

Estimating the potential number of daily visitors in urban areas. Methodical evaluation of the points of interest of open data services using the example of transport mobility in Cologne

Georg Moog , Frauke Kraas 

Received: 20 December 2024 ■ Accepted: 3 December 2025 ■ Published online: 21 January 2026

Abstract

Accurate estimates of the potential number of daily visitors are an important parameter in the planning of transportation infrastructure, allowing the development of appropriate solutions for specific planning cases. In this paper, previous methods for recording or estimating potential daily visitor volumes in urban areas in spatially differentiated analysis are extended by points of interest from open data services and supplemented by cartographic representations. The methodology addresses existing research gaps and areas for further development. For the first time, the explicit spatial contextualization of the calculated visitor volumes is addressed. The results of the calculated daily visitor volumes enable case-based planning of transport infrastructure in urban areas. Moreover, spatial visualizations of high-density hotspots can be produced to support planners lacking local knowledge. In order to critically reflect on the methodology used, the transport model of the City of Cologne serves as a benchmark for the study. Nevertheless, the relevance of the further develop-

ment of the procedure is emphasized, especially with regard to the application of adequate estimates of the potential number of daily visitors.

Keywords: Traffic modelling ■ OpenStreetMap ■ Open Data ■ estimation of visitors ■ GIS ■ Cologne

Schätzung potenziellen Personenaufkommens im urbanen Raum. Methodische Evaluierung der “Points of Interest” von Open-Data-Angeboten am Beispiel der Verkehrsmobilität in Köln

Zusammenfassung

Genaue Schätzungen des täglichen Personenaufkommens stellen im Rahmen der Planung der Verkehrsinfrastruktur einen wichtigen Kennwert dar, der es erlaubt, adäquate Lösungen für spezifische Planfälle zu entwickeln. In diesem Beitrag werden bisherige Methoden zur Erfassung bzw. Schätzung täglicher potenzieller Personenaufkommen im urbanen Raum in räumlich differenzierter Analyse um “Points of Interest” aus Open-Data-Angeboten erweitert und um kartographische Darstellungen ergänzt. Die verwendete Methodik greift hierbei bestehende Forschungs- und Entwicklungslücken auf. Erstmals wird das berechnete Aufkommen in einen räumlichen Kontext gesetzt. Die Ergebnisse der berechneten täglichen Personenaufkommen ermöglichen eine fallgerechte Planung der Verkehrsinfrastruktur in urbanen Räumen. Zusätzlich können räumliche Visualisierungen von Hotspots hoher Personenaufkommen durchgeführt werden, welche insbesondere ortsfremde Planer unterstützen können. Um die angewandte Methodik kritisch an einem Beispiel zu reflektieren, dient das Verkehrsmodell der Stadt Köln als Referenzgröße. Die Rele-

Link zum Onlinematerial: <https://rur.oekom.de/index.php/rur/article/view/3091/7251>

✉ **Georg Moog**, SSP Consult Beratende Ingenieure GmbH, Waltherstraße 49-51, 51069 Köln, Germany
moog@ssp-consult.de

Prof. Dr. Frauke Kraas, Institut für Geographie, Universität zu Köln, Albertus-Magnus-Platz, 50923 Köln, Germany
f.kraas@uni-koeln.de



© 2025 by the author(s); licensee oekom. This Open Access article is published under a Creative Commons Attribution 4.0 International Licence (CC BY).

vanz der Weiterentwicklung des Vorgehens wird insbesondere in Hinblick auf die Anwendung von adäquaten Schätzwerten des potenziellen Personenaufkommens betont.

Schlüsselwörter: Verkehrsmodellierung ■ OpenStreetMap ■ Open Data ■ potenzielles Personenaufkommen ■ GIS ■ Köln

1 Introduction

In the context of climate change, advancing a sustainable transport transition, especially in urban areas, plays an important role (Dangschat 2022: 3). Cities produce around 70 % of greenhouse gas emissions worldwide (UN-Habitat 2022: 181; Hertig/Keck 2023: 5); a significant proportion of this is attributable to transport, which generates 23 % of emissions worldwide, 75 % of which come from road traffic (Loo/Tsoi 2018: 961; Bechtel/Hüser 2023: 12). As both companies (e.g. retailers) and city residents are dependent on motorized private transport (Göttsche/Brinkmann 2023: 21), the car remains highly relevant (Topp 2023: 31). In order to develop sustainable and climate-resilient cities, there is therefore a great need for action, particularly at the level of transportation (UN-Habitat 2022: 149). Reducing the importance of the car is just as important as strengthening more sustainable alternatives, such as the environmental network of public transport, cycling and walking (Kagerbauer 2022: 31).

Scientists and politicians consider the improved use of suitable digital planning tools in the context of transport planning to be essential (Chmielewski/Kempa 2020: 103; BMDV 2023: 38). Such tools should check whether measures taken are bringing about positive changes (Ziemke/Kaddoura/Nagel 2019: 870) or help to sound out planning development options. Transport models play a special role, as they can map traffic for different sized study areas, different modes of transport and different options for action (Kagerbauer 2022: 82). Transport modelling is continuously being improved from various perspectives.

A current scientific discourse focuses on the analysis and implementation of open data services and their effects on traffic models (Klinkhardt/Kühnel/Heilig et al. 2023: 661). The focus is particularly on points of interest (POI). Points of interest can be, e.g. locations of retailers, leisure facilities and bus stops, which can be used to estimate traffic and passenger volumes (Klinkhardt/Woerle/Briem et al. 2021). These and other data from open data services can be used both to supplement existing models and to calculate independent models based on open data (Surahman/Wegner 2022: 61).

One service that is particularly discussed in scientific publications is OpenStreetMap (OSM). The database is con-

sidered to have great potential in the context of traffic modelling and other planning processes (Steiniger/Poorazizi/Scott et al. 2016: 275; Camargo/Bright/Hale 2019: 11). For this reason, OSM has already been examined in a number of publications from different scientific perspectives with regard to its suitability for traffic modelling (Steiniger/Poorazizi/Scott et al. 2016; Camargo/Bright/Hale et al. 2019; Keler/Grigoropoulos/Mussack 2019). The importance of this geodata service lies in the fact that it can be downloaded free of charge on a daily basis and updated ad hoc by thousands of users in parallel (Stengel/Pomplun 2011: 115). Compared to providers who sell data collected on a specific date for a fee or to municipal datasets, the aforementioned properties of the OSM database are of particular importance (Mahajan/Kühnel/Intzevidou et al. 2022: 430).

In this paper, the following objectives and questions are examined within the framework of the considerations outlined above. The overarching research question addresses (1) whether points of interest from open data services (such as OSM) can be used to improve the mapping of passenger transportation, taking into account all modes of transport that occur in urban areas. To this end, the state of research must first be analysed. (2) Furthermore, the relevance and potential of open data services, including the extent to which they can contribute to research on the estimation of passenger transport, will be examined. The following methodological section (3) examines whether point-of-interest datasets can be used to plausibly and verifiably localize spatial hotspots of potential daily visitor volumes within urban areas for selected points of interest. This is followed (4) by investigation of whether a distinction can be made between explicit areas of origin and destination traffic and (5) discussion of the implications that can be derived from this study for further research in the field of open data for traffic modelling and sustainable transport research.

2 State of research

2.1 Traffic modelling

In order to develop suitable adaptation and prevention strategies to reduce the effects of climate change, holistic approaches that bring together different disciplines are helpful (Rau/Scheiner 2020: 1); with regard to transport, the main relevant disciplines are spatial planning, management science and engineering (Busch-Geertsema/Klinger/Lanzendorf 2019: 1016). This paper therefore uses a methodology that combines these three perspectives (Köhler 2014: 32) in so-called transport or transport-demand models. These models are capable of determining

and mapping both current and future transport demand within a study area (Köhler 2014: 32). They represent the resulting changes in location based on the decision-making processes of road users, for both passenger and commercial/freight traffic (Köhler 2014: 32). Both traffic categories can be examined within traffic models. The present study focuses on passenger transport, which can be described as “the transportation of people between different places; it connects people and thus creates an important prerequisite for work, education, shopping, leisure and much more”.¹

Transport models are based on four basic components: the spatial distribution of the population, the locations of important activity sites (e.g. industrial areas, green spaces) and important facilities (e.g. retailers, universities), and the level of service (transport routes, frequency of public transport journeys, etc.) (Köhler 2014: 32). Based on these, two types of simplified representations of reality can be created (Mahajan/Kühnel/Intzevidou et al. 2022: 420; Surahman/Wegner 2022: 5). Firstly, microscopic models can be used, which usually examine the path of each individual road user and its effects on traffic flow in smaller study areas, e.g. at an intersection of two roads (Treiber/Kesting 2010: 53; Köhler 2014: 56; Kagerbauer 2022: 82). Secondly, macroscopic models can be created in which – for a study area such as a municipality, a federal state or even a national territory – the entirety of the routes of a group of road users is examined (Treiber/Kesting 2010: 53; Köhler 2014: 56; Kagerbauer 2022: 82). The latter models additionally offer the option of an agent-based model in which agent profiles are created for certain population groups, storing relevant characteristics, e.g. mobility behaviour, whereby the individual agent types differ.² In order to enable a description of transport demand, it is necessary to divide the respective study area into transport cells (Köhler 2014: 51). The information already mentioned is stored in these cells and thus characterizes the respective traffic cell (Köhler 2014: 51). In addition to demographic information (e.g. number of inhabitants), infrastructural information also plays a role in characterizing these traffic cells. Information such as the locations of grocery stores or leisure facilities is relevant here. For example, among other indicators, points of interest are relevant for describing and enriching the characteristics of a traffic cell; they are usually set up by the traffic planner and need to be created specifically for each study area. In addition to points of interest, roads and public transport

routes, for example, can also be used from both commercial and open data services such as OpenStreetMap.

2.2 OpenStreetMap

OpenStreetMap (OSM) is an open data service that offers geodata free of charge as part of its own geodatabase (Stengel/Pomplun 2011: 115), referred to as “Volunteered Geographic Information” (VGI) (de Lange 2020: 233; Yan/Feng/Huang et al. 2020: 1766; Mahajan/Kühnel/Intzevidou et al. 2022: 429). Volunteered Geographic Information is created through crowdsourcing, i.e. the collection of data by people who want to participate voluntarily. Volunteered Geographic Information emerged in the 2010s, with OSM being the largest Volunteered Geographic Information measured by active users of the service. OSM obtains data from its users and can therefore be regarded as an open community data platform (OCD) (Mahajan/Kühnel/Intzevidou et al. 2022: 426). Users may be private individuals who view mapping in OSM as a hobby, or experts and professionals who map for professional purposes (Reynard 2018: 2). State and municipal institutions are also of great importance. They fill the OSM database with their own collected data, e.g. from the real estate cadastre, and provide important data such as building outlines and road courses, which are collected by state surveying authorities and are therefore of high quality (Yeow/Low/Tan et al. 2021: 21). Other map services also populate the database with their own data; a well-known example of this is the Bing map service from the US company Microsoft (Stengel/Pomplun 2011: 116). Such contributions can be discussed within the OSM community before an import is carried out. Due to the variety of participants who fill the OSM database with information, it changes every minute on a global level, so that it is highly up-to-date, especially in densely populated areas (Stengel/Pomplun 2011: 115; Klinkhardt/Kühnel/Heilig et al. 2023: 672), which makes it possible to obtain daily updated datasets.

A look at the documentation maintained by the OCD project³ helps to determine which data is collected as part of OSM and displayed cartographically in a web application. All possible geo-objects recorded in OSM are listed under “Map Features”. The list of several hundred objects contains a variety of different categories.⁴ These start with different types of paths and roads and range from geological formations and different types of buildings to the mapping

¹ https://www.destatis.de/DE/Themen/Branchen-Unternehmen/Transport-Verkehr/Personenverkehr/_inhalt.html (13.11.2025), own translation.

² <https://www.ptvgroup.com/de/anwendungsfaelle/verkehrsmodellierung> (13.11.2025).

³ https://wiki.openstreetmap.org/wiki/Main_Page (13.11.2025); https://wiki.openstreetmap.org/wiki/DE:Map_Features (13.11.2025).

⁴ https://wiki.openstreetmap.org/wiki/DE:Map_Features (13.11.2025).

of individual vending machines.⁵ These different geo-objects can be stored either as points (nodes in the context of OSM), polylines (ways in the context of OSM), polygons (closed ways where the beginning and end vertex is identical) or as complex structures or relations (relations in the context of OSM) (Zhang/Pfoser 2019: 4; Klinkhardt/Woerle/Briem et al. 2021: 296). Complex objects are either polygons that represent e.g. buildings or meaningful connections between individual objects, e.g. a polygon typifying a store that is located within a polygon of a shopping centre (Xu/Lin/Lu et al. 2016: 37; Klinkhardt/Woerle/Briem et al. 2021: 296). Information about the object is stored in the database as part of a key-value principle. An example of a key would be “store” and for a matching value “bakery”.⁶ This principle is applied continuously across all categories in order to structure the stored information.

Thanks to this simple principle, OpenStreetMap is considered to be very beginner-friendly; initial mapping by inexperienced users is quite simple (Stengel/Pomplun 2011: 115). The complexity involved in contributing information may differ based on the platform utilized. However, this is also one of the main points of criticism. For example, Mahajan/Kühnel/Intzevidou et al. (2022: 429) criticize the lack of quality control on the part of OSM. Since explicit definitions are not available for all potential values, geo-objects may be incorrectly mapped. At the same time, general mapping errors can also occur without missing definitions (Martinelli 2018; Surahman/Wegner 2022: 60). Since anyone can support the project by mapping, there is also a risk of vandalism through deliberately incorrect mapping (Liu/Long 2016: 355). Inconsistencies between the time of upload and data access also represent a potential error for studies that require a high degree of timeliness, e.g. if certain regions have not been processed for a long time (Liu/Long 2016: 355). This is also linked to an issue described by Yeow/Low/Tan et al. (2021: 22), namely the problem of “spatial bias” (Klinkhardt/Kühnel/Heilig et al. 2023: 662). There are preferred regions in which regular updates are made. These can be cities, for example, which experience more frequent changes than rural areas due to the higher frequency of potential users and higher population densities.

Thus, in contrast to proprietary data sources such as Google, OSM provides only limited assurance with respect to the accuracy of its data (Stengel/Pomplun 2011: 117). Despite these limitations, various advantages can be ob-

served, particularly in the context of transport modelling, as outlined below.

2.3 Open data in traffic modelling

In order to enable sustainable transport planning, it is necessary to adequately map transport supply and demand in actual and target states (forecast case) (Chmielewski/Kempa 2020: 103; Kagerbauer 2022: 1) and to have current and updated datasets (Kagerbauer 2022: 55; Klinkhardt/Kühnel/Heilig et al. 2023: 661). The dynamics in small-scale retail, gastronomy, leisure facilities etc. in particular are very high in urban areas, which can have a lasting impact on traffic flows. Municipal datasets in the context of irregular but large-scale mapping cannot reflect these dynamics in a timely manner (Stengel/Pomplun 2011: 117), so at the beginning of the 2010s the first tests were conducted to see whether datasets from open data services, primarily OSM, could fill this gap (Zilske/Neumann/Nagel et al. 2011: 126). Since the late 2010s and early 2020s, such services have been increasingly used in transportation models (Surahman/Wegner 2022: 23–26).

Microscopic traffic models increasingly employ data on traffic infrastructure with regard to traffic light systems (Ziemke/Braun 2021: 745), littering of road sections (Cai/Lee/Luo et al. 2020: 4025), accessibility of footpaths (Cohen/Dalyot 2020: 1264) or cycling infrastructure (Fenster/Fischer/Managh et al. 2020: 64; Vierø/Vybornova/Szell 2024: 512). The primary aim is to enrich existing models with additional data (Keler/Grigoropoulos/Mussack 2019: 1) or to create new, more detailed models so that accident blackspots can be identified retrospectively or preventively (Arase/Wu/Migita et al. 2022). Vierø/Vybornova/Szell et al. (2024: 525) emphasize that municipal data on cycling infrastructure in Denmark could be enriched and gaps in the network closed with the help of OSM.

Macroscopic modelling is conducted in a variety of topics worldwide. For example, Balac/Hörl (2021) used open data to generate a synthetic population model based on mapped residential buildings in San Francisco and San Diego (USA), while Surahman/Wegner (2022) used procedures in Uppsala (Sweden). The mobility behaviour of individual agents was generated in the macroscopic, synthetic agent model, which coincides with values from population data collected by municipalities (Balac/Hörl 2021: 13). The research group comes to similar conclusions using the example of São Paulo (Brazil) (Sallard/Balac/Hörl 2020: 18).

However, German-speaking countries are pioneers in the implementation of OSM in macroscopic models. In addition to the aforementioned research group at the Federal Institute of Technology Zurich (ETH), the German Research

⁵ https://wiki.openstreetmap.org/wiki/DE:Map_Features (13.11.2025).

⁶ https://wiki.openstreetmap.org/wiki/DE:Map_Features (13.11.2025).

Foundation (DFG) is funding projects.⁷ Two research papers are particularly noteworthy with regard to the quality of the OSM dataset (Klinkhardt/Kühnel/Heilig et al. 2023) and the calculation of the number of people using OSM-POI (Klinkhardt/Woerle/Briem et al. 2021).

The MATSim (Multi-Agent Transport Simulation) project Open Berlin created an agent-based traffic model based solely on open data (Ziemke/Kaddoura/Nagel 2019). In addition to OSM data, freely accessible data from other providers was also used, such as municipal data, data from Berlin's public transport companies and census data. It was possible to create a model that can be considered realistic and accurately reflects the traffic flows within the Berlin metropolitan region. For example, the morning and evening peak times are realistically modelled, and a distinction can be made between working days, public holidays and weekends (Ziemke/Kaddoura/Nagel 2019: 874). Open data services could therefore serve as a useful alternative to calculating traffic models.

Studies by teams from the Karlsruhe Institute of Technology, such as Briem/Heilig/Klinkhardt et al. (2019), Klinkhardt/Woerle/Briem et al. (2021) and Klinkhardt/Kühnel/Heilig et al. (2023), take a more differentiated approach to the quality of the datasets used. Here, the relevance of points of interest from OSM for the creation of traffic models is emphasized, their importance for urban and regional development is underlined and, in particular, the quality of the OSM database is discussed with regard to its impact on the authenticity of the calculated models (Briem/Heilig/Klinkhardt et al. 2019: 112; Klinkhardt/Kühnel/Heilig et al. 2023: 671). In particular, the coverage of retail facilities, restaurants and leisure activities is highlighted positively. Deficits within OSM relate in particular to the fact that businesses such as offices, doctors or small trades are not accurately recorded due to their low public presence, which makes it less likely that they are visible within OSM. This was established in a Germany-wide comparison by Klinkhardt/Kühnel/Heilig et al. (2023) with the help of ground truth data, but also previously in a detailed study on Karlsruhe (Briem/Heilig/Klinkhardt et al. 2019). Nevertheless, it is emphasized that much more accurate OSM data is available for certain POI categories, such as retail, especially in comparison to municipally collected datasets that cannot be updated regularly (Stengel/Pomplun 2011: 115; Klinkhardt/Woerle/Briem et al. 2021: 300). The results underline the great potential of OSM for creating

traffic models, especially in combination with supplementary geodata (Klinkhardt/Woerle/Briem et al. 2021: 301). This is particularly evident in the methodology used to calculate the potential number of daily visitors per point of interest in Karlsruhe (Klinkhardt/Woerle/Briem et al. 2021: 296–299). Here, the points of interest from OSM, including the building polygons, were used to estimate or model the potential number of visitors using factors according to Bosserhoff (2022) based on the area per point of interest. This collection of conversion factors for estimating the potential number of visitors is Germany's most comprehensive and covers a wide range of facilities (Klinkhardt/Woerle/Briem et al. 2021: 299). The same authors criticize the determination of the exact area per facility, especially with regard to the sales area (Klinkhardt/Woerle/Briem et al. 2021: 298). As this can extend over several floors, the authors suggest using supplementary data such as height models to identify the number of floors and thus determine more accurate sales areas.

Even if the quality of the OSM dataset can be further improved, as can the still incomplete estimate of the number of visitors, the calculated estimates represent major improvements. Full automation of the methodology would also be of interest, as this would enable dynamic consideration of temporal changes for modelled potential person volumes (Klinkhardt/Woerle/Briem et al. 2021: 301–302). However, the current licensing of the OSM database still poses a hurdle, as all details would have to be published to allow an automated process. Commercial implementation with OSM data is therefore largely ruled out for the time being, except in specific cases (Klinkhardt/Woerle/Briem et al. 2021: 302).

The methodology briefly described here and utilized in this research project is similar to that used by Surahman/Wegner (2022) for an open data traffic model in Uppsala, Sweden. Klinkhardt/Woerle/Briem et al. (2021: 301) highlight a research gap in the spatial analysis of estimated passenger volumes using OSM data, emphasizing that points of interest of the same category, such as swimming pools, must be evaluated differently depending on their location within the study area, as this can significantly affect their assigned value. From a geographical and planning perspective, it would be interesting to determine whether areas of high attractiveness can be identified within the study areas. For this purpose, it would make sense to divide the study area into smaller traffic cells in order to gain detailed insights (Peter 2021: 151). Chmielewski/Kempa (2020) tested this by placing a hexagon grid over the study area and calculated the potential traffic volume for each hexagon using points of interest from OSM.

As discussed above, the use of point-of-interest data has already been addressed in a number of scientific studies.

⁷ e.g. <https://gepris.dfg.de/gepris/projekt/415208373?context=projekt&task=showDetail&id=415208373&> (13.11.2025) or <https://www.mos.ed.tum.de/vvs/forschung/projekte/past-projects/trampa/> (13.11.2025).

In particular, the work of Klinkhardt/Woerle/Briem et al. (2021) and Surahman/Wegner (2022) can be highlighted. They primarily focus on the numerical dimension of POI-based analyses regarding the calculated potential number of visitors. Qualitative investigations of OSM data have also been undertaken from various perspectives (Klinkhardt/Kühnel/Heilig et al. 2023).

From the standpoint of spatial, urban and transport planning, the spatial distribution of the calculated visitor volumes is of particular relevance. But cartographic representation and analysis of these results have not yet been systematically addressed in the literature. It may be assumed that spatial depiction could reveal distinct hotspots which may hold significant potential for supporting planning processes.

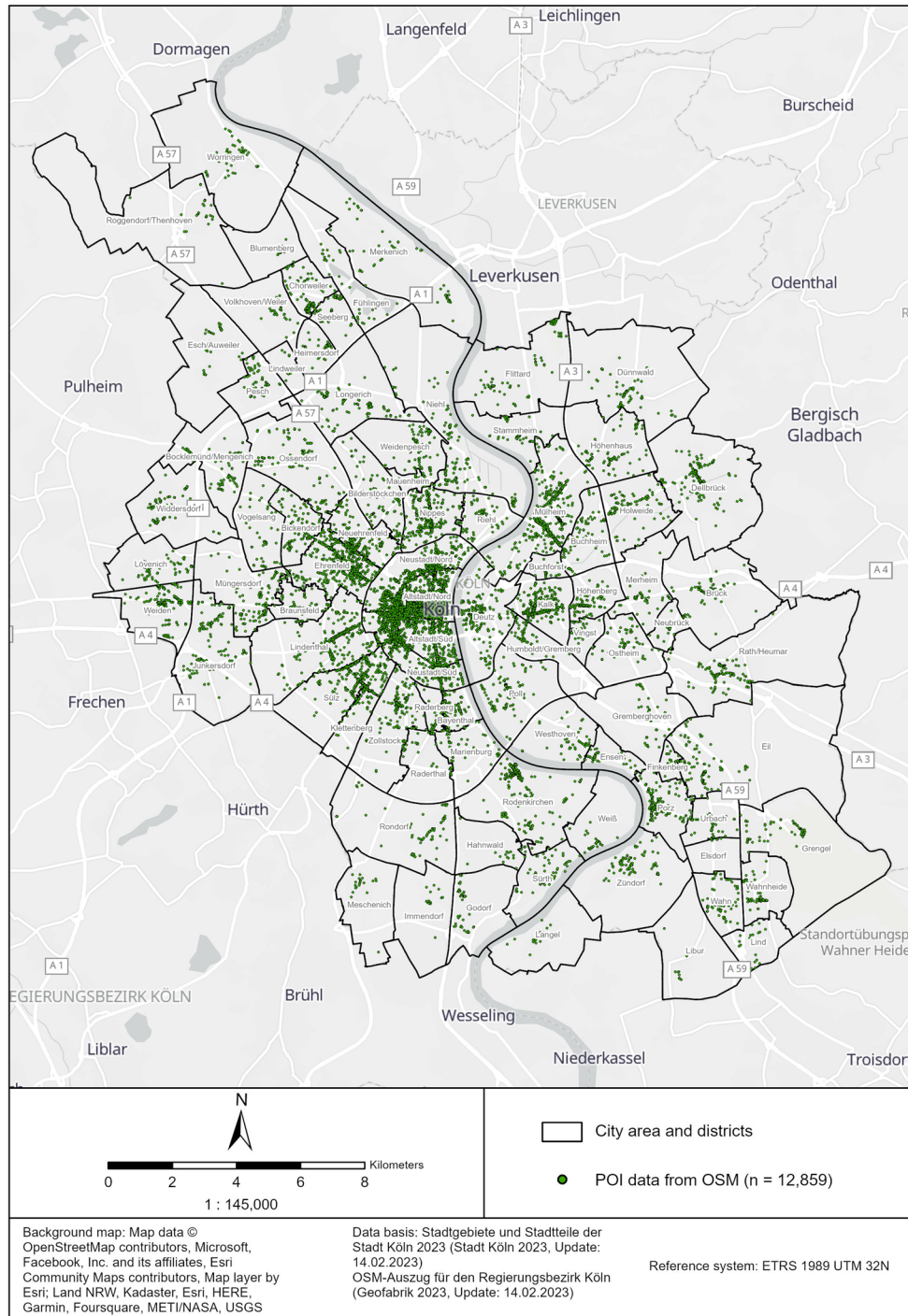


Figure 1 Overview map of the study area: the City of Cologne

3 Study area, data and methodology

3.1 Study area

The City of Cologne served as the study area for determining the potential number of visitors with the help of OSM. With 1,095,500 inhabitants (2023)⁸ in an urban area of approximately 405 km² (Stadt Köln 2022: 3), Cologne is crossed by numerous transport axes of different modes of transport, many of which are highly relevant for the transportation of people and goods (Figure 1). Around 308,000 people commute to the growing business location every day (2023).⁹ The City of Cologne is also a major residential location, with 139,023 people commuting to the surrounding area every day.¹⁰ Residential areas in high-density and central locations (such as the Altstadt-Nord or Altstadt-Süd districts) reach densities of up to 13,173 people/km²; in suburban locations (such as Roggendorf/Thenhoven in the northwest) up to 348 people/km² in 2022 (Stadt Köln 2023a: 4–6). Urban transport plays a key role in efforts to develop Cologne in a more sustainable and future-oriented way (Stadt Köln 2020), with the City of Cologne aligning itself with the sustainability goals of the United Nations and the goals of the Leipzig Charta (Stadt Köln 2020: 48, 107). To this end, environmentally friendly means of transportation are to be strengthened and street spaces redesigned for improved quality of life (Stadt Köln 2020: 127). From the perspective of the City of Cologne, valid and up-to-date tools are required for the extensive planning of transport infrastructure (Stadt Köln 2023b: 6–11). This is where determining potential visitor volumes based on the most up-to-date geodata comes into play.

3.2 Software and data

The two geographic information systems ArcGIS Pro and QGIS were used to calculate the potential number of daily visitors; the final cartographic products were created using ArcGIS Pro. In addition, the use of Microsoft Excel supported calculation of the number of daily visitors per point of interest. The OSMOSIS tool was used to process the

OSM database extracts,¹¹ similarly to other publications, e.g. Klinkhardt/Woerle/Briem et al. (2021: 297). This tool can perform SQL-like queries to extract specific geospatial objects from the OSM database.

Data was drawn primarily from an extract from the OSM database, which was downloaded from the Geofabrik website for the Cologne district government (OSM-PBF file and shapefiles).¹² As in Klinkhardt/Woerle/Briem et al. (2021: 299), the values according to Bosserhoff (2022) were used as estimated values to calculate the potential number of visitors. These can be determined with the help of a factor and the corresponding area of a certain facility. In addition, data from SSP Consult was used¹³ in which the number of people calculated by this company for each 100 m x 100 m INSPIRE grid cell was recorded for the areas of work and living. This dataset was used by the company to create the North Rhine-Westphalia state traffic model and the traffic model for the City of Cologne. With regard to research question 4, it clarifies whether the OSM points of interest can be used to explicitly differentiate between source and destination traffic according to passenger traffic and residential data. In addition to this INSPIRE raster, which is already filled with structural data, an empty INSPIRE raster dataset from the Federal Agency for Cartography and Geodesy (Bundesamt für Kartographie und Geodäsie) was used.¹⁴

Further datasets from the City of Cologne (Stadt Köln) were used to check the plausibility of the calculated results, namely, the current traffic model of the City of Cologne (Stadt Köln 2023c; provided by the Office for Sustainable Mobility Development (Amt für nachhaltige Mobilitätsentwicklung)). In addition to this, several freely available administrative boundaries of the City of Cologne were used for the subsequent visualization and analysis of the results,¹⁵ as well as an excerpt from the point-of-interest database of

⁸ [https://www.stadt-koeln.de/politik-und-verwaltung/presse/mitteilungen/26657/index.html#:~:text=Die%20Stadt%20K%C3%B6ln%20hat%20das,1.088.964\)%20angestiegen%20ist](https://www.stadt-koeln.de/politik-und-verwaltung/presse/mitteilungen/26657/index.html#:~:text=Die%20Stadt%20K%C3%B6ln%20hat%20das,1.088.964)%20angestiegen%20ist) (17.11.2025).

⁹ <https://statistik.arbeitsagentur.de/DE/Navigation/Statistiken/Interaktive-Statistiken/Pendleratlas/Pendleratlas-Nav.html> (17.11.2025).

¹⁰ <https://statistik.arbeitsagentur.de/DE/Navigation/Statistiken/Interaktive-Statistiken/Pendleratlas/Pendleratlas-Nav.html> (17.11.2025).

¹¹ <https://github.com/openstreetmap/osmosis/releases/tag/0.48.3> (17.11.2025).

¹² <http://download.geofabrik.de/europe/germany/nordrhein-westfalen/koeln-regbez.html> (17.11.2025).

¹³ Bevölkerung je INSPIRE-Rasterzelle, Beschäftigte je INSPIRE-Rasterzelle, Bildungseinrichtungen und deren Nutzerinnen/Nutzer für das Land Nordrhein-Westfalen (vgl. SSP Consult 2021a; SSP Consult 2021b; SSP Consult 2021c).

¹⁴ <https://gdz.bkg.bund.de/index.php/default/inspire/sonstige-inspire-themen/geographische-gitter-fur-deutschland-in-lambert-projektion-geogitter-inspire.html> (17.11.2025).

¹⁵ <https://www.offenedaten-koeln.de/dataset/stadtgebiet-k%C3%B6ln> (17.11.2025); <https://www.offenedaten-koeln.de/dataset/stadtbezirke-k%C3%B6ln> (17.11.2025); <https://www.offenedaten-koeln.de/dataset/stadtteile-k%C3%B6ln> (17.11.2025).

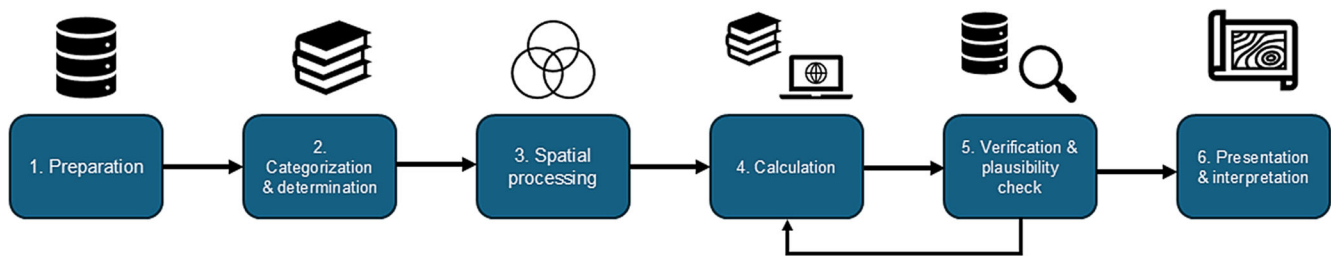


Figure 2 Workflow of the methodology

the Office for Real Estate, Surveying and Cadastre¹⁶ (Amt für Liegenschaften, Vermessung und Kataster) and location data from the City of Cologne’s business register.¹⁷

Thus, the data records mentioned first are used to calculate the potential number of daily visitors per point of interest; the data records mentioned last check the plausibility of the results. With regard to the data protection relevant to the business register, it should be emphasized that this dataset was processed in the offices of the City of Cologne. Additionally, scientific and other supplementary literature was also used to justify the methodological approach.

3.3 Methodology

The methodology can be divided into six overarching steps (Figure 2). Several steps based on Klinkhardt/Woerle/Briem et al. (2021) were adapted by the authors where necessary to solve any problem that arose. The first two phases of “preparation” and “categorization & determination” were interlinked. During the preparation of the two OSM raw datasets in shapefile and OSM-PBF format, irrelevant object keys (e.g. “aeroway” for airport runways) and further OSM keys were filtered out. The relevance of a category was determined on the basis of a higher-level categorization made in advance, and 14 potential travel purposes were determined, drawing on Klinkhardt/Woerle/Briem et al. (2021: 298) and FGSV (2022: 96, 120–125) (Table 1).

If an OSM key and its value could be assigned to one of these categories, points of interest belonging to this key with the respective value were used. For example, the combination “amenity=bar” is assigned to trip purpose 5 “Restaurants & cafés”. Following this step, OSMOSIS was used to query the selected points of interest from the OSM PBF file.¹⁸ An interim check using an OSM background map

then revealed that not all points of interest, such as well-known large supermarkets, were extracted by the OSMOSIS queries, likely due to incorrect queries within OSMOSIS. All relevant points of interest were also extracted from the downloadable POI shapefile from Geofabrik.¹⁹ The points of interest from both datasets were then transferred to the same layer. Duplicates were removed using identical object ID. Several plausibility checks were carried out to ensure that the points of interest were correct. For example, in some cases swimming pools on private properties had been categorized as public swimming pools in OSM. Points of interest with these or similar errors were removed or manually corrected if necessary. In addition, superfluous points of interest, such as individual animal enclosures in the zoo, were also removed manually, as the entire zoo’s potential number of visitors was calculated rather than the potential number of visitors to any one enclosure.²⁰ This procedure resulted in the identification of 12,859 points of interest for which the number of visitors had to be estimated (Figure 1).

Following the plausibility check of the points of interest, a third step involved their spatial processing. To estimate the potential number of individuals associated with each point of interest, information such as sales area, retail space or similar metrics is required. These values were derived using building polygons available in the OSM dataset. In addition, other spatial units, such as polygons representing parks, were incorporated where relevant. A spatial intersection was performed to determine which points of interest are located within which polygons, thereby enabling the assignment of area values. Based on the resulting area and the type of point of interest, the potential number of visitors was calculated using POI-specific factors as proposed by Bosserhoff (2022). In cases where multiple points of interest were located within the same polygon, such as a building containing both a café and a bakery, the area was dis-

¹⁶ <https://www.stadt-koeln.de/service/adressen/00070/index.html> (17.11.2025).

¹⁷ Unternehmensregister Köln (Registerabzug September 2021).

¹⁸ A list of all queries used in OSMOSIS to extract data from the OSM PBF file can be found in supplement material S-1, including additional OSM categories.

¹⁹ <http://download.geofabrik.de/europe/germany/nordrhein-westfalen/koeln-regbez.html> (17.11.2025).

²⁰ A list of such cases, as mentioned, can be found in supplement material S-2 with screenshots and brief explanations.

Table 1 Travel purposes

No. Category	Purpose of journey	Number of points of interest (n =)	Average value for the calculated potential number of daily visitors	Median of the potential number of daily visitors
1	Daily requirement (supermarkets, dis-counters, kiosks, chemists, etc.)	1,889	696	500
2	Short to long-term requirements (books, clothing, DIY stores, furniture, etc.)	2,691	133	57
3	Petrol stations and motor vehicle needs	185	486	176
4	Sports grounds and green spaces	1,405	309	15
5	Restaurants and cafés (gastronomy)	3,242	310	192
6	Hotels and other accommodation	352	107	47
7	Cultural events	147	212	45
8	Cultural education	234	233	61
9	Other leisure activities	111	119	48
10	Religious facilities	410	136	14
11	Medical facilities including pharmacies	840	260	50
12	Post offices	99	1,006	992
13	Hairdressers and similar service providers	1,071	261	205
14	Banks and other financial services	183	50	50
	Total	12,859		

Source: authors' table, based on Klinkhardt/Woerle/Briem et al. (2021: 298) and FGSV (2022: 96–125)

tributed proportionally among them.²¹ Here too, plausibility checks were necessary to ensure that the assigned area values were realistic and appropriate for the respective points of interest. The fourth step involved calculating the number of daily visitors. For this, the correct estimating values first had to be assigned to the respective OSM point of interest. Bosserhoff (2022) provides an estimated value or factor to be multiplied by the area of a facility and for all points of interest.²² For example, if the calculated area of a supermarket is 2,480 m², this value must be multiplied by a factor of 0.30–0.45, which suggests an estimated number of 744 to 1,116 customers within 24 hours.

Nevertheless, the approach used by Bosserhoff (2022) also has clear limitations, which is why approximations had to be made using researched or specially determined estimates. For hospitals whose point of interest was determined by OSM, the respective quality reports were used. These contained the respective number of outpatients per year, so that the daily volume per hospital could be estimated. The same applies to the hospitality industry. Here, a methodology had to be used to obtain an estimate. In addition to medical facilities of all kinds and restaurants, the authors developed their own methodologies to estimate the

number of potential visitors to hotels, banks and kiosks.²³ In the fifth step (linked to step 4 in an iterative process), a verification and plausibility check of unrealistically calculated passenger volumes took place. Irregularities were, for example, unrealistically high volumes of visitors to individual facilities, such as a bakery with several tens of thousands of potential customers per day. The sixth and final step involved the presentation and interpretation of the results. For this purpose, the calculated visitor volumes were transferred to the points of interest in the GIS, then the points of interest per 100 m x 100 m grid cell were added up with the help of the INSPIRE grid, enabling a spatial representation of the potential number of daily visitors. In addition, the statistics of residents and workers per INSPIRE grid cell from SSP Consult were used to spatially determine source and destination traffic.²⁴ The traffic cells of the traffic model of the City of Cologne (Stadt Köln 2023c) were supplemented with the calculated passenger volume in order to be able to compare the values of the traffic model and the calculated values. Finally, the OSM points of interest for all districts of Cologne were compared with the points of interest of the

²¹ A visual example can be found in supplement material S-2.

²² The assignment of values according to Bosserhoff (2022) for five example OSM points of interest can be found in supplement material S-3.

²³ An explanatory description of the provisional approximation can be found in supplement material S-4.

²⁴ Bevölkerung je INSPIRE-Rasterzelle, Beschäftigte je INSPIRE-Rasterzelle, Bildungseinrichtungen und deren Nutzerinnen/Nutzer für das Land Nordrhein-Westfalen (vgl. SSP Consult 2021a; SSP Consult 2021b; SSP Consult 2021c).

City of Cologne (Stadt Köln 2021; Stadt Köln 2023d) in order to be able to compare the timeliness and precision of the OSM dataset with the municipal dataset. The resulting cartographic results are presented in the following section.

4 Results

A total of 12,859 points of interest from the OSM database were included in the calculation of the number of daily visitors (Table 1). The largest number of points of interest was assigned to the Restaurants & Cafés category (3,242); the lowest number was in the Post Office category with 99 points of interest. Despite the low number, these points of interest had the highest median number of potential visitors at 992. Points of interest in the Daily Needs category, such as supermarkets or drugstores, were the second most common with 500 daily visitors. Points of interest in the Religious Facilities category, e.g. churches, had the lowest median number of potential daily visitors.

The number of people was aggregated into the INSPIRE 100 m x 100 m grid (Figure 3). Only grid cells with number of daily visitors > 0 were displayed. The maximum calculated value for a grid cell is 50,000 visitors within 24 hours. In Figure 3, it can be seen that clear clusters of grid cells form throughout Cologne's urban area, with a high number of daily visitors marked in dark blue shades. Some clusters have linear structures that correspond to streets.

The points of interest representing the trip purpose Daily Needs and Short to Medium-term Needs are located within these clusters; these are inner-city local shopping centres along streets, recognizable by the red outlines and names of the respective inner-city centres. This assignment was performed manually by the authors, drawing on their contextual knowledge of the location and their interpretation of the calculated map results. One example is Venloer Straße in the district of Ehrenfeld. Furthermore, the large central clustering of numerous grid cells in the upper quantiles with a high potential number of daily visitors in Cologne's densely populated city centre is evident.

In summary, it can be said that with the help of the calculated number of daily visitors, clear local supply centres can be identified. However, these may also be leisure areas, such as the "Kwartier Latäng" student quarter bordering the city centre to the west or the Müngersdorf sports park in the west of Cologne.

To address the question of whether it is possible to identify a distinction between origin and destination traffic, a bivariate comparison was made between the combined number of people living, working and studying in each grid cell in blue and the calculated number of potential daily visitors to the points of interest in red (Figure 4). A clear distinction

can be made between locations with a high volume of people related to living, working and education, and locations with a high calculated volume of daily visitors to points of interest. Although no widespread red-coloured areas can be observed, smaller red clusters emerge that correspond to the clusters depicted in Figure 3. However, it is clear that particularly in the city centre and within the Altstadt-Nord and Altstadt-Süd districts, there are high and almost equivalent characteristics, recognizable in the large accumulation of many raster cells in shades of grey. These clusters of many grid cells coloured grey can also be assigned to the clusters in Figure 3. It should be emphasized that in many cases isolated red clusters represent larger leisure facilities, such as Fühlinger See in the north or Sportpark Müngersdorf in the west. The latter is located between areas with grid cells in shades of blue, i.e. for living, working and education. In conclusion, many of the area-wide blue shading can be assigned to residential areas, commercial areas or larger educational facilities. Clear source and destination traffic areas (Figure 3 and 4) can be matched to the clusters with approximate volumes with the help of the calculated passenger volumes based on the OSM points of interest.

The difference between the calculated passenger volume to points of interest and the destination traffic from the traffic model of the City of Cologne per traffic cell can be depicted, as seen in Figure 5. For this purpose, the calculated passenger volume was subtracted from the destination traffic per traffic cell, with a lower volume shown in red and a higher volume in blue. Traffic cells with no difference are shown in white; those in which no points of interest were present are not taken into account. The largest difference in terms of an undercalculated traffic volume can be seen in the centre of Cologne with a difference of -11,847 visitors; the largest difference for an overcalculated traffic volume based on the OSM points of interest is in the west of Cologne with a difference of +55,852 visitors. The first example is the Schildergasse shopping street and the second case is the Müngersdorfer Sportpark, where the RheinEnergieSTADION and a popular swimming pool are located. Particularly in the peripheral districts, the estimated passenger volumes closely correspond to the results of the traffic model. In contrast, there are clear differences in the central inner-city locations in particular, with significantly higher calculated values, recognizable by the large number of dark blue traffic cells.

Further, the difference between the number of points of interest from the OSM database and the aggregated data for the City of Cologne are calculated (Figure 6). Positive values are shown in blue (more points of interest in OSM) and negative values in red (more objects in the aggregated data). Due to the larger number of districts in red tones, it can be seen that in many cases the aggregated data of the

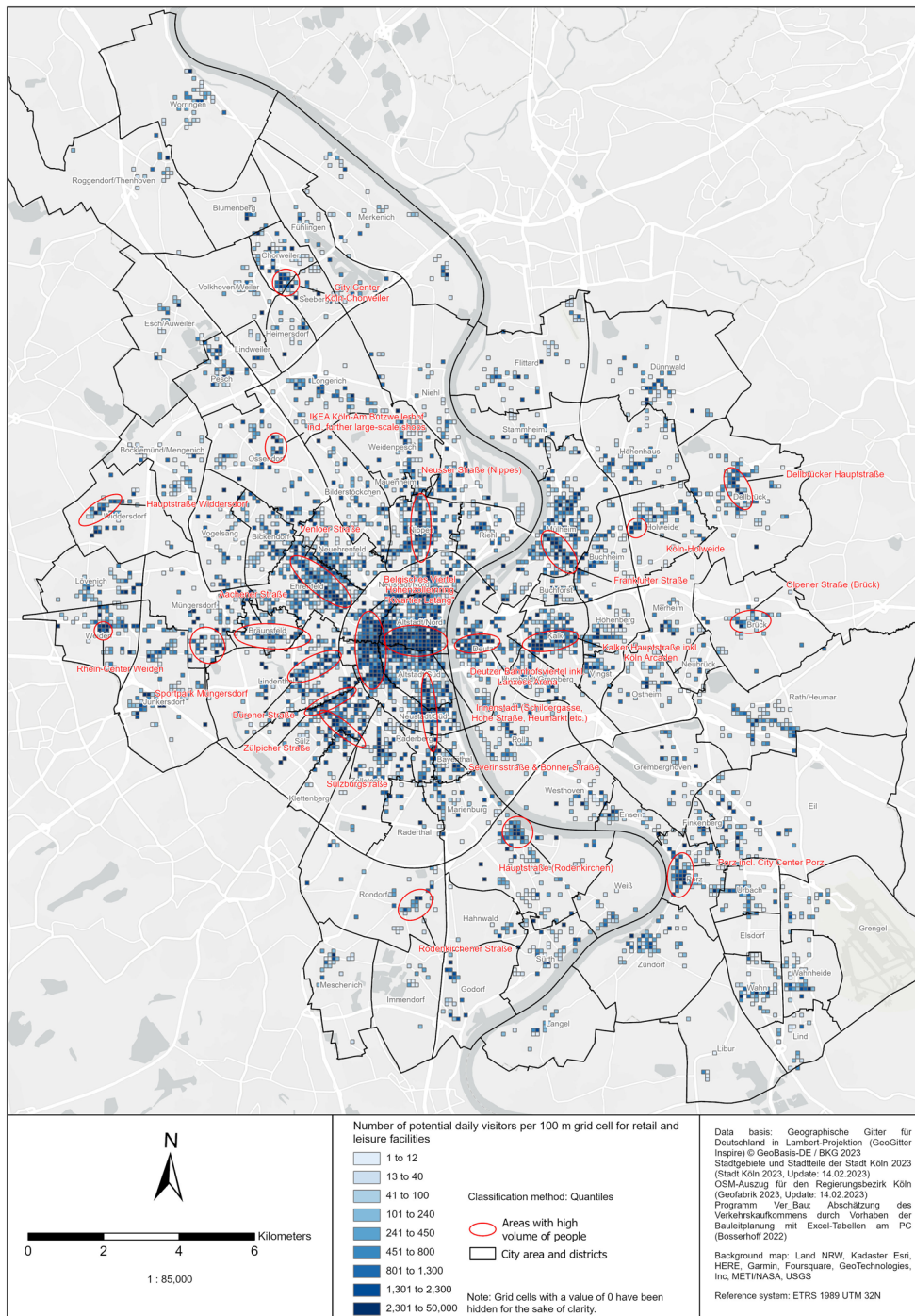


Figure 3 Recognizable structures based on the calculated potential number of visitors in the area of the City of Cologne

City of Cologne provides more detailed information than the OSM database regarding the selected travel purposes from Table 1. In some cases, up to 325 additional objects can be found in the City of Cologne's datasets. Nevertheless, six districts have a higher number of OSM points of interest such as the inner-city districts of Altstadt-Nord and

Deutz with values of up to 232 facilities. In the north-west and south-east of Cologne, several districts also show slight differences, ranging from 1 to 17 mentions.

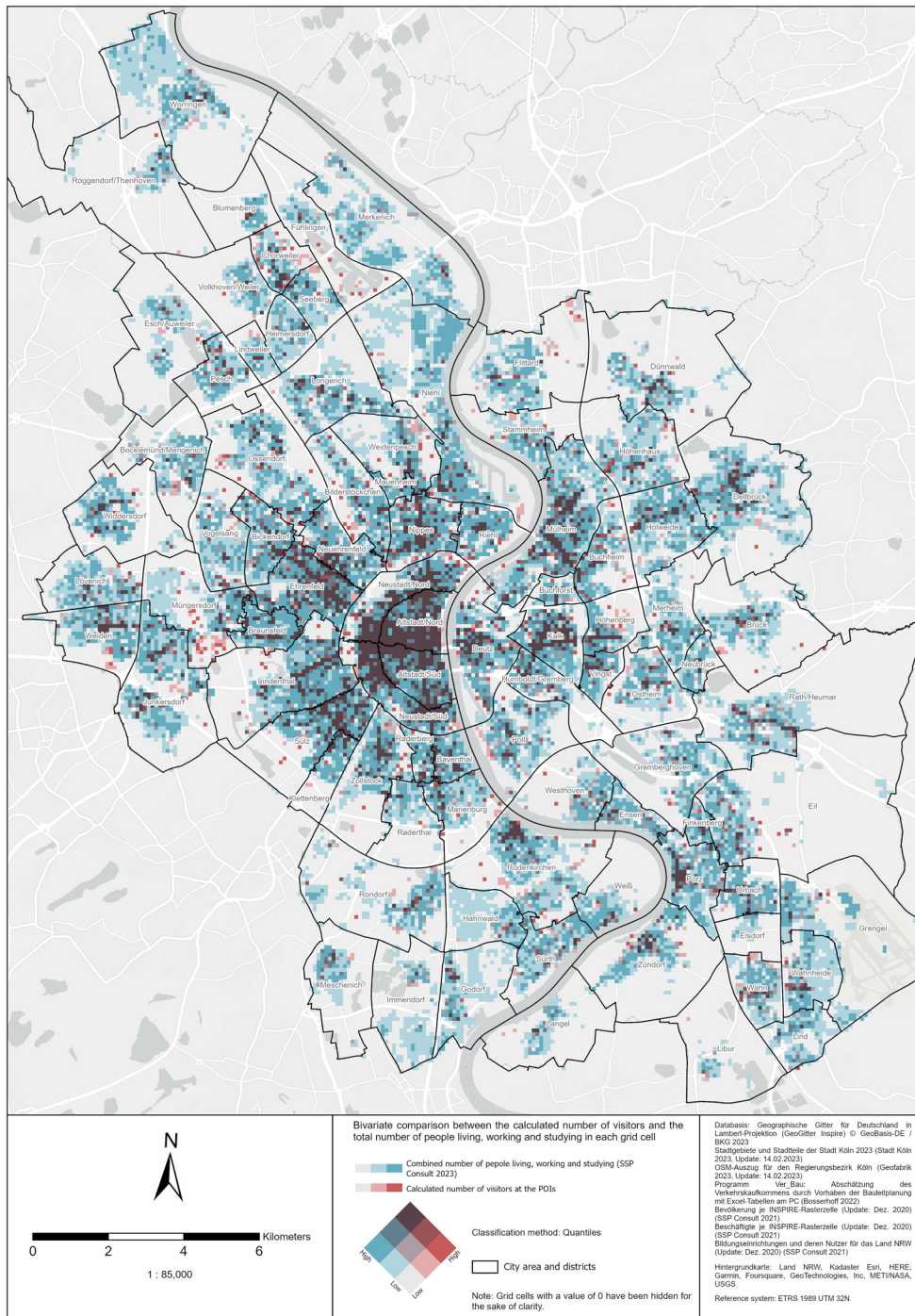


Figure 4 Bivariate comparison between the calculated potential number of daily visitors and the combined number of people living, working and studying in the City of Cologne

5 Discussion

When analysing the areas with different calculated numbers of daily visitors (Figure 3), it was possible to determine potential hotspots with the help of the GIS-based visualization. These hotspots could be attributed to specific streets, which

became visually apparent due to the concentration of multiple grid cells with high visitor volumes. Such clusters can be identified throughout the entire urban area and include high-density as well as less densely built-up areas. With the help of the applied methodology, hotspots can be identified that can be specified in more detail and, if necessary,

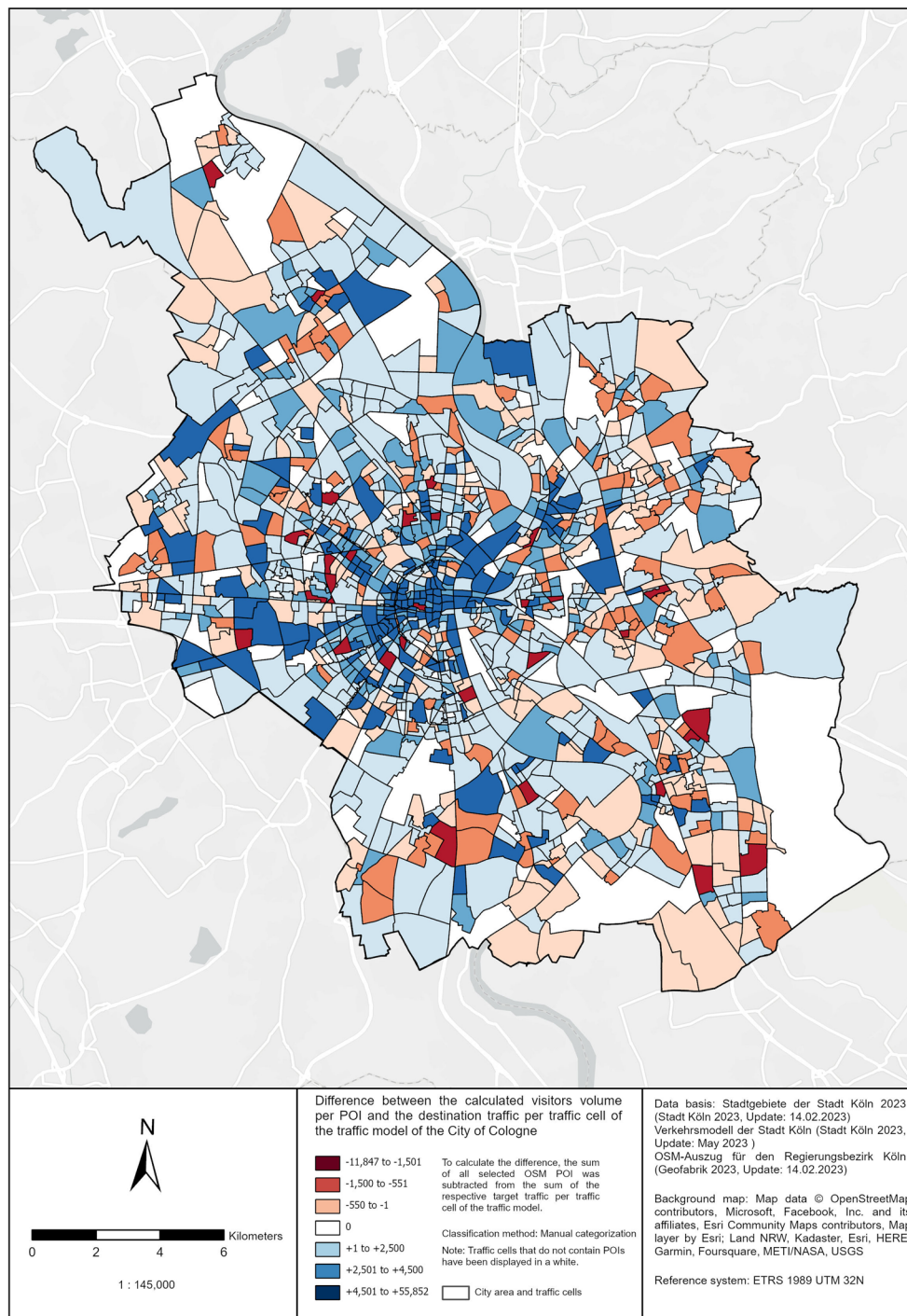


Figure 5 Difference between the calculated volume of visitors to the points of interest and the destination traffic of the traffic model of the City of Cologne

examined individually and in greater depth with regard to their potential number of daily visitors. In accordance with other studies (Chmielewski/Kempa 2020: 103; Kagerbauer 2022: 1), it is thus also feasible to generate cartographic representations that can be analysed in-depth using an es-

tablished scientific methodology. Likewise, addressing the fourth question with the help of the applied methodology, areas for living, working and education, central local supply and nightlife could be distinguished, recognizable by the bivariate representation (Figure 4). The selected grid

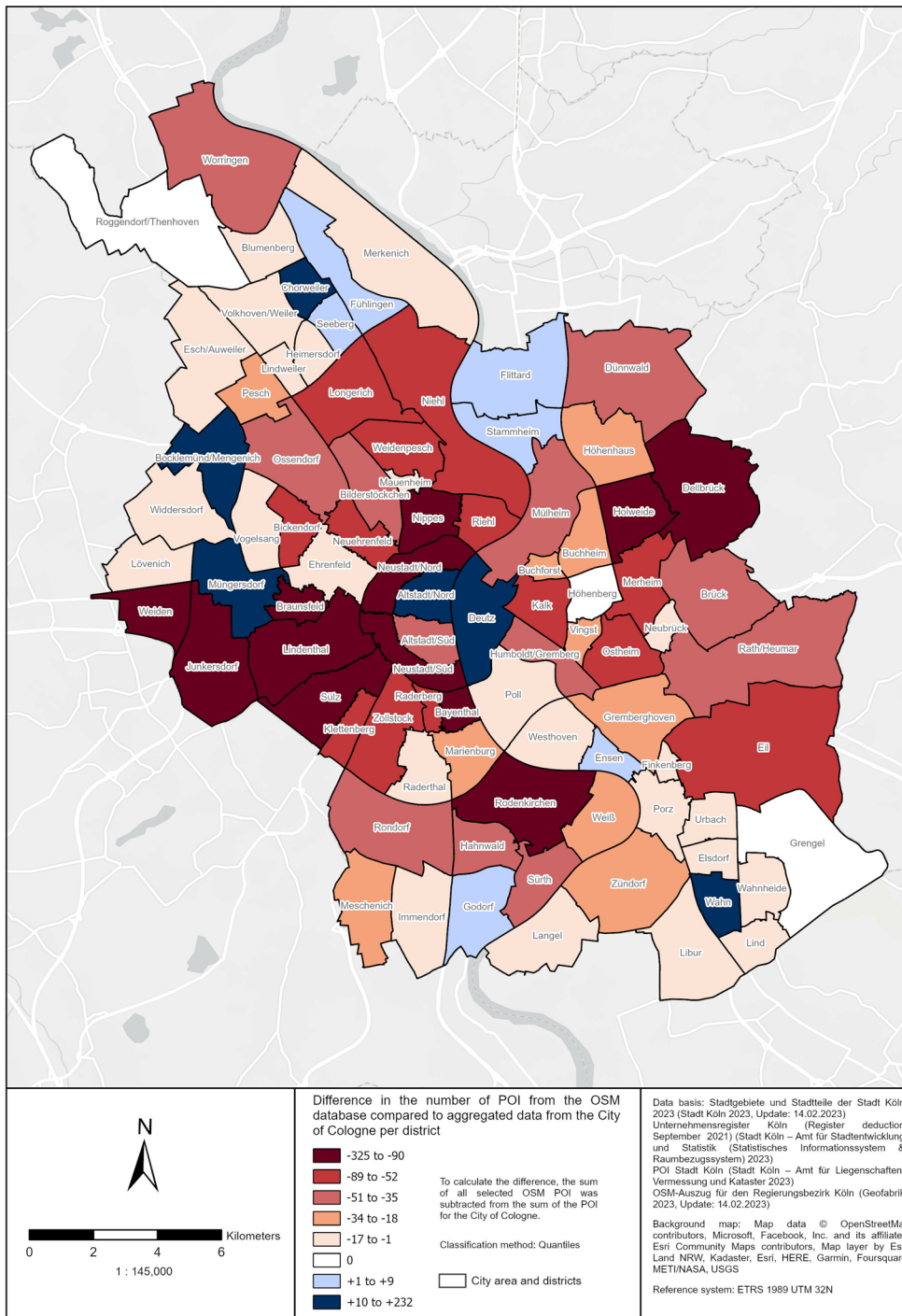


Figure 6 Difference between the number of points of interest objects in the OSM dataset compared to aggregated data for the City of Cologne

cell size of 100 m thus makes it possible to determine differences between individual neighbourhoods within the districts. However, this requires the availability of the supplementary datasets used here. An even more detailed investigation of such areas could be carried out as part of

a more comprehensive study using only OSM data, as in Surahman/Wegner's (2022) work on Uppsala. It can be seen that, in addition to the spatial hexagonal subdivision used by Chmielewski/Kempa (2020: 110), a subdivision using

raster cells also enables the spatial representation and detailed interpretation of traffic-related spatial data.

However, with regard to the numerical values of the calculated daily visitor volumes, both the OSM dataset for the Cologne case study and the applied methodology warrant critical reflection. Clear numerical differences between the traffic model and the number of daily visitors were calculated. In many cases, the number of daily visitors was overestimated. The main reasons for this are probably the quality and quantity of the OSM data, but also the approximation attempts for estimation values which needed to be used because of missing values within Bosserhoff (2022). With regard to the quality of the OSM data, it can be seen that – as described in Yeow/Low/Tan et al. (2021: 22) – there are some deficits regarding points of interest compared to the municipal dataset (Figure 6). Specifically, in many districts the municipal data contain more points of interest in many districts than the OSM database. One reason for this is that the OSM dataset is a VGI dataset that is heavily dependent on the activities and precision of OSM users, which could be insufficient in the City of Cologne. Furthermore, it must be acknowledged that the selected points of interest may be incomplete or inaccurate. This applies in both directions, either due to an overrepresentation of non-comparable points of interest in the dataset provided by the City of Cologne, or due to an underrepresentation of points of interest in the OSM dataset. Additionally, both data sources may contain instances of incorrect categorization of individual points of interest. With regard to the usability of points of interest from open data services for the improved representation of passenger transport, there may therefore be some doubts concerning the suitability of the OSM dataset. In contrast to the work of Klinkhardt/Woerle/Briem et al. (2021: 301), Surahman/Wegner (2022: 61) and Ziemke/Kaddoura/Nagel (2019: 875), it was only possible to calculate approximately similar values for passenger transport compared to the transport model of the City of Cologne (Stadt Köln 2023c) to a limited extent. The aforementioned studies achieved a higher level of accuracy with the help of OSM points of interest relative to their respective reference data. It can be assumed that more extensive adjustments were made to the OSM dataset as part of these previous studies. It might be argued that the calculated results using the OSM points of interest for the City of Cologne are perhaps more precise. In light of the extensive and resource-intensive process involved in creating a traffic model, however, this assumption cannot be supported. Numerous plausibility checks and correction processes are of great relevance in the creation of such a traffic model; in comparison, exclusively calculating the potential number of daily visitors with the help of points of interest can only provide initial approximations, if at all.

Taking into account the points already mentioned, the

methodology used needs to be adapted for more precise results. In particular, the deviations presented may well be related to the approximation and research of supplementary estimates to fill the gaps in the estimation values drawn from Bosserhoff (2022). Gaps concerning the hospitality and medical sectors may have had a major impact on the results, as approximations and estimates had to be collected to fill them. As the number of points of interest used in the catering category in particular was very high, it can be assumed that the approximation resulted in inaccurate values being calculated with regard to the number of daily visitors. This could therefore be one of the main reasons for the deviation between the values calculated and the values of the traffic model. It should also be emphasized that areas of housing, work and education were deliberately excluded, as external data was already available for these. This was also taken into account in this study, in addition to the OSM points of interest, in order to carry out an investigation as similar as possible to the previous scientific discourse. This helps explain the different results achieved here compared to the work by Klinkhardt/Woerle/Briem et al. (2021: 298) and Surahman/Wegner (2022: 61), as, among other aspects, no separate calculation was carried out for residents. Rather, the focus of this study was on the retail, gastronomy and leisure activity sectors, as specific values based on the OSM points of interest were only calculated for these areas.

There may also be errors in the municipal comparison data used, such as points of interest that no longer exist and were removed from the OSM dataset. The business register of the City of Cologne from 2021 had to be used for the points of interest quantity comparison, as no more recent dataset was available. The time difference of two years compared to the points of interest dataset leads, for example, to missing points of interest in highly dynamic areas, such as gastronomy, and to more difficult comparability in the context of newly developed areas and thus new points of interest. In addition, the COVID-19 pandemic led to extensive and rapid changes in points of interest in a relatively short period of time, the dynamics of which would require a separate investigation. Above all, data such as the municipal data used here can only be determined after great effort and at irregular intervals (Stengel/Pomplun 2011: 117). The constant updating of the OSM database by users should be positively acknowledged. In order to be able to make differentiated statements regarding up-to-datedness, extensive comparisons with ground truth data would have to be carried out in future work, as mentioned by Klinkhardt/Kühnel/Heilig et al. (2023). The use of points of interest from the OSM database requires extensive preparation of the data, taking into account the quality and timeliness of the data export from the database used in order to achieve the most valid results possible (see Klinkhardt/Woerle/Briem et al. 2021:

301; Surahman/Wegner 2022: 61; Ziemke/Kaddoura/Nagel 2019: 875). Considering the limitations mentioned, the data is suitable for traffic modelling, but this is accompanied by time-consuming preparation of the open data because potential mapping and updating errors within the OSM data need to be removed, as they can impact the relevant results.

Regarding assignment of the estimation values according to Bosserhoff (2022) and the gaps in estimated values mentioned, the collection of further values for several areas is necessary in order to achieve representative results. Gastronomy, the medical sector (excluding pharmacies and medical supply stores, as values are available for these), hotels, petrol stations and banks should be mentioned explicitly. One suggestion for data collection would be for the respective industry associations to collect data as part of member surveys; this could then be used to derive estimated values in accordance with Bosserhoff (2022). In addition, it would be desirable to use further supplementary datasets in future studies in order to be able to calculate the number of daily visitors as accurately as possible. This may refer, for example, to area-related data, which in the present study were derived from polygons provided by the OSM database but which could be complemented by additional datasets in future research. Finally, with regard to calculation of the number of daily visitors, the automation of this methodology, as mentioned by Klinkhardt/Woerle/Briem et al. (2021: 302), can be regarded as an overriding goal. Automation would involve the OSM data being loaded into an algorithm and the potential number of daily visitors made visible. However, this requires overcoming the current limitations, e.g., the lack of estimated values for certain point-of-interest categories and the lack of data quality and quantity in some cases. Nonetheless, striving for an exact representation seems an unrealistic aim, as traffic models fundamentally remain simplified representations of reality.

Regarding the fifth research question, it is apparent that extensive traffic modelling can be shortened with the help of calculated passenger volumes. Similar methods can be used to determine pedestrian volumes with the calculated value for streets in order to enable potential replanning proposals for road cross-sections or similar. Extensive traffic modelling remains indispensable. However, it can be assumed that for smaller planning processes for changes in road space, an indicative estimate of the number of daily visitors based on OSM points of interest can provide initial helpful indications.

6 Conclusion

In respect to the first research question, points of interest of the OSM open data show some limitations with regard to the mapping of passenger traffic. Although it is possible to calculate the number of daily visitors, this requires very extensive processing of the datasets. This begins with identifying and selecting the relevant points of interest. The actual calculation is an essential expansion point for future studies. The estimation values according to Bosserhoff (2022) can be used for German-speaking countries but should be extensively supplemented in the future, for example to include the areas of gastronomy, medical services and petrol stations. To this end, assumptions were made that may distort the calculated results but that allowed initial approximations regarding potential visitor traffic. For the selected study area, clear hotspots and local supply centres could be identified for which a potential volume of daily visitors could be calculated. Furthermore, it was possible to create cartographic illustrations which can be interpreted and analysed regarding the calculated potential number of visitors. For individuals who are unfamiliar with the local area, such as external stakeholders or consultants, this method provides comprehensive understanding of the spatial arrangement of critical local service areas within the defined study area. From a traffic planning perspective, it can be assumed that the values for the cluster areas can also be used to derive the potential volume of daily visitors in order to initiate planning processes for the potential redesign of street space. This may support the more sustainable design of traffic, such as the expansion of cycling infrastructure or the construction of infrastructure for alternative drive systems, such as e-cars. Taking into account planning guidelines and instructions, such as those of the Road and Transportation Research Association (Forschungsgesellschaft für Straßen- und Verkehrswesen), different implementation options for transport infrastructure planning can be derived with the help of passenger volumes. Furthermore, the methodology could be applied and further developed in areas such as transport supply modelling, particularly in the context of accessibility analysis and land-use–transport interaction models. However, this would require more comprehensive development of the approach to meet the specific requirements of these additional use cases. With regard to the fourth research question, a distinction could be made between specific areas within a selected urban area. Residential, occupational and educational areas could be clearly differentiated from zones designated for retail and leisure activities, as well as locations where these functional areas intersect. Concerning the fifth research question, it is important to emphasize that, despite existing limitations and challenges, the methodology remains applicable for planning processes

and traffic modelling. Potential solutions have also been identified in the discussion. In order to use the methodology presented more effectively, the gaps in the estimates provided by Bosserhoff (2022) must first be closed to ensure comprehensive coverage of all points of interest. This would allow large areas such as the hospitality industry to be better mapped in terms of the number of daily visitors. The quality and up-to-datedness of the open-data points of interest used must then be verified and corrected if necessary. In this case, a list of common errors that need to be corrected would facilitate a more accurate estimate using the methodology presented. The use of points of interest from open data offers – as proven – great potential to estimate passenger volumes for traffic modelling. However, further improvements to the methodology and additional datasets are essential for precise calculations.

Competing Interests The authors declare no competing interests.

Acknowledgements We would like to thank two anonymous reviewers for their helpful comments. We are very grateful for the detailed and very helpful advice as well for the time they took. Furthermore, we would also like to thank the City of Cologne and SSP Consult Beratende Ingenieure GmbH for providing the data. We would also like to thank the City of Cologne for allowing some of the data processing to take place in the city's offices for data protection reasons. We would also like to thank the editors for their meticulous and communicative support with regard to the manuscript.

Funding This work received no external funding.

References

- Arase, K.; Wu, Z.; Migita, T.; Takahashi, N. (2022): Deep Learning of OpenStreetMap Images Labelled Using Road Traffic Accident Data. In: Institute of Electrical and Electronics Engineers (ed.): Proceedings of 2022 IEEE Region 10 International Conference, TENCON 2022. Hong Kong, 1–6. <https://doi.org/10.1109/TENCON55691.2022.9977529>
- Balac, M.; Hörl, S. (2021): Synthetic population for the state of California based on open data: examples of San Francisco Bay area and San Diego County. Zürich. <https://doi.org/10.3929/ethz-b-000481954>
- Bechtel, B.; Hüser, C. (2023): Das Stadtklima: Ursachen, Effekte und Erfassung. In: Geographische Rundschau 75, 7/8, 10–15.
- BMDV – Bundesministerium für Digitales und Verkehr (2023): Nationaler Radverkehrsplan 3.0. Berlin.
- Bosserhoff, D. (2022): Programm Ver_Bau: Abschätzung des Verkehrsaufkommens durch Vorhaben der Bauleitplanung mit Excel-Tabellen am PC. <https://www.dietmarbosserhoff.de/Programm.html> (13.11.2025).
- Briem, L.; Heilig, M.; Klinkhardt, C.; Vortisch, P. (2019): Analyzing OpenStreetMap as data source for travel demand models case study in Karlsruhe. A case study in Karlsruhe. In: Transportation Research Procedia 41, 104–112. <https://doi.org/10.1016/j.trpro.2019.09.021>
- Busch-Geertsema, A.; Klinger, T.; Lanzendorf, M. (2019): Geographien der Mobilität. In: Gebhardt, H.; Glaser, R.; Radtke, U.; Reuber, P.; Vött, A. (eds.): Geographie – Physische Geographie und Humangeographie. Berlin, 1015–1032.
- Cai, P.; Lee, Y.; Luo, Y.; Hsu, D. (2020): SUMMIT: A Simulator for Urban Driving in Massive Mixed Traffic. In: 2020 IEEE International Conference on Robotics and Automation. Paris, 4023–4029. <https://doi.org/10.1109/ICRA40945.2020.9197228>
- Camargo, C.Q.; Bright, J.; Hale, S.A. (2019): Diagnosing the performance of human mobility models at small spatial scales using volunteered geographical information. In: Royal Society Open Science 6, 11, 1–15. <https://doi.org/10.1098/rsos.191034>
- Chmielewski, J.; Kempa, J. (2020): Hexagonal Zones in Transport Demand Models. In: KnE Engineering: International Congress on Engineering — Engineering for Evolution, 103–116. <https://doi.org/10.18502/keg.v5i6.7025>
- Cohen, A.; Dalyot, S. (2020): Machine-learning prediction models for pedestrian traffic flow levels: Towards optimizing walking routes for blind pedestrians. In: Transactions in GIS 24, 5, 1264–1279. <https://doi.org/10.1111/tgis.12674>
- Dangschat, J.S. (2022): Verkehrswende – sozial und räumlich ausgewogen. In: Journal für Mobilität und Verkehr 14, 2–10. <https://doi.org/10.34647/jmv.nr14.id87>
- Ferster, C.; Fischer, J.; Manaugh, K.; Nelson, T.; Winters, M. (2020): Using OpenStreetMap to inventory bicycle infrastructure: A comparison with open data from cities. In: International Journal of Sustainable Transportation 14, 1, 64–73. <https://doi.org/10.1080/15568318.2018.1519746>
- FGSV – Forschungsgesellschaft für Straßen- und Verkehrswesen (2022): Empfehlungen zum Einsatz von Verkehrsnachfragemodellen für den Personenverkehr (EVNM-PV). Köln.
- Göttsche, F.; Brinkmann, J. (2023): Wie passen wir unsere Städte in Zeiten des Klimawandels an Hitze an? In: Geographische Rundschau 75, 7/8, 16–21.
- Hertig, E.; Keck, M. (2023): Deutschlands Städte im Klimawandel. In: Geographische Rundschau 75, 7/8, 4–9.
- Kagerbauer, M. (2022): Integration von neuen Mobilitätsformen in Verkehrserhebungen und Verkehrsmodellierung. Karlsruhe. = Schriftenreihe des Instituts für Verkehrswesen 77. <https://doi.org/10.5445/KSP/1000144791>
- Keler, A.; Grigoropoulos, G.; Mussack, D. (2019): Enrich-

- ing complex road intersections from OSM with traffic-related behavioral information. In: 29th International Cartographic Conference 2, 61. <https://doi.org/10.5194/ica-proc-2-61-2019>
- Klinkhardt, C.; Woerle, T.; Briem, L.; Heilig, M.; Kagerbauer, M.; Vortisch, P. (2021): Using OpenStreetMap as a Data Source for Attractiveness in Travel Demand Models. In: *Transportations Research Record: Journal of the Transportation Research Board* 2675, 8, 294–303. <https://doi.org/10.1177/0361198121997415>
- Klinkhardt, C.; Kühnel, F.; Heilig, M.; Lautenbach, S.; Wörle, T.; Vortisch, P.; Kuhnimhof, T. (2023): Quality Assessment of OpenStreetMap's Points of Interest with Large-Scale Real Data. In: *Transportation Research Record: Journal of the Transportation Research Board* 2677, 12, 661–674. <https://doi.org/10.1177/03611981231169280>
- Köhler, U. (2014): Einführung in die Verkehrsplanung. Stuttgart.
- de Lange, N. (2020): Geoinformatik in Theorie und Praxis. Grundlagen von Geoinformationssystemen, Fernerkundung und digitaler Bildverarbeitung. Berlin. <https://doi.org/10.1007/978-3-662-60709-1>
- Liu, X.; Long, Y. (2016): Automated identification and characterization of parcels with OpenStreetMap and points of interest. In: *Environment and Planning B: Urban Analytics and City Science* 43, 2, 341–360. <https://doi.org/10.1177/0265813515604767>
- Loo, B.; Tsoi, K.H. (2018): The sustainable transport pathway: A holistic strategy of Five Transformations. In: *Journal of Transport and Land Use* 11, 1, 961–980. <https://doi.org/10.5198/jtlu.2018.1354>
- Mahajan, V.; Kühnel, N.; Intzevidou, A.; Cantelmo, G.; Moeckel, R.; Antoniou, C. (2022): Data to the people: a review of public and proprietary data for transport models. In: *Transport Reviews* 42, 4, 415–440. <https://doi.org/10.1080/01441647.2021.1977414>
- Martinelli, L. (2018): Can we validate every change on OSM? https://2018.stateofthemap.org/2018/T079-Can_we_validate_every_change_on_OSM/ (13.11.2025).
- Peter, M. (2021): Die Berechnung kleinräumiger und multimodaler Erreichbarkeiten auf regionaler Ebene. Hamburg. = Harburger Berichte zur Verkehrsplanung und Logistik 22. <https://doi.org/10.15480/882.3673>
- Rau, H.; Scheiner, J. (2020): Sustainable Mobility: Interdisciplinary Approaches. In: *Sustainability* 12, 23, 9995. <https://doi.org/10.3390/su12239995>
- Reynard, D. (2018): Five classes of geospatial data and the barriers to using them. In: *Geography Compass* 12, 4, e12364. <https://doi.org/10.1111/gec3.12364>
- Sallard, A.; Balac, M.; Hörl, S. (2020): A synthetic population for the greater São Paulo metropolitan region. Zürich. = Arbeitsberichte Verkehrs- und Raumplanung 1545. <https://doi.org/10.3929/ethz-b-000429951>
- SSP Consult – SSP Consult Beratende Ingenieure GmbH (2021a): Bevölkerung je INSPIRE-Rasterzelle. (Dezember 2020).
- SSP Consult – SSP Consult Beratende Ingenieure GmbH (2021b): Beschäftigte je INSPIRE-Rasterzelle. (Dezember 2020).
- SSP Consult – SSP Consult Beratende Ingenieure GmbH (2021c): Bildungseinrichtungen und deren Nutzer für das Land NRW. (Dezember 2020).
- Stadt Köln (2020): Kölner Perspektiven 2030+. Köln.
- Stadt Köln (2021): Unternehmensregister Köln (Registerabzug September 2021). Köln.
- Stadt Köln (2022): Die Kommunale Gebietsgliederung. Ein räumlicher Bezug für statistische Daten. Köln.
- Stadt Köln (2023a): Kölner Stadtteilinformationen. Bevölkerungszahlen 2022. Köln. = Kölner Statistische Nachrichten 5/2023.
- Stadt Köln (2023b): Netzentwicklung Mobilität – Attraktive Verkehrsnetze für Köln. Köln.
- Stadt Köln (2023c): Verkehrsmodell der Stadt Köln. Stand: Mai 2023. Köln.
- Stadt Köln (2023d): Points of Interest der Stadt Köln. Stand: Februar 2023. Köln.
- Steiniger, S.; Poorazizi, M.E.; Scott, D.R.; Fuentes, C.; Crespo, R. (2016): Can we use OpenStreetMap POIs for the Evaluation of Urban Accessibility? In: *International Conference on GIScience Short Paper Proceedings* 1, 1, 272–275. <https://doi.org/10.21433/B31167f0678p>
- Stengel, S.; Pumplun, S. (2011): Die freie Weltkarte OpenStreetMap – Potenziale und Risiken. In: *Kartographische Nachrichten – Journal of Cartography and Geographic Information* 61, 3, 115–120. <https://doi.org/10.1007/BF03544072>
- Surahman, I.; Wegner, G. (2022): Integration of Open Data in Disaggregate Transport Modelling. A Case Study of Uppsala. Stockholm.
- Topp, H. (2023): Von der autogerechten Stadt zur menschengerechten Stadt. In: *Straßenverkehrstechnik* 67, 1, 31–37.
- Treiber, M.; Kesting, A. (2010): Verkehrsdynamik und -simulation. Daten, Modelle und Anwendungen der Verkehrsflussdynamik. Heidelberg.
- UN-Habitat – United Nations Human Settlements Programme (2022): World Cities Report 2022. Envisaging the Future of Cities. Nairobi.
- Vierø, A.R.; Vybornova, A.; Szell, M. (2024): BikeDNA: A tool for bicycle infrastructure data and network assessment. In: *Environment and Planning B: Urban Analytics and City Science* 51, 2, 512–528. <https://doi.org/10.1177/23998083231184471>
- Xu, F.F.; Lin, B.Y.; Lu, Q.; Huang, Y.; Zhu, K.Q. (2016):

- Cross-region Traffic Prediction for China on OpenStreetMap. In: Winter, S. (ed.): IWCTS '16: Proceedings of the 9th ACM SIGSPATIAL International Workshop on Computational Transportation Science. New York, 37–42. <https://doi.org/10.1145/3003965.3003972>
- Yan, Y.; Feng, C.-C.; Huang, W.; Fan, H.; Wang, Y.-C.; Zipf, A. (2020): Volunteered geographic information research in the first decade: a narrative review of selected journal articles in GIScience. In: International Journal of Geographical Information Science 34, 9, 1765–1791. <https://doi.org/10.1080/13658816.2020.1730848>
- Yeow, L.W.; Low, R.; Tan, Y.X.; Cheah, L. (2021): Points-of-Interest (POI) Data Validation Methods: An Urban Case Study. In: International Journal of Geo-Information 10, 11, 735. <https://doi.org/10.3390/ijgi10110735>
- Zhang, L.; Pfoser, D. (2019): Using OpenStreetMap point-of-interest data to model urban change – A feasibility study. In: PLoS One 14, 2, e0212606. <https://doi.org/10.1371/journal.pone.0212606>
- Ziemke, T.; Braun, S. (2021): Automated generation of traffic signals and lanes for MATSim based on OpenStreetMap. In: Procedia Computer Science 184, 745–752. <https://doi.org/10.1016/j.procs.2021.03.093>
- Ziemke, D.; Kaddoura, I.; Nagel, K. (2019): The MATSim Open Berlin Scenario: A multimodal agent-based transport simulation scenario based on synthetic demand modeling and open data. In: Procedia Computer Science 151, 870–877. <https://doi.org/10.1016/j.procs.2019.04.120>
- Zilske, M.; Neumann, A.; Nagel, K. (2011): OpenStreetMap for traffic simulation. In: Proceedings of the 1st European state of the map: OpenStreetMap conference. Wien, 126–134.