ANALYSIS AND EVALUATION OF GEOSPATIAL FACTORS IN SMART CITIES: A STUDY OF OFF-STREET PARKING IN MAINZ, GERMANY

A. Rolwes *, K. Böhm *

* i3mainz, Institute for Spatial Information and Surveying Technology Mainz University of Applied Sciences, D-55128 Mainz, Germany - (alexander.rolwes, klaus.boehm)@hs-mainz.de

Commission IV, WG IV/3

KEY WORDS: geospatial analysis, metric of geospatial impact, smart city planning, big data, geo analytics.

ABSTRACT:

Geospatial data often build the basis for planning decisions in smart cities. In the decision-making process, geospatial relationships have not yet been fully considered and represented. The aim of our research is to investigate these geospatial relationships. This paper presents a four-step process to identify geospatial key factors in smart city use cases. We further develop and evaluate an existing metric to measure the impact of geospatial factors in urban area. For this, we use three variables to characterise distance decay, opening hours and an attractiveness weight for customers depends on the use case. The use case of this study is off-street parking over a period of four years in Mainz, Germany. The results show temporal relationships between parking and geospatial factors. Consequently, our research indicates the impact of different factors on car parks. This knowledge enriches decisions for sustainable planning in cities.

1. INTRODUCTION

Increasing urbanisation promotes the need of application-oriented analysis options for the development of intelligent city concepts (Woods, 2019). The European Commission pronounced to make cities “more efficient [...] for the benefit of its inhabitants and business.” (n.d.). Technology enables sustainable planning in cities. Therefore, the mass of raw data generated every day requires a better understanding and evaluation in the future (Moustaka et al., 2018; Etezadzadeh, 2020). Smart geospatial data form the backbone of future cities (Coors, 2015). The increasing mass of availability extends the diversity of analysis to take planning decisions. Urban planning needs impulses from science and business in order to recognise relation and determine the spatial impact (Engelke et al., 2019; Soike et al., 2019). In these analyses data scientist usually work together with domain experts. They tend to put personal assessments on geospatial factors, even though they are not in the position to evaluate them neutrally (Hong et al., 2020). Therefore, it is important to find a way to consider the opinion of the domain expert, but also to keep a non-biased view on the issue. In order to gain knowledge, geospatial factors needs specific identification and analysis in the context of the use case. In the following, geospatial factors are geospatial information, which have a direct and an indirect impact on the use case. Examples for geospatial information are health care, leisure activities or restaurants. In our paper, we present a metric of geospatial impact to measure the impact of geospatial factors in smart city use cases. In the analysis and evaluation, we consider different parameters and combine existing approaches. Numerous research exist in literature that determine the significance of geospatial data on different mobility use cases in smart cities. For example, Wagner et al. (2013) and Wagner et al. (2014) investigate the relationship between potential e-mobility charging stations and geospatial data for smart city planning. The approach uses regression analysis involving walking distances, points of interest (POI) with distance decay (van der Goot, 1982), census data and other parameters. The results show the impact of different geospatial factors on the use of charging stations. These investigations provide a basis for future planning of charging stations through the optimisation of their locations. The research of Bendler et al. (2014) uses regression analysis to explain the relation between different crimes and geospatial factors in San Francisco. They use tweets as geo social media data to understand characteristics of crime incidents. The work of Klemmer et al. (2016) also uses regression analysis to identify the space-time relationship between car sharing mobility behaviour and POIs in Amsterdam. They subdivide the study area into cells and merge the datasets of car sharing endpoints and POIs with kernel density estimation (KDE). The results show different key factors on car sharing behaviour in different time slots. Willing et al. (2017) extend the investigation of car sharing in Amsterdam. They select geospatial factors with gradient boosting machine and obtain the ten most important predictor variables for the distribution of different times. The results of the regression analysis generate knowledge, which they transfer on a further city (Berlin) to predict car sharing behaviour and approve their approach. Wang and Chen (2020), Schimohr and Scheiner (2021) and Wang et al. (2021) apply the approaches to bike sharing use cases on different continents. They identify a space-time relationship between bike sharing and geospatial factors with regression analysis and diverse methods of machine learning.

The main weakness of many existing studies is the insufficient consideration of POI opening hours and a weight of each POI in the use case. Opening hours may provide causal explanations and avoid positive relationships at times when POIs are closed. In addition, the attractiveness of each POI vary within the use case. Therefore, it is important to process both information. In order to gain a holistic view about the issue, we investigate POI opening hours with distance decay and the weight of each POI in the use case in combination. Our method extends existing approaches and equations, so that the impact of geospatial factors is measurable. In this paper, we aim to respond the following research question: How the impact of geospatial factors can be measured in smart city use cases?

* Corresponding author
Our approach shows a process to identify geospatial key factors and expands knowledge on the impact in smart cities to support planning decisions.

2. USE CASE AND DATA
The research area of this study is Mainz, Germany. Our data set contains two parts, parking data and geospatial data. A local parking company provides us data on off street parking over a period of four years, 2015 - 2019. In our use case, the parking company with its management represents the domain expert. The data preparation identified 13 car parks with continuous data about the parking occupancy in 60-min intervals. In order to investigate temporal relationships, we need to split the data into the week intervals ‘Weekday’ (Monday-Friday), ‘Saturday’ and ‘Sunday’. We treat public holidays as Sundays. Further, every week interval we split into four day intervals, night (00:00 - 07:00), morning (07:00 - 12:00), afternoon (12:00 - 18:00) and evening hours (18:00 - 00:00). In addition, we consider geospatial data. Our data basis for geospatial information is OpenStreetMap\(^1\). In our study, POIs are possible destinations of a parking customer like restaurants, shopping stores or cinemas.

3. METHODOLOGY
In order to identify geospatial key factors we pass through the following process (Figure 1). We divide the process into four steps. The most important part is step (c). Subsequent, we explain the different steps in detail and relate them to our use case.

(a) Definition and selection of geospatial factors

(b) Clustering and Categorisation

(c) Metric of geospatial impact

(d) (Geo) analytics

Figure 1: Process to identify geospatial key factors

3.1 Definition and selection of geospatial factors (a)
We define and select geospatial factors with the knowledge of the domain expert and the existing investigations about parking decision made by car drivers (Forschungsvereinigung Automobiltechnik e.V., 2015; Stadtplanungsamt der Stadt Regensburg, 2017; Valizade-Funder et al., 2018). For the customer the attractiveness of each destination varies depending on the use case. For example, an ATM in the city centre is a very attractive destination to park the car in a car park garage. In contrast, a finance appointment at the bank is an attractive destination. Therefore, it is important to select factors, which are attractive for parking customers. The attractiveness of the geospatial factors also change across the time intervals of the week and the day. This information affects the selection process and reinforces the determination of the intervals.

In addition to this, an exploratory data analysis provides initial characteristics of car parks. Figure 2 shows the increasing parking occupancy in the car park ‘Cinestar’ before film starts in a nearby cinema. The analysis indicates an impact from the cinema on the car park. The example supports the importance of consideration of attractiveness and characteristics on car parks.

Figure 2: Increasing parking occupancy before film starts

In the following, we investigate the car park ‘Kronberger Hof’, because it allows different perspectives for the use case. This car park is in the city centre and in the surrounding area are many shops, restaurants and leisure time activities. Overall, we define and select more than 6,000 POIs in 255 different OpenStreetMap amenities, who are relevant for the parking scenario.

3.2 Clustering and Categorisation (b)
In the second step, we cluster the amenities and categorise eleven geospatial factors according to Wagner et al. (2014), Bendler et al. (2016), Klemmer et al. (2016), Valizade-Funder et al. (2018) and Schimohr and Scheiner (2021). The defined categories are shopping, health, food services, leisure time, grocery, services and specialty retail, finance and insurance, education, public sector, religion and others. This categorisation allows us to determine a relationship between them and the parking occupancy in the following steps.

3.3 Metric of geospatial impact (c)
Based on the results in the first two steps, we develop our metric of geospatial impact. Our metric extends existing approaches and the equation of Wagner et al. (2014) by the additional variables for opening hours and weights for POIs in the use case. It shows the sum of all relevant values for a defined category \(j\), which corresponds with the density of a POI category \(j\) at a point \(i\) (1).

\[
x_q = \rho_j = \sum_{i} p_i \cdot r_{ij} \cdot \omega_j \cdot \omega_k
\]  

where \(x_q\) = sum of all relevant values for a particular POI category \(j\) at point \(i\), \(\rho_j\) = density of POI category \(j\) at point \(i\), \(p_i\) = specific POI \(k\), \(P_j\) = set of POI that belong to category \(j\), \(r_{ij}\) = distance decay of a POI \(k\) to point \(i\), \(\omega_j\) = opening hour of POI category \(j\), \(\omega_k\) = weight of POI \(k\) in the use case

Point \(i\) represents the destination, on which we measure the impact. In our case, it represents an individual car park. KDE is not a solution for us to calculate the density, because car parks or

---

\(^1\) https://www.openstreetmap.de/
their capacities do not vary in their location at different times. The equation considers three work steps with three variables as result. First, $r_j$ represents the weights of a reachability analysis. Starting by a point $i$, the attractiveness of a POI decreases with the distance (van der Goot, 1982). A customer prefers short walking distances to reach his destination. Isochrones and hot spots on Figure 3 show the reachable POIs from the car park ‘Kronberger Hof’.

We determine $\omega_j$ as weighted opening hours for each category at the corresponding time, because we do not have this information for each POI. Therefore, we not assign a POI by True or False. Third, the variable $\omega_j$ represents the weight of an individual POI in the use case. As mentioned before, the attractiveness of a POI also change with the use case and thus its impact varies. For example, a bakery is unlikely the only destination of parking customers. A bakery is a potential secondary destination. In contrast, a large supermarket is a potential primary destination for car drivers. Nevertheless, with the knowledge that many supermarkets provide their own parking spaces. The example shows the need of weighting POIs according to the use case. We classify POIs in primary and secondary destinations and assign the weight $\omega_j$. The classification bases on Valizade-Funder et al. (2018) and domain expertise.

In summary, with calculation of these three variables we complete our metric of geospatial impact. We receive a common density of each POI category at every car park. This value represents one of the explanatory variable in the following analytics step.

### 3.4 (Geo) analytics (d)

In the last step of the process, we analyse and identify geospatial key factors with (geo) analytics. In order to elaborate relations between parking and geospatial factors, we use slot wise multiple linear regression analysis. The parking occupancy represents the dependent variable, while the density of each POI category at every car park represents the explanatory variable. Further predictors like weather, events and holidays can enrich the analysis. The following chapter describes and discuss the results of the analysis.

### 4. RESULTS AND DISCUSSION

In this part of our paper we analyse and identify geospatial key factors with slot wise multiple linear regression analysis. As mentioned before, our data analysis of the categories with mapped opening hours indicates a weakness, because some categories show a low availability of POIs with mapped opening hours. Table 2 shows the percentage of the number of POIs of each category with the mapped opening hours.

| Category | Percentage of POI with mapped opening hours |
|----------|---------------------------------------------|
| leisure | 22.64 %                                      |
| health | 14.21 %                                      |
| food services | 19.92 %                                     |
| shopping | 11.89 %                                     |
| religion | 5.90 %                                      |
| education | 4.18 %                                      |
| others | 7.98 %                                      |
| leisure time | 1.18 %                                    |
| religion | 0.51 %                                      |

The main weakness are the categories with lower percentage. Nevertheless, we also calculate the ratio for them and extrapolate it to the entire category. For example, the category religion includes churches, rectories or mosques. This category only has $0.51 \%$ of the POIs with mapped opening hours. If individual outlier POIs are available here, they represent the opening period of an entire category. This indicates spurious correlations. In addition, multicollinearity is another problem, because several categories are similar or even identical to each other. For this reason, we opt the first five categories with more than $10 \%$ of POIs with mapped opening hours for our regression analysis. The following results focus on the car park ‘Kronberger Hof’. This car park contains 43,815 parking occupancy data. The
reachability analysis identifies 1,104 POIs surrounding it in a maximum radius of 800 m by foot. Table 3 presents the standardized regression results for the different time slots of the car park ‘Kronberger Hof’.

Table 3: Standardized regression results of slot wise multiple linear regression analysis of the car park ‘Kronberger Hof’

| Time of day | Services and specialty retail | Grocery | Health | Food services | Shopping | Adjusted R-squared |
|-------------|-------------------------------|---------|--------|---------------|----------|-------------------|
| 06:00 - 07:00 | 0.11*** | 0.000 | 0.000 | 0.000 | -0.071 | 0.035 |
| 07:00 - 08:00 | 0.02*** | 0.034 | 0.126 | 0.244*** | 0.171* | 0.439 |
| 08:00 - 10:00 | -0.10 | 0.062 | 0.230** | 0.168*** | -0.113*** | 0.405*** |
| 18:00 - 20:00 | -0.09 | 0.020 | 0.049 | 0.403*** | 0.112 | 0.335 |
| 20:00 - 22:00 | 0.00 | 0.000 | 0.000 | -0.109 | 0.114 | 0.033 |
| 22:00 - 00:00 | 0.00 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |

On weekdays before and during the working hours in the morning, the geospatial factors services and specialty retail (b = 0.255; p < .001) and food services (b = 0.244; p < .001) show a highly positive significant effect on the parking occupancy. We explain this with the location of the car park. For example, many hairdressers or opticians and café, coffee stores or bakeries surround the car park. Hairdresser or optician appointments are often in the morning and workers take advantage of short breaks for a snack. Furthermore, shopping (b = 0.171; p < .05) shows a significantly positive effect in the same time slot. Parking customers use shopping stores for purchases. On weekday afternoons, health (b = 0.240; p < .001), and food services (b = 0.168; p < .001) show a highly positive significant effect. In addition, shopping (b = 0.133; p < .001) demonstrates a highly negative significant effect. Due to a low adjusted R-squared (0.058), we partly question this correlation. On weekday evenings, food services shows a highly positive significant effect on the parking occupancy in ‘Kronberger Hof’ (b = 0.405; p < .001). Consequently, this category represents a geospatial key factor. We justify it by many open bars, restaurants and pubs near to the car park, where parking customers spend their evenings for dinner or a drink. On Saturday mornings, food services (b = 0.274; p < .001) and shopping (b = 0.227; p < .05) influence the occupancy highly positive significant and positive significant. Many people use days off to visit the city centre, go shopping or meet friends and family in restaurants or café. In the same time slot, the highly negative significant correlation for grocery (b = -0.195; p < .001) appears doubtful. Especially, the popular farmer’s market or a large supermarket are in a short walking distance to the car park. It requires a category review and additional investigation. On Saturday afternoons, only shopping (b = 1.616; p < .05) displays a positive significant effect. Saturday as the shopping day in Germany explains this effect. Due to a weak adjusted R-squared (0.141), we handle this statement with caution. On Sundays or public holidays, food services demonstrates a highly positive significance in the morning (b = 0.217; p < .001) and in the evening (b = 0.207; p < .001). In contrast, this category shows a highly negative significant effect in the afternoon (b = -0.145; p < .001). It confirms our expectations, but needs additional investigation due to the low adjusted R-squared.

For evaluation, we check the results for causality and discuss them with our external domain experts. As described above, the research results often show expected correlations. We confirm them through implicit knowledge of the city. Nevertheless, some correlations require further investigation. In order to eliminate the spurious correlations mentioned at the beginning, we recommend additional research.

5. CONCLUSION AND FUTURE WORK

Growing cities come along with a rising demand for application-oriented analysis options to address urban challenges. The mobility transformation turns research of mobility use cases like car parking in the political spotlight. Our approach shows a process to identify geospatial key factors in smart cities. The presented metric of geospatial impact allows measuring the effect of geospatial factors. It considers existing distance decay in combination with an approach to integrate opening hours and an attractiveness weight of the POI depending on the use case. First, regression results identify different key factors on various car parks. An evaluation with domain experts confirms the characteristics of the individual car parks, but at the same time it shows the need for action. Based on existing open source data, our solution contributes to sustainable urban planning.

In future work, we recommend developing this approach further on two ways. First, we will optimise the metric of geospatial impact. We improve the discussed problem with opening hours by adding opening hours of POIs from outside the study area. Other cities have a higher number of POIs with mapped opening hours. Similar German cities enrich categories with limited information on opening hours (e.g. religion), so that we get more reliable results without spurious correlations. In addition, the occupancy information of the POI can generate a benefit for this research area. An open POI represents a potential destination, but it does not provide information about the occupancy at different times. An occupancy weight can optimise the metric. Moreover, we will transfer this approach to another mobility use case in a different city to investigate the applicability of the developed method. Artificial intelligence offers further analysis methods to analyse and identify geospatial key factors, e.g. classification methods. Second, we will visualise the results of the examination to make it comprehensible for the domain expert by using visual methods. Currently, the results show relationships between urban use cases and geospatial factors by using statistical tables and figures. In order to support the work of the domain experts in combination with the data scientist, we need customised visualisations for knowledge transfer.

ACKNOWLEDGEMENTS

The research for this paper is part of the project ‘BAM – Big-data-analytics in environmental and structural monitoring’ at Mainz University of Applied Sciences. The Carl Zeiss Foundation funds it. The authors like to thank the local parking company for the excellent cooperation and provision of data. The authors also like to thank Pauline Radu for supporting the project with excellent data analyses during her student internship. The authors thank for proofreading the article. This study is part of an ongoing Ph.D. research at RheinMain University of Applied Sciences and Mainz University of Applied Sciences.

REFERENCES

Benderl, J., Brandt, T., Neumann, D., 2016. An Open Door May Tempt a Saint - Data Analytics for Spatial Criminology. 24th European Conference on Information Systems, ECIS 2016, Research Paper 98.

Benderl, J., Ratku, A., Neumann, D., 2014. Crime Mapping through Geo-Spatial Social Media Activity. Thirty Fifth International Conference on Information Systems.

This contribution has been peer-reviewed.

https://doi.org/10.5194/isprs-archives-XLVI-4-W1-2021-97-2021 | © Author(s) 2021. CC BY 4.0 License.
Coors, V., 2015. Ohne smarte Geodaten keine smarten Städte. 
zfv – Zeitschrift für Geodäsie, Geoinformation und 
Landmanagement (4/2015), 244–248.

Engelke, D., Hagedorn, C., Schmitt, H.-M., Buechel, C., 2019.
Raumwirksamkeit der Digitalisierung. Ergebnisse einer 
breit angelegten Delphi Umfrage.

Etezadzadeh, C., 2020. Smart City – Made in Germany. 
Springer Fachmedien Wiesbaden, Wiesbaden.

European Commission, n.d. Smart cities, 
https://ec.europa.eu/info/eu-regional-and-urban-
development/topics/cities-and-urban-development/city-
initiatives/smарт-cities_en. (Accessed 5 May, 2021).

Forschungsvereinigung Automobiltechnik e.V., 2015. Auskunft 
über verfügbare Parkplätze in Städten.

Hong, S.R., Hullman, J., Bertini, E., 2020. Human Factors in 
Model Interpretability: Industry Practices, Challenges, and 
Needs. Proceedings of the ACM on Human-Computer 
Interaction 4 (CSCW1), 1–26.

Klemmer, K., Willing, C., Wagner, S., Brandt, T., 2016.
Explaining Spatio-Temporal Dynamics in Carsharing: A 
Case Study of Amsterdam. Twenty-second Americas 
Conference on Information Systems, San Diego.

Moustaka, V., Vakali, A., Anthopoulos, L.G., 2018. A 
Systematic Review for Smart City Data Analytics. ACM 
Computing Surveys 51 (5), 1–41.

Schimohr, K., Scheiner, J., 2021. Spatial and temporal analysis 
of bike-sharing use in Cologne taking into account a public 
transit disruption. Journal of Transport Geography 92, 
103017.

Soike, R., Libbe, J., Konieczeck-Woger, M., Plate, E., 2019.
Räumliche Dimensionen der Digitalisierung.
Handlungsbedarfe für die Stadtentwicklungsplanung. Ein 
Thesenpapier. Difu-Sonderveröffentlichung.

Valizade-Funder, S., Funder, J., Sehi, R., 2018. Situation des 
innerstädtischen Einzelhandels in Mainz.

van der Goot, D., 1982. A model to describe the choice of 
parking places. Transportation Research Part A: General 
16, 109–115.

Wagner, S., Götzinger, M., Neumann, D., 2013. Optimal 
location of charging stations in smart cities. A points of 
interest based approach. Thirty Fourth International 
Conference on Information System.

Wagner, S., Neumann, D., Brandt, T., 2014. Smart city planning 
- Developing an urban charging infrastructure for electric 
vehicles. Twenty Second European Conference on 
Information Systems, Tel Aviv.

Wang, K., Chen, Y.-J., 2020. Joint analysis of the impacts of 
built environment on bikeshare station capacity and trip 
attractions. Journal of Transport Geography 82, 102603.

Wang, X., Cheng, Z., Trépanier, M., Sun, L., 2021. Modeling 
bike-sharing demand using a regression model with 
spatially varying coefficients. Journal of Transport 
Geography 93, 103059.

Willing, C., Klemmer, K., Brandt, T., Neumann, D., 2017. 
Moving in time and space – Location intelligence for 
carsharing decision support. Decision Support Systems 99, 
75–85.

Woods, E., 2019. Data analytics: The key to delivering smart 
cities. Smart Energy International (5), 104–105.