Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 41 (2019) 104–112

Transportation Research Procedia

www.elsevier.com/locate/procedia

mobil.TUM 2018 "Urban Mobility - Shaping the Future Together" - International Scientific Conference on Mobility and Transport

Analyzing OpenStreetMap as data source for travel demand models A case study in Karlsruhe

Lars Briem[a,*], Michael Heilig[a], Christian Klinkhardt[a], Peter Vortisch[a]

[a] *Institute for Transport Studies, Karlsruhe Institute of Technology, Kaiserstrae 12, 76131 Karlsruhe, Germany*

Abstract

Microscopic destination choice models allow for a choice of individual destinations instead of travel analysis zones (TAZ). Consequently, for such a case, information about each individual destination on a more detailed spatial level is needed. This information can be gathered manually or, if available, through the classical approach from public authorities. However, these approaches are often time consuming and expensive. An alternative way is to gather the information automatically from open data portals. OpenStreetMap (OSM) is such an open data portal driven by volunteers. The data is publicly available, but as volunteers provide it, the quality of the data is not controlled by a single instance. Therefore, we analyze the potential and risks of using OSM as a data source for microscopic destination choice models.

Our analysis took place in the city of Karlsruhe. We analyzed the available information for destinations regarding the education and work activity type. Our analysis is split into two parts. Firstly, we analyzed the trend of the information available on OSM over the last years. Secondly, we compared the currently available information for one area of Karlsruhe. By comparing OSM data with manually gathered data, we show that for the education activity type, OSM data is suitable to determine location attributes, whereas, for the work activity type, OSM data is still lacking and therefore not suitable to determine location attributes.

Keywords: Open street map; destination choice model; travel demand model;

1. Introduction

Travel demand models are essential tools to evaluate impacts of transport policy measures, innovations and new developments. In the past, mainly macroscopic models have been used for this purpose. Structural data was primarily collected by the use of classical data sources (e.g., statistic departments of local authorities, commercial data). With the increasing power of modern computers, it became possible to use microscopic models. Besides modeling the population on an individual level, microscopic models also offer the possibility of using accurate locations for destination choice instead of travel demand zones. This implies that the provided structural data (e.g., networks, workplaces,

* Corresponding author. Tel.: +49-721-608-47772 ; fax: +49-721-608-46777.
E-mail address: lars.briem@kit.edu

10.1016/j.trpro.2019.09.021

shops, schools) needs to contain much more detail. Collecting these data by hand can be time consuming and expensive. The upcoming movement towards open data portals and publicly available data provides a promising approach to new data sources, e.g., OpenStreetMap or GovData. An automatic or semi-automatic process collecting the required data from one or more than one of such sources would speed up the creation of new travel demand models a lot. Besides the high potential, open data sources can provide, there are also drawbacks compared to classical data sources. OSM data usually does not provide information typically used for destination choice models, like the number of employees for offices or shops, the number of students for schools or the size of shops.

In the past few years, several studies investigated the quality of OSM data in general, regarding road networks and points of interests (POI). Haklay (2010) was one of the first comparing OSM data with an ordnance survey (OS) dataset. He found that in 2010, the quality for OSM data in Great Britain was lacking in some points, but increased rapidly from one year to another.

Girres and Touya (2010) extend the work of Haklay and state that on the one hand, OSM is very good regarding responsiveness and flexibility. On the other hand, they highlight that the heterogeneity of OSM data profoundly limits the possible application.

Arsanjani et al. (2015) analyzed the process of collaborative contribution to the OSM project spatially and temporally for the years 2007 to 2012. They found that once the area was basically mapped seven years after the start of OSM in 2004, a densification process started. Land-use types, such as transport units, leisure facilities, urban fabrics, commercial units, and forest areas, are of high interest for mappers in OSM. By the use of a cellular automata model, they tried to forecast the further development and stated that by 2014, contributions on all of the urban fabrics would be received.

Mashhadi et al. (2012) analyzed the quality of POIs in OSM by comparing OSM data and commercial data from Navteq and Yelp for London (UK) and Rome (Italy). They found that for these urban areas, the accuracy concerning the geographic position of POIs existing in both datasets is very high. Zielstra and Zipf (2010) compared the road networks of OSM and the commercial TomTom Multinet dataset. They found that in densely populated urban areas, OSM data can be an alternative to commercial data sources. However, in rural areas, the quality of the OSM road network was lacking.

Zilske et al. (2011) finally show that the geographical data contained in OpenStreetMap (OSM), such as streets, is suitable to setup agent-based travel simulations. It provides not only information about the size and location of those elements but also contains details about the shops, offices or companies in one of those elements. These details make the use of OSM data for destination choice models using individual destinations very interesting, as every place or building has exact coordinates and can be used as a possible destination. However, data quality was not part of this work. Gil (2015) also showed that a multimodal urban network model could be built based on OSM data, but lacking data quality is still an issue that needs to be solved.

Most studies investigating the quality of OSM data were conducted in 2015 or earlier. However, data quality might have improved over time, as also stated by Arsanjani et al. (2015). This paper aims at evaluating the data quality as well as the potential and risks of OSM as a single open data source for destination choice models. Therefore, the data provided by OSM is initially analyzed regarding completeness and validity. The provided information of OSM is then analyzed, whether or not they are beneficial to approximate the attributes used in destination choice models. As a starting point, the activity types *work* and *education* are investigated in the city of Karlsruhe.

This paper is structured as follows. First, we give a short overview of destination choice on a microscopic level and OSM data. We then describe the methodology followed by the presentation of results. We complete our paper with a conclusion and an outlook of future work.

2. Destination choice on a microscopic level

Typically, macroscopic travel demand models aggregate people in homogeneous person groups in so-called travel analysis zones (TAZ) of 500 to 2000 inhabitants. However, depending on the nature of the model the zone size can vary widely (Ortúzar and Willumsen, 2011). Destination choice in such cases is made on the level of the TAZs. The TAZs are selected as destinations based on different attributes of each TAZ (examples for attributes can be found in Horni (2013), p.35) as well as on trip attributes (e.g., distance, travel time). With the introduction of microscopic models for simulating travel demand, it became (1) possible to model the missing details of demand data by modeling every person as an agent, (2) possible to consider person attributes of the single agents and (3) possible to also determine an exact location as destination for the single agent. These advantages allow for more accurate models and though, more accurate predictions for policy measures. Modeling individual destinations on an aggregate level is complex because the available data is usually missing many details.

Consequently, using individual destinations in microscopic travel demand models has several drawbacks. First, the computational effort is much higher (Horni, 2013). Second, location data needs also to be available on a disaggregated level. In addition to the locations of potential destinations, it is also necessary to have information about the location attributes (e.g., size of a shop or number of students at a school). Although the individual locations may be available nowadays on online map services like Google maps or OSM, the additional information about the location attributes are often missing.

On one side, data collected by volunteers as on OSM is quite heterogeneous and thus may not provide the accuracy and completeness like the one from classical data sources. On the other side, there is also the potential for enthusiastic volunteers to provide even more data than official sites. However, data about individual locations gained from online sources tend to be incomplete and therefore have to be reviewed before using them for travel demand modeling. In case of missing information, synthetic methods are essential to generate the missing information (Nagel and Axhausen, 2001). A possible synthetic way to generate missing location attributes could be to approximate the number of employees or students based on the size of a building, including all floors of the building, as described in UNESCO Institute for Statistics (UIS) (2012) for students and more general Vogt (2006) for Germany.

3. Data in OSM

OSM is built upon a few fundamental elements, namely *node*, *way*, *relation* and *tag* (OpenStreetMap Wiki, 2018a; Ramm and Topf, 2010). A node is a point on the earth's surface specified by its latitude, longitude, and ID. The ID is used by other elements to refer to that node. A way is a polyline defined by at least two referenced nodes. Ways are also used to define areas, such as buildings or forests. Ways are limited to 2000 referenced nodes. For elements that are more complex, relations are used. A relation describes a relationship between multiple elements, e.g., a multipolygon can be used to model a building with holes using an outer way to model the boundary of the building and inner ways to model the holes. Ways and relations also contain IDs for reference purposes.

As nodes, ways, and relations describe the spatial distribution of the elements, *tags* contain all non-spatial information. A tag is a combination of a *key* and a *value* which are both arbitrary strings. It can be attached to nodes, ways, and relations. Assigning tags to nodes, ways or relations is optional. Therefore, it cannot be guaranteed, that for example a building that in reality contains a shop has the appropriate tags attached. This leads to inaccuracies in the data which are analyzed in this paper.

The OSM wiki describes tag lists for often-used elements, called map features. Contributors should use them when reporting new data to OSM, see Ramm and Topf (2010). The editors of OSM maintain the list as an informal standard. New tags are added as map features after a vote for or against them.

When adding new elements to OSM, the contributor decides to which element a tag is attached. E.g., the contributor wants to add a new school. In this case, the school can be reported as a way or relation, given that the volunteer is aware of the building layout. Otherwise, the school can be reported as a node placed in the center of the school. This results in heterogeneous data, where potential destinations are reported as different OSM elements with more or less levels of detail. Depending on the level of detail, an element from OSM can be more or less useful for a microscopic destination choice model. Therefore, the analysis considers this aspect.

4. Methodology

To examine the potential of OSM, we analyze the activity types most relevant for travel demand model - work and education. The data extraction is done using osmosis and JOSM (OpenStreetMap Wiki, 2018b; JOSM, 2018).

4.1. Data extraction

Osmosis is an open source command line application to process OSM data. It is built in a modular way to combine basic building blocks to larger processing chains. Osmosis has powerful filter capabilities, which are used in our case to find the required elements inside a geographical region. Tag lists specify the required elements. Thus, for each activity type, we define a list of tags. The tags used for the education activity type are listed in Table 1 and for work activity type in Table 2.

Key	Value
amenity	school, university
building	school, university
landuse	university, school

Table 1: Used tags for education activity type.

Key	Value
aeroway	aerodrome, spaceport, terminal
amenity	bar, biergarten, cafe, fast_food, food_court, ice_cream, pub, restaurant, library, archive, college, kindergarten, school, music_school, driving_school, language_school, university, bicycle_rental, boat_rental, car_rental, bank, bureau_de_change, baby_hatch, clinic, dentist, doctors, hospital, nursing_home, pharmacy, social_facility, veterinary, blood_donation, arts_centre, brothel, casino, cinema, community_centre, gambling, nightclub, planetarium, social_centre, stripclub, studio, swingerclub, theatre, courthouse, animal_boarding, animal_shelter, coworking_space, crematorium, dive_centre, embassy, fire_station, gym, internet_cafe, marketplace, place_of_worship, police, post_office, prison, public_building, ranger_station, recycling_station, townhall, hotel, farm, commercial, office, industrial, retail, warehouse, kiosk
building	religious, cathedral, chapel, church, mosque, temple, synagogue, shrine, bakehouse, kindergarten, civic, hospital, school, stadium, train_station, transportation, university, public, conservatory, construction, hanger, industrial, commercial, retail, mall
craft	*
emergency	ambulance_station, lifeguard_base
landuse	cementary, commercial, construction, depot, industrial, landfill, military, port, quarry, religious, retail
leisure	adult_gaming_centre, amusement_arcade, beach_resort, dance, fitness_centre, swimming_pool, water_park, swimming_pool
man_made	wastewater_plant, works
military	barracks, naval_base, office
office	*
power	plant
shop	*
tourism	apartment, aquarium, camp_site, caravan_site, chalet, gallery, guest_house, hostel, hotel, motel, museum, theme_park, zoo

Table 2: Used tags for work activity type.

As mentioned in Section 3, the tags typically used are an informal standard. Thus, both tag lists are based on these map features. This allows transferring the lists to other cities without checking the tags that are only locally available in a town.

After extracting the elements of interest, we further process them using JOSM editor. The JOSM editor is used to maintain OSM data but can be used to calculate statistics like number and the size of the area of the elements.

4.2. Analyzed criteria

The analysis is divided into two parts. First, only information available on OSM is used to analyze different years, called historical analysis. Second, a small part of the information on OSM is compared to the real world.

In the historical analysis, it is evaluated how reliable the data is. Therefore, two criteria are used. First, the stability over the last years is analyzed. This is done by examining the number of nodes, ways, and relations over the previous years. The ratio between the years is used as the criterion. Meaning that a small change between two years is treated as stable and a considerable change is treated as unstable. Second, the informational content of the years is analyzed. Therefore, the ratio between elements containing area information and those which do not is analyzed. The more elements include area information, the higher the informational content is.

Additionally, this is set in relation to the aggregated area of all elements. The area is calculated based on the extracted ways and relations using the measurement plugin of JOSM. Nodes are not assigned to ways or relations and thus do not increase the area.

The comparison of OSM with the real world uses the extracted elements and filters the available tags to a subset. As each element in OSM can have an unlimited number of tags, comparing all tags with the real world will be too time-consuming. We decided to use only those tags that are more relevant for travel demand models or are necessary to identify a destination. The tags we considered are: `addr:city`, `addr:country`, `addr:housenumber`, `addr:postcode`, `addr:street`, `amenity`, `isced:level`, `name`, `office`, `opening_hours`, `post_office`, `service_times`, `shop`, and `tourism`.

The comparison is made in both directions. All elements available on OSM are searched in the real world. For existing elements, the tags are compared. Elements available on OSM, but missing in the real world are treated as incorrect. All elements which are available in the real world, but are missing in OSM, are treated as missing. The two activity types work and education are analyzed differently. For the education activity type, the elements from OSM are compared to a list of all state schools given by the education authority. For the work activity type, one student headed out to a small part of Karlsruhe to manually collect the data needed for the comparison.

5. Analysis of the city of Karlsruhe

We examined the potential for OSM being used as data source for destination choice in the city of Karlsruhe. The area of investigation ranges from 8.3056641 in the west to 8.5185242 in the east and from 49.0545202 in the north to 48.9617365 in the south. For the various types of activities, the stability and informational content are analyzed. Selected analyses are processed over the last years.

5.1. Historical analysis

For the historical analyses, we use the data provided by geofabrik.de for the years 2014 until 2018. The OSM file of January 1th is used from each year. Table 3 and Table 4 show that the number of elements available on OSM increases from year to year. This is plausible when assuming, that elements are only removed due to closing, for example, a shop or by removing incorrect data. In contrast, new elements are added to OSM by opening shops or adding missing data. We assume the last one to be mainly responsible for the continuously increasing number of elements.

5.1.1. Education activity type

In case of the education activity type, as shown in Table 3, the number of nodes decreases, while the number of ways and relations increases. As ways and relations are used to model areas, this could be interpreted as an increase in the level of detail. One reason for this is that schools have been modeled as nodes in the past and are now modeled as ways or relations.

Taking a look at the increase of the area size compared to the rise in the number of ways supports this theory. The area increases by about 5% from 2014 to 2018 while the number of ways increases by about 17%. We assume

	2014	2015	2016	2017	2018
Relations	5	14	17	19	24
Ways	367	406	402	430	431
Nodes	36	31	30	26	24
Sum	408	451	449	475	479
Ways per Node	10.194	13.097	13.4	16.538	17.958
Area in km^2	3.97	3.997	4.09	4.156	4.162
Factor to previous year		1.007	1.023	1.016	1.001

Table 3: Analyzed elements for the education activity type.

that the large schools have already been modeled as ways in the past, but the smaller ones are now transformed from nodes to ways, and also schools are modeled in greater detail as the number of relations increases. This supports the assumption of a subsequent densification process after having completed the basic mapping, as stated by Arsanjani et al. (2015).

5.1.2. *Work activity type*

For the work activity type, as shown in Table 4, we see an increase in the number of nodes, ways, and relations during the analyzed period. The number of nodes and ways increases by 80% while the number of relations increases by 275%. We assume that this vast increase stems from many missing locations in the past. When propagating this into the future, we also assume that nowadays there are still many locations missing.

	2014	2015	2016	2017	2018
Relations	20	44	60	59	75
Ways	1550	1927	2404	2697	2816
Nodes	1996	2185	2694	3164	3588
Sum	3566	4156	5158	5920	6479
Ways per Node	0.777	0.882	0.892	0.852	0.785
Area in km^2	17.775	18.513	19.763	20.433	21.289
Factor to previous year		1.042	1.068	1.034	1.042

Table 4: Analyzed elements for the work activity type.

Analyzing the area size of those locations, it increases much slower than the number of ways. The area increases by 20% from 2014 to 2018, being much lower than 80% for the number of ways. As for the education activity type, we also assume here, that the larger work locations have been already modeled as ways in the past and now, the smaller ones are added. Because the number of nodes increases between the years, it is uncertain whether nodes are transformed into ways, or new locations are added directly as ways. Therefore, there is no evidence if the existing buildings are modeled more in detail or only new ones are added. Taking a detailed look at each node is out of the scope of this paper. However, we assume, that the level of detail increases with the number of added nodes and ways.

5.1.3. *Comparing education and work activity type*

The long-term goal of this analysis is to estimate the quality of OSM data to make it possible to calculate reliable location attributes for unique destinations. As described in Section 2, one way to calculate location attributes is the size of a location. As more locations are available with their size, the model could be more reliable. One indicator for this is the difference between the number of ways relative to the number of nodes.

For the education activity type in 2014, the number of ways is ten times the number of nodes. This factor increases more or less constant to a factor of 18 in 2018, indicating a shift from destinations without information about their size to destinations with information about their size or at least to destinations with a more accurate size.

The factor for the number of ways compared to the number of nodes for the work activity type is 0.78 and is more or less the same for 2014 and 2018, but is higher in between. On one side, this indicates that it is more complex to estimate the location attributes for destinations, as the size is currently not available for all destinations. On the other side, this could also indicate, that for the activity type work, the number of destinations in the same building is much higher than for education. In this case, one destination is represented by a way, while all other destinations in the same building are represented only by nodes.

Another measure for the reliability of the model is annual changes in the provided data. For the education activity type, only small changes regarding the area size have been made since 2014. Therefore, we conclude that these locations are already mapped in final quality, and hence, it is possible to build a destination choice model for education activity type on these data.

For the work activity type, it does not look as promising as for education. There is still a rapidly increasing number of elements, and also the area size used by these elements has grown more since 2014. Therefore, we conclude that mapping these locations is still ongoing and therefore, OSM data is currently not reliable enough to build a destination choice model for the work activity type using only OSM data.

5.2. *Comparison with real world*

After analyzing the raw data over the last years, a snapshot of the current OSM data is compared with the real world. We compare the available and relevant tags for each activity type.

5.2.1. *Education activity type*

The analysis for the education activity type uses tags specified in Section 4 to filter all schools. The schools mentioned in OSM are compared with a list of all schools given by local authorities. Nearly all secondary schools of Karlsruhe can be found using this filter. One example for a missing school is the *Nebenius Realschule*. It is located in a building of another school. This is the main issue when extracting schools from OSM because there are often combinations of several schools in one building or area. Typically, the building is assigned to one of the schools and the others are only modeled as nodes.

Estimating the location attributes of a school can be done by a combination of the size of a school and the information provided by Vogt (2006). Vogt (2006) differentiates education in three levels - primary, secondary and tertiary schools - and provides an estimate of students per m^2 of a building's area size. Table 5 shows the levels of schools and the possible tags to classify OSM elements for one of the levels.

Level	Students per $100m^2$	Key `amenity` and `building`	Key `isced:level`	Key `school`
Primary	9 - 12	`school`	1	`primary`
Secondary	8 - 10	`school`	2, 3, 4	`secondary`
Tertiary	4,5 - 9	`university`	5, 6, 7, 8	-

Table 5: Assignment of OSM tags to education levels and students per m^2 specified by Vogt (2006).

Calculating the location attributes for schools this way for multiple schools in one building would result in too many students for one of the schools and too few for the others. In travel demand models, this is only a small issue, because the students head to the same location. Whether the students choose school *a* or *b* in the same building or at the same address is not relevant, as long as the sum of the students for all schools at one address is correct.

The tags provided by OSM can be used to differentiate the schools into the right level. This works well for university buildings, as the keys `amenity` and `building` can be used with the value `university` to filter OSM elements. Due to this, all university elements extracted from OSM contain the appropriate tag and can be used as tertiary education destinations. The differentiation between primary and secondary schools does not work equally well, because the keys `isced:level` and `school` are only available at a small number of schools.

5.2.2. *Work activity type*

The analysis for the work activity type is more complicated as the list of used tags is larger, see Table 2. Therefore, it is also more difficult to estimate the number of employees based on the size of a building or place. There are vast differences between shops, big companies or offices (open or closed) to the public. Estimations might improve when combining the OSM data with data from Eurostat (2008).

To reduce the complexity, we first analyzed the elements that are definite work locations. Therefore, the elements with the keys `amenity`, `shop`, `office`, and `tourism` are extracted from OSM. Due to the reduced complexity, not all possible work locations are analyzed further on.

Estimating the location attributes for work locations could also be done by a combination of the area and statistical data from Eurostat (2008) and Vogt (2006). A critical aspect concerning the area size of a shop or office is the number of floors of a building. If it is given and higher than one, it further increases the area. The number of floors is available for about 11% of all ways and relations. For all other elements, it is assumed that the there is only one floor.

Another interesting aspect for agent-based and microscopic travel demand models are the opening hours of a destination. As this is not relevant for macroscopic models, modeling the time microscopically allows limiting the destinations based on their opening hours. The opening hours can be found for about 33% of all elements. The opening hours are available in various formats, but extracting all opening hours should allow building a parser to automate the extraction.

Further on, we took this data and headed out to compare it for one area of Karlsruhe. Figure 1 shows the area of investigation. In this area, 48 work locations are available on OSM. 44 of those still exist, while two have been renamed and two have been closed. Another 44 work locations, which were identified in the manual survey, could not be found on OSM. This results in about 50% missing work locations, which we assume is too much to be able to estimate a reliable destination choice model. However, there is a difference between work locations open to the public, like shops, and those that are not open to the public. The open ones are mostly all available, while the closed ones are mainly missing. We compared the tags of all the elements available on OSM as specified in Section 4. A more in-depth look at the opening hours revealed that eleven of the 48 elements extracted from OSM contain opening hours. Five of those opening hours are correct. The rest has only a minor deviation from the correct values.

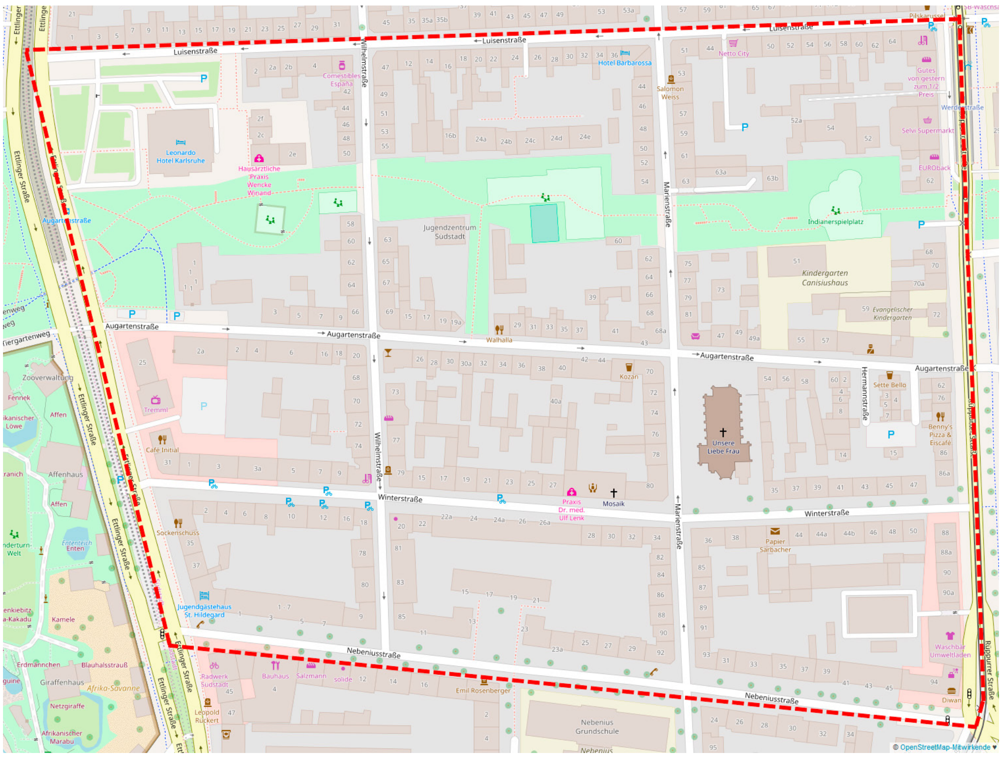

Fig. 1: Area of investigation in the south of Karlsruhe.

6. Conclusion and Future Work

Whether the information provided by OSM is usable for a microscopic destination choice model or not depends on the activity type. For education, the provided data (especially the size of the buildings or lots) are suited for an approximation. There are still some minor issues with multiple schools in the same building, but an algorithm could solve this. The stability of the data concerning schools is also well. The data has changed only a little over the last years. The informational content is well, as there are many elements containing area information. Therefore, we assume that elements for the education activity type are reliable and valid.

For determining work locations, the suitability of OSM is lacking. There are still too many work locations missing. Especially work locations not open to customers are missing. The data about work locations is also more variable than the data for education, as the number of elements has nearly doubled during the investigation period. We assume that this data will improve over the years, but can currently not be the single source for a microscopic destination choice model. The informational content is lower compared to the education activity type. This could be a result of the usage of buildings for different purposes. E.g., multiple shops are often combined in a mall. In case the mall is modeled as a single building, and the shops are modeled as nodes, the informational content is lower compared to each shop being modeled separately. Taking the whole size of the mall into account could lead to a good approximation. We recommend checking the plausibility of automatically extracted work locations from OSM faithfully.

Overcoming these deficiencies, OSM data promises to make a significant contribution to the data collecting process depending on the activity type. Relevant data with very good spatial detail can be extracted automatically and always up to date. However, the data still has to be checked for reasonableness to ensure a small gap in structural data between the model and reality. Therefore, our further research heads towards an automatic and reliable process to gather the data from OSM. This includes assigning node elements to the next matching area but also disaggregating bigger areas to single destinations by using, for example, land use information.

References

Arsanjani, J.J., Helbich, M., Bakillah, M., Loos, L., 2015. The emergence and evolution of openstreetmap: a cellular automata approach. International Journal of Digital Earth 8, 76–90. URL: https://doi.org/10.1080/17538947.2013.847125, doi:10.1080/17538947.2013.847125, arXiv:https://doi.org/10.1080/17538947.2013.847125.

Eurostat, N., 2008. Rev. 2–statistical classification of economic activities in the european community. Office for Official Publications of the European Communities, Luxemburg .

Gil, J., 2015. Building a Multimodal Urban Network Model Using OpenStreetMap Data for the Analysis of Sustainable Accessibility. Springer International Publishing, Cham. pp. 229–251. URL: https://doi.org/10.1007/978-3-319-14280-7_12, doi:10.1007/978-3-319-14280-7_12.

Girres, J.F., Touya, G., 2010. Quality assessment of the french openstreetmap dataset. Transactions in GIS 14, 435–459. URL: http:https://doi.org/10.1111/j.1467-9671.2010.01203.x, doi:10.1111/j.1467-9671.2010.01203.x.

Haklay, M., 2010. How good is volunteered geographical information? a comparative study of openstreetmap and ordnance survey datasets. Environment and Planning B: Planning and Design 37, 682–703. URL: https://doi.org/10.1068/b35097, doi:10.1068/b35097, arXiv:https://doi.org/10.1068/b35097.

Horni, A., 2013. Destination choice modeling of discretionary activities in transport microsimulations. Ph.D. thesis. ETH Zurich. doi:10.3929/ethz-a-010006641.

JOSM, 2018. Josm. URL: https://josm.openstreetmap.de/.

Mashhadi, A., Quattrone, G., Capra, L., Mooney, P., 2012. On the accuracy of urban crowd-sourcing for maintaining large-scale geospatial databases, in: Proceedings of the Eighth Annual International Symposium on Wikis and Open Collaboration, ACM. p. 15.

Nagel, K., Axhausen, K.W., 2001. Workshop report: Microsimulation. doi:10.3929/ethz-a-004241835. draft - Contribution for D.A. Hensher (Ed.) (2001) The leading edge in travel behaviour research, Pergamon, Oxford.

OpenStreetMap Wiki, 2018a. Openstreetmap wiki. URL: https://wiki.openstreetmap.org/wiki/Elements.

OpenStreetMap Wiki, 2018b. Openstreetmap wiki. URL: https://wiki.openstreetmap.org/wiki/DE:Osmosis.

Ortúzar, J.d.D., Willumsen, L.G., 2011. Data and space, in: Modelling Transport. John Wiley & Sons, Ltd, pp. 55–137. doi:10.1002/9781119993308.ch3.

Ramm, F., Topf, J., 2010. OpenStreetMap: Die freie Weltkarte nutzen und mitgestalten. Lehmanns Media.

UNESCO Institute for Statistics (UIS), 2012. International Standard Classification of Education: ISCED 2011. UIS, Montreal, Quebec.

Vogt, W., 2006. Hinweise zur Schätzung des Verkehrsaufkommens von Gebietstypen. volume 147. fgSV Verlag.

Zielstra, D., Zipf, A., 2010. A comparative study of proprietary geodata and volunteered geographic information for germany, in: 13th AGILE international conference on geographic information science.

Zilske, M., Neumann, A., Nagel, K., 2011. Openstreetmap for traffic simulation .