodes were created from GIS path data using NetLogo datasets. These objects store all parameters from GIS data. Buildings are represented by GIS data without further modification. In the next step, agents need to be added. They will move and interact in the exact copy of the city thanks to the GIS data.

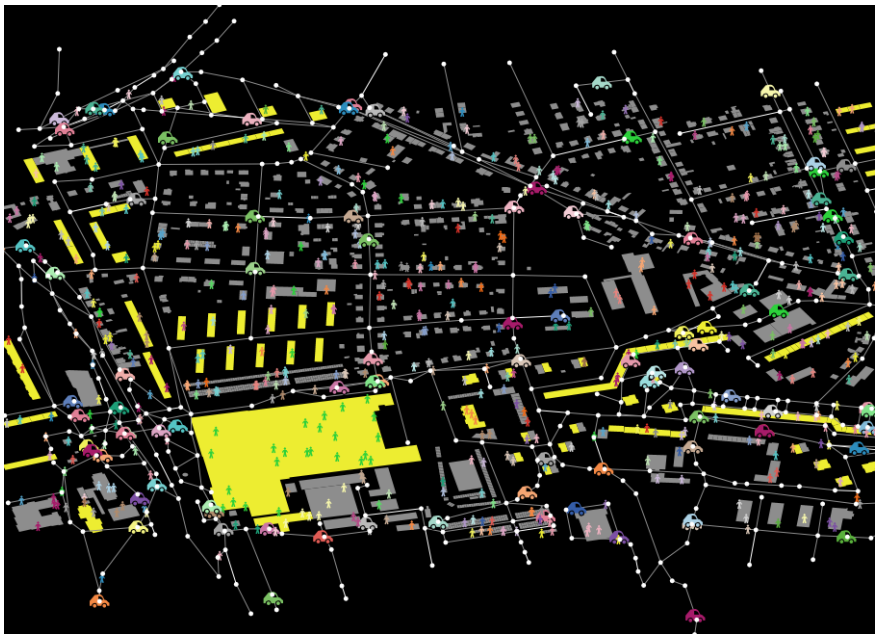

Figure 5. Traffic model

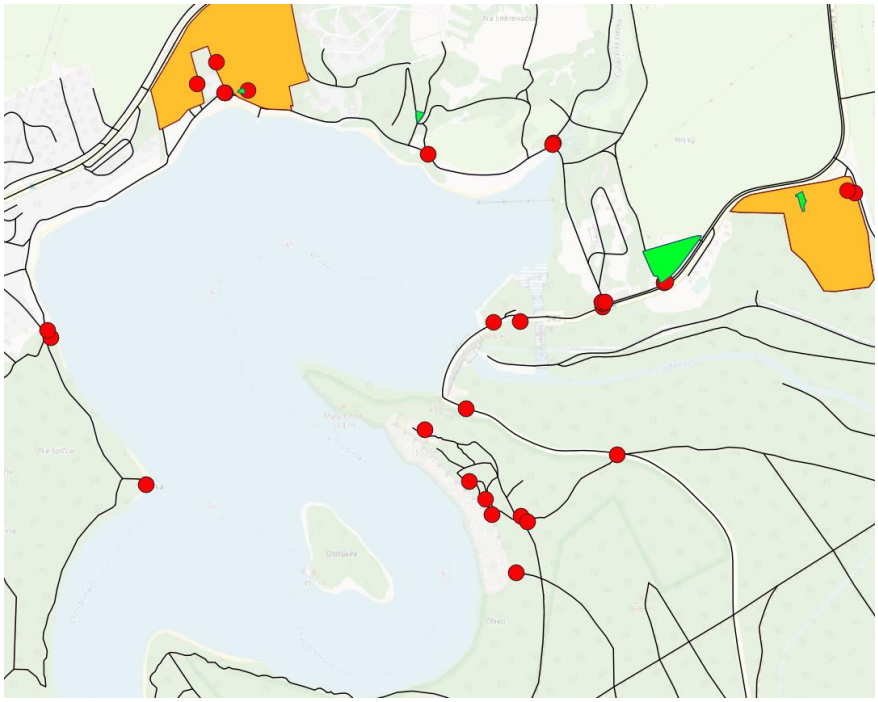

Figure 6. Extracted data of the area around the castle Oheb. Tourist places (red), parking places (green) and camps with wheelchair access (orange).

Figure 5 shows created traffic model with buildings and roads information extracted from GIS data.

Similarly, there is a simulation of tourist areas with extracted information about tourist places (**Figure 6**). Creating simulations is all the easier, because many important parameters are already contained in the data.

4. Discussion

The realistic simulations are important tool in smart cities research. The specifications of the environment have to build on GIS data. OpenStreetMap and QGIS were successfully adopted in sample model of urban area. NetLogo simulation was developed to demonstrate the applicability of agent-based approach for development of digital twin of smart city.

The next research will be focused on refinement of agents´ behavior and extending the simulation with smart city elements such as intelligent traffic lights, information systems and mobile applications.

Acknowledgments: The financial support of the Specific Research Project "Information and knowledge management and cognitive science in tourism 6".

Conflict of interest: none

References

AnyLogic. (2022). *AnyLogic.* https://www.anylogic.com

ArcGIS Hub. (2022). *ArcGIS. Hub* https://hub.arcgis.com

Autodesk. (2021). *AutoCAD Map 3D*. https://www.autodesk.cz/products/autocad/included-toolsets/autocad-map-3d

Barricelli, B. R., Casiraghi, E., & Fogli, D. (2019). A Survey on Digital Twin: Definitions, Characteristics, Applications, and Design Implications. *IEEE Access*, *7*, 167653-167671. https://doi.org/10.1109/ACCESS.2019.2953499

Boje, C., Guerriero, A., Kubicki, S., & Rezgui, Y. (2020). Towards a semantic Construction Digital Twin: Directions for future research. *Automation in Construction*, *114*, 103179. https://doi.org/10.1016/j.autcon.2020.103179

Clemen, T., Ahmady-Moghaddam, N., Lenfers, U. A., Ocker, F., Osterholz, D., Ströbele, J., & Glake, D. (2021). Multi-Agent Systems and Digital Twins for Smarter Cities. In *Proceedings of the 2021 ACM SIGSIM Conference on Principles of Advanced Discrete Simulation* (pp. 45–55). https://doi.org/10.1145/3437959.3459254

Crooks, A., Heppenstall, A., Malleson, N., & Manley, E. (2021). Agent-Based Modeling and the City: A Gallery of Applications. In W. Shi, M. F. Goodchild, M. Batty, M.-P. Kwan, & A. Zhang (Eds.), *Urban Informatics* (pp. 885–910). Springer Singapore. https://doi.org/10.1007/978-981-15-8983-6_46

Decker, D. (2001). *GIS data sources*. John Wiley & Sons.

Esri. (2022). *ArcGIS*. https://www.esri.com/en-us/arcgis/about-arcgis/overview

Google. (2022a). *Google Earth*. https://earth.google.com

Google. (2022b). *Google Maps*. https://maps.google.com

Grass GIS. (2022). *Grass GIS.* https://grass.osgeo.org

Grieves, M., & Vickers, J. (2017). Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems. In F.-J. Kahlen, S. Flumerfelt, & A. Alves (Eds.), *Transdisciplinary Perspectives on Complex Systems* (pp. 85–113). Springer International Publishing. https://doi.org/10.1007/978-3-319-38756-7_4

gvSIG. (2009). *gvSIG.* http://www.gvsig.com

Huang, J., Cui, Y., Zhang, L., Tong, W., Shi, Y., & Liu, Z. (2022). An Overview of Agent-Based Models for Transport Simulation and Analysis. *Journal of Advanced Transportation, 2022,* 1–17. https://doi.org/10.1155/2022/1252534

Khan, A., Aslam, S., Aurangzeb, K., Alhussein, M., & Javaid, N. (2022). Multiscale modeling in smart cities: A survey on applications, current trends, and challenges. *Sustainable Cities and Society*, *78*, 103517. https://doi.org/10.1016/j.scs.2021.103517

Kořínek, M., Tázlar, O., & Štekerová, K. (2021). Digital Twin Models: BIM Meets NetLogo. In *Hradec Economic Days 2021* (394–400). https://doi.org/10.36689/uhk/hed/2021-01-040

Kumar, H., Singh, M. K., Gupta, M. P., & Madaan, J. (2020). Moving towards smart cities: Solutions that lead to the Smart City Transformation Framework. *Technological Forecasting and Social Change, 153*, 119281. https://doi.org/10.1016/j.techfore.2018.04.024

Li, W., Batty, M., & Goodchild, M. F. (2020). Real-time GIS for smart cities. *International Journal of Geographical Information Science*, *34*(2), 311–324. https://doi.org/10.1080/13658816.2019.1673397

Lom, M., & Pribyl, O. (2021). Smart city model based on systems theory. *International Journal of Information Management, 56*, 102092. https://doi.org/10.1016/j.ijinfomgt.2020.102092

Manzoor, U., Kavak, H., Kim, J-S., Crooks, A.T., Pfoser, D., Zufle, A., & Wenk, C. (2021). Towards Large-Scale Agent-Based Geospatial Simulation. In *International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation*.

Morvaj, B., Lugaric, L., & Krajcar, S. (2011). Demonstrating smart buildings and smart grid features in a smart energy city. In *Proceedings of the 2011 3rd International Youth Conference on Energetics (IYCE)* (pp. 1–8).

Mykoniatis, K., & Harris, G. A. (2021). A digital twin emulator of a modular production system using a data-driven hybrid modeling and simulation approach. *Journal of Intelligent Manufacturing, 32*(7), 1899–1911. https://doi.org/10.1007/s10845-020-01724-5

Myeong, S., Park, J., & Lee, M. (2022). Research Models and Methodologies on the Smart City: A Systematic Literature Review. *Sustainability*, *14*, 1687. https://doi.org/10.3390/su14031687

Nam, T., & Pardo, T. A. (2011). Smart City as Urban Innovation: Focusing on Management, Policy, and Context. In *Proceedings of the 5th International Conference on Theory and Practice of Electronic Governance* (pp. 185–194). https://doi.org/10.1145/2072069.2072100

Natural Earth. (2022). *Natural Earth.* https://www.naturalearthdata.com

NetLogo. (2016). *NetLogo.* Available at: https://ccl.northwestern.edu/netlogo

OpenStreetMap. (2022). *OpenStreetMap*. https://www.openstreetmap.org

QGIS. (2022). *QGIS.* https://www.qgis.org

Repast. (2021). *The Repast Suit*. https://repast.github.io/index.html

Santana, E. F. Z., Lago, N., Kon, F., & Milojicic, D. S. (2018). InterSCSimulator: Large-Scale Traffic Simulation in Smart Cities Using Erlang. In G. P. Dimuro, & L. Antunes (Eds.), *Multi-Agent Based Simulation XVIII* (Vol. 10798, pp. 211–227). Springer International Publishing. https://doi.org/10.1007/978-3-319-91587-6_15

University of Delaware. (2022). *File Formats for GIS*. https://sites.udel.edu/gis/file-formats-for-gis

Winkowska, J., Szpilko, D., & Pejić, S. (2019). Smart city concept in the light of the literature review. *Engineering Management in Production and Services, 11*(2), 70–86. https://doi.org/10.2478/emj-2019-0012