Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 00 (2024) 000–000

Procedia Computer Science

www.elsevier.com/locate/procedia

The 15th International Conference on Ambient Systems, Networks and Technologies (ANT)
April 23-25, 2024, Hasselt, Belgium

Estimation of informal park-and-ride

Luk Knapen[a,*], Tom Bellemans[a], Elenna Dugundji[b], Rob van der Mei[c]

[a]Hasselt University, Martelarenlaan 42, 3500 Hasselt, Belgium
[b]MIT Center for Transportation and Logistics, 1 Amherst St, Bldg E40 2nd fl, Cambridge, MA 02142 USA
[c]CWI, Science Park 123, 1098 XG Amsterdam, The Netherlands

Abstract

Curbside parking in a residential area may be induced by the presence of a public transport stop. Travelers may park their car near the stop and continue their trip to the city center by other means. This is called *informal park-and-ride*. The magnitude of the phenomenon is estimated by simulation. Parking demand is derived from the history of parking time (tickets) sold at vending machines. For each ticket, an activity location is determined by stochastic sampling from a buildings (facilities) database based on the position of the vending machine. The activity timing is derived from the parking duration specified by the ticket. Suitable parking spots for an activity are determined for the cases (i) *drive-park-walk* and (ii) *drive-park-publicTransport-walk* respectively. The generalized cost (based on money and travel duration) is determined for both options. The decision is sampled by means of a behavioural model. Several scenarios are considered and the results allow to evaluate the complaints issued by residents of a study area because microsimulation enables the generation of probability densities for parking occupation in such area in each scenario. This paper reports how the method is applied to a study area in Amsterdam.

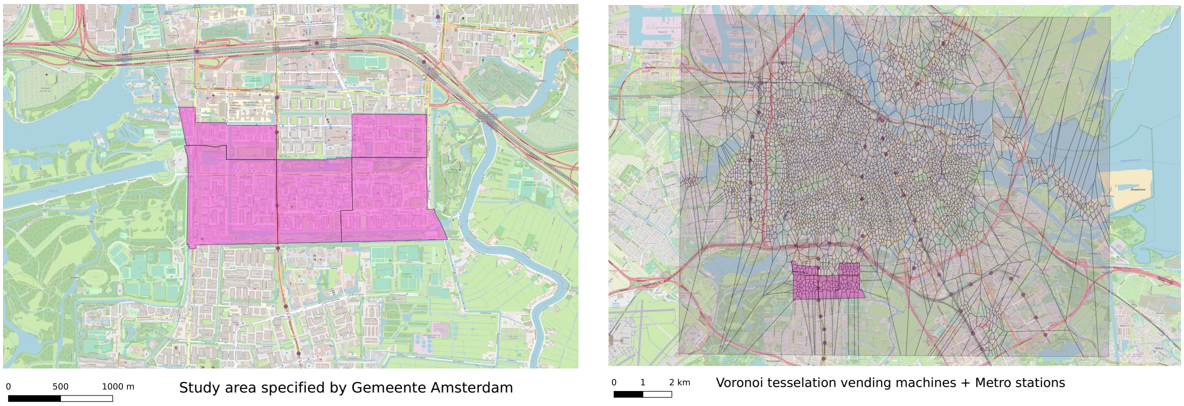
Keywords: park and ride; stochastic microsimulation; scenario analysis; I-Splines

1. Introduction - Problem Statement

Modifications in transportation infrastructure often cause people to complain about unwanted side effects. This paper presents a tool to evaluate spatial changes in parking pressure caused by changed accessibility and parking fare policy. The method is applied to Amsterdam-Zuid. Many inhabitants of Amsterdam-Zuid suggest increased parking pressure in their neighbourhood during weekends due to the opening of the new Noord-Zuid metro line and the raise of parking fares in the city center. They suggest that the phenomenon is caused by *informal park and ride* (iP+R). In Amsterdam *scan cars* continuously identify parked cars in all streets where parking is not free of charge. However, no parking observations are available for Amsterdam-Zuid where parking is free during weekends. This raises the

* Corresponding author. Tel.: +32-11-26.91.11
E-mail address: luk.knapen@uhasselt.be

1877-0509



(a) Study Area in Amsterdam

(b) Metro lines and Voronoi tessels based on vending machines.

Fig. 1: *Activity* area (enclosed by ring road), *study* area (pink), metro stations and Voronoi tessels each containing one vending machine.

research question: "*To what measure do the availability of new travel options and new parking policies in particular regions affect the situation elsewhere ?*"

The study area (stu-area) (shown in Figure 1a) is the residential area for which overload by iP+R needs to be investigated. Parking is free on Saturday in this area.

We consider the set of neighbourhoods (Dutch: buurten) having paid parking on Saturday as the *activity area* (act-area). Buildings in this area constitute the destinations for trips for which a parking ticket is purchased at one of the vending machines in the area. Figure 1b shows both the stu-area and the act-area (the city center for which parking data on Saturday are available).

2. Related Research

2.1. Parking Duration

Several researchers focus on models to predict parking duration. Some examples are given below.

[2] approximate parking duration by a log-normal distribution fitted to a 246 respondents survey for mall visitors but the authors do not provide an indication of goodness-of-fit. For long duration parking a Weibull distribution based on synthetic data is used.

[1] evaluate the feasibility of using EV parked in a commercial P+R parking lot as an energy store in a V2G context. A dataset listing all arrivals and departures in the parking lot is available. The time of arrival and the parking duration are relevant parameters. A gaussian kernel is used to approximate the probability density function (PDF) for the parking duration. However, the crucial *bandwidth* parameter h is not specified in the paper.

[8] study a parking lot using 370 wireless IoT sensors to monitor car presence on each parking place. Parking event start and stop times are used to determine a distribution for parking duration. The authors claim and show in a graph that the Weibull distribution nicely fits to their observations.

[7] collected data using a survey for several parking lots in Delhi, India. They use a three layer ANN to predict the parking duration class ($< 2h, 2h-4h, 4h-6h, 6h-8h, > 8h$) from income (6 classes), classes, profession (6 classes), hourly parking fee, travel distance, trip purpose and travel time.

Several papers (e.g. [12]) predict parking lot occupancy. Few papers mention the parking duration distribution probably due to lack of access to commercial data.

2.2. Park-and-Ride

[11] present a discrete choice model that maps (i) *personal* attributes (such as gender and income), (ii) *travel* attributes (parking cost, trip duration and cost for private car and public transport, transfer walk distance and wait

time) and (iii) *attitude* w.r.t. park-and-ride (P+R) onto three travel mode options: *car*, *PT* and *P+R*. The study aims to support the solution of the trip distribution problem and to estimate the required capacity for P+R parking lots. Several examples in Chinese cities are mentioned. A multinomial model is estimated and discussed. Such model is essential for the daily travel plan generator in order to determine which fraction of the trip demand shall be assigned to P+R. However, the ratio between the reported Chinese unit costs for travel (gasoline, parking) to the monthly income differs too much from the European context so that the numeric results cannot be used.

[3] state that multimodal trips have been frequently researched but never specifically focusing on P+R. The research covers cross-regional trips in the Toronto-Hamilton region and focuses on commuting work trips found in a five-yearly survey that mentioned *transit* as the main mode and *car* as access mode. The authors combine the 2006 travel survey with additional data collected from transit service operators and data describing the properties of the transfer stations. An MNL model is built to support the selection of the P+R station. Both regional transit trains (GO network) (53 P+R stations) and the Toronto Transit Commission (TTC) subway (15 P+R stations) are considered. The choice set for an individual consists of the 5 *GO stations* and 3 *TTC stations* closest to the traveller's home. One of the independent variables is the *relative station direction* which is the angle between the *home-station* and *home-workLocation* lines. This model design decision is not argued in the paper and may have been chosen for technical simplicity. Euclidean distance from origin to stations is used.
Unfortunately, the resulting models do not fit our purpose because (i) trips for which the main mode was reported to be *transit* were considered and (ii) the concept of *relative station direction* may be inappropriate to study European cities providing P+R parking lots near to the city border in order to promote *last-mile transit* and (iii) the model is estimated for commuting trips.

[4] describe the activity based travel demand generator *GTAModel V4* [1] for Toronto. Locations are predicted at TAZ (Traffic Analysis Zone) level by means of a time prism and based on the auto network travel times. Mode choice follows location choice in the prediction procedure. The PT station(s) choice model is tour based i.e. it takes the PT options for the return trip into account during station(s) selection.

[6] start from a survey among P+R users in Austin, USA and develop choice models taking into account (i) person attributes (ii) car trip attributes, (iii) transit attributes and (iv) P+R lot attributes. They estimated models are: (i) overall MNL model, (ii) observed heterogeneity (personal attributes), (iii) mixed model to account for unobserved heterogeneity and (iv) a mixed model similar to the previous one but in which correlation among observed quantities is specified by a correlation matrix. Unfortunately, the parking cost is not included (due to lack of data).

[10] apply both random utility maximization (RUM) and random regret minimization (RRM) to create choice models for P+R locations (both formal and informal). Models are estimated using properties of morning peak trips only although the time dependence of the general cost is discussed.

Our research considers both the forward and backward trips.

3. Solution Model and Procedure

Consider a study area near a public transport (PT) station where free curbside parking is allowed. We present a model to assess the contribution to parking pressure in the study area caused by people entering the city by car and executing activities in areas with paid parking. Such individuals choose between (1) a *pure car trip* with paid parking near the destination and (2) a *multimodal car-PT trip* involving free parking near a PT stop. The proposed method uses a general model that determines the set of suitable parking options for an activity of a particular duration at a specific location.
The available PT options depend on the time-of-day and need to be considered for both the forward (pre-activity) and backward (post-activity) PT trips. Walk segments and PT wait times play an important role in the model (via *value-of-time*). Furthermore, the parking pressure affects the distance between the parking spot and the PT stop. Hence, we opted for micro-simulation which allows for modeling these phenomena.

Starting from the parking demand observed from ticket sales data, *parking substitution* is investigated. This is equivalent to a multi-modal route choice problem as shown in Figure 2. For each parking event in the act-area an

[1] https://tmg.utoronto.ca/doc/1.6/gtamodel/

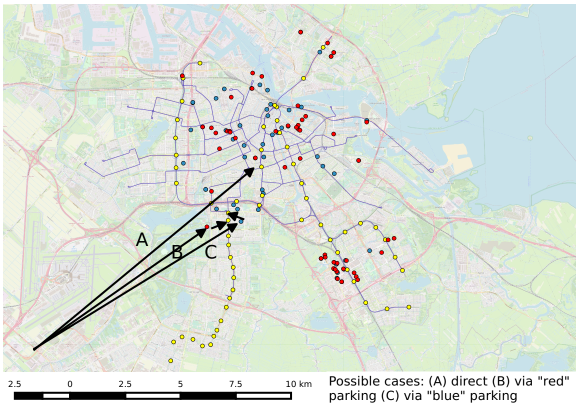


Fig. 2: Multi-modal route choice.

alternative in the stu-area is evaluated. Ultimately, the *difference in parking pressure* (as opposed to the total parking pressure) on the stu-area caused by changes in the act-area is investigated.

3.1. Assumptions

1. The demand is estimated from tickets purchased to park in the act-area on Saturday. Hence, the corresponding individuals are assumed to be occasional visitors (no residents, no workers, no students).
2. Metro stations near the stu-area are selected in advance. This does not affect the results but saves computing resources by options known to be infeasible.
3. In the stu-area, the *background parking pressure* by residents and local facility visitors is constant during the day.

3.2. Microsimulation steps

The microsimulator keeps track of the parking spot occupation as a function of time. The walking distance from each parking spot to each candidate metro station is computed in advance (because it is time independent). The initial parking pressure by residents in the stu-area is given as a scenario parameter. It specifies the probability that a parking spot is occupied for the complete day i.e. is not available for iP+R.

The number of city center visitors on Saturdays using the stu-area for iP+R is determined by microsimulation. The parking demand in the paid parking zones covered by vending machines is derived from the subset of NPR (Nationaal Parkeer Register[2]). Each NPR record specifies the number of tickets and the total parking duration sold for each vending machine (VM) and for each 1-hour period of the day. A distribution for the parking duration has been derived from NPR records reporting a single transaction only.

Each NPR record leads to a set of parking events for which the start time is determined by uniform sampling within the specified 1-hour period.

The car occupation and trip distance are sampled from the respective distributions derived from the app based travel survey ODiN[3] considering trips to *several* large cities (Amsterdam, Utrecht, Rotterdam and s-Gravenhage) on Saturday because of data sparsity in ODiN. The sampled distance is used to determine a set of candidate *trip origin* neighbourhoods by means of the distance along the road to the center of Amsterdam. The set of all residential

[2] https://www.nationaalparkeerregister.nl/
[3] https://www.cbs.nl/nl-nl/longread/rapportages/2021/onderweg-in-nederland–odin—2020-onderzoeksbeschrijving/2-onderweg-in-nederland–odin–

addresses belonging to one of the selected neighbourhoods is used for uniform sampling of an origin address. As a consequence, variation in population density is taken into account.

Voronoi tesselation is used to partition the act-area based on the set of VMs. For each parking event, an activity location is sampled from a set of addresses associated with the VM. Addresses suitable for activity execution (on a Saturday) are assigned to zero or more vending machines as follows:

1. the set of potential trip destinations does not include addresses labeled by *residential* or by work related labels such as *industry*, *office* and *education* (exclusion is used since an address may have multiple labels)

2. an address A is assigned to the vending machine $VM(A)$ in whose Voronoi tessel it is located (its nearest vending machine) or in one of the adjacent tessels

3. and only if the distance along the road $d_r(VM(A), A) \leq \overline{\delta}$ where $\overline{\delta}$ is the maximum egress walk distance specified by the 0.95 quantile value ($1551[m]$) for walking distance derived from the MOBIS dataset.

Once the attributes for the trip (origin, destination, number of persons in the car) and for the activity (location, duration) are known, two options are considered: (1) pure car trip with parking near the destination and (2) iP+R trip, parking in stu-area and continuing by metro. For the iP+R case. each metro station associated with the stu-area is considered together with its nearest available parking spot. The parking spot delivering the lowest generalized cost for the combined *forward* and *backward* trips is selected. Both trips (to the destination and back to the parking spot) are combined because the PT service level depends on time-of-day and may differ between stations served by different metro lines. This mechanism determines both the parking spot and the metro station used.

The monetary cost for a multimodal trip is based on the distance driven by car (and estimated fuel consumption), the parking cost (based on the duration and the hourly tariff) and the cost for the metro (based on the number of people in the car). Parking cost in the act-area is more expensive than in the stu-area but the parking duration in the latter is longer (because it includes the duration of the PT trips).

Finally, a basic choice model is used to select either the *direct* or the iP+R trip.

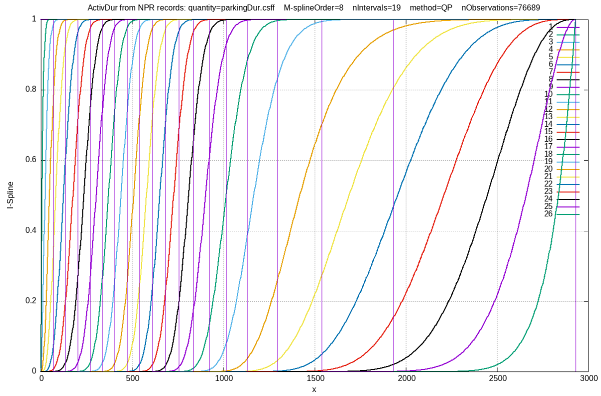
4. Data Sets Used - Parking Duration Distribution

4.1. Datasets Used

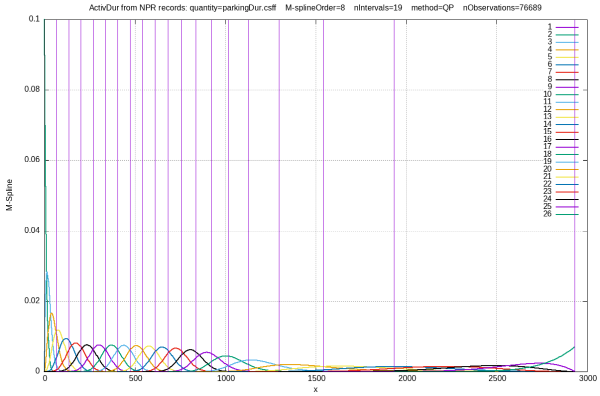
NPR	The extract for Amsterdam applies to 3610 vending machines and contains $22.5M$ hourly data records from 2018-Mar-01 to 2020-Apr-30. Approximately $1.4M$ records (containing 76689 different duration values) specify a single event only: these are used to estimate the parking duration distribution.
BAG	Building address database specifying coordinates an building functions.
Walk	Accurate data could not be extracted from ODiN (dataset too small, suffers from discretization due to self-reporting). An extract from the Swiss MOBIS dataset [5] describing the post parking walking leg was obtained from ETHZ.
VM	Vending machines: positions extracted from NPR.
pSpots	Parking spots: geometry and properties extracted from shape file obtained from Gemeente Amsterdam
GTFS	Metro services, lines, timetables downloaded from http://gtfs.ovapi.nl/
PT fares	Ticket price extracted from https://www.travelguide.amsterdam/en/publictransport-gvb/tickets-fares/
ODiN	Saturday trips to large cities to extract distance, timing and number of persons in the car.

4.2. Distribution for parking duration

The density seems to be composed of an exponential part and a gamma or log-normal part. Hence parameter estimations for these distributions turn out to lead to insufficient *goodness of fit*. Therefore, an empirical parking duration distribution was computed. The numerically specified cumulative distribution was approximated by a family of I-splines. These have the following interesting properties (see [9]):



(a) I-Spline family



(b) M-Spline family. Each function integrates to one over the interval.

Fig. 3: I-Spline basis functions for Csff (cumulative sample frequency function) approximation and underlying M-Splines.

1. the I-spline is a linear combination of M-splines which in turn are recursively defined as piecewise polynomials over the domain. Let k denote the order of the spline, N denote the number of family members, i denote the *family member* and t denote the set of knots (nodes). Then $|t| = k + N$ and

$$M_i(x|k,t) = \begin{cases} 1/(t_{i+1} - t_i) & if\ k = 1 \\ \dfrac{k[(x - t_i)M_i(x|k-1,t) + (t_{i+1} - x)M_{i+1}(x|k-1,t)]}{(k-1)(t_{i+1} - t_i} & if\ k > 1 \end{cases} \tag{1}$$

$$t_j \leq x < t_{j+1} \Rightarrow I_i(x|k,t) = \begin{cases} 0 & if \quad i > j \\ 1 & if \quad i < j - k + 1 \\ \sum_{m=i}^{j} (t_{m+k+1} - t_m)M_m(x|k+1,t)/(k+1) & otherwise \end{cases} \tag{2}$$

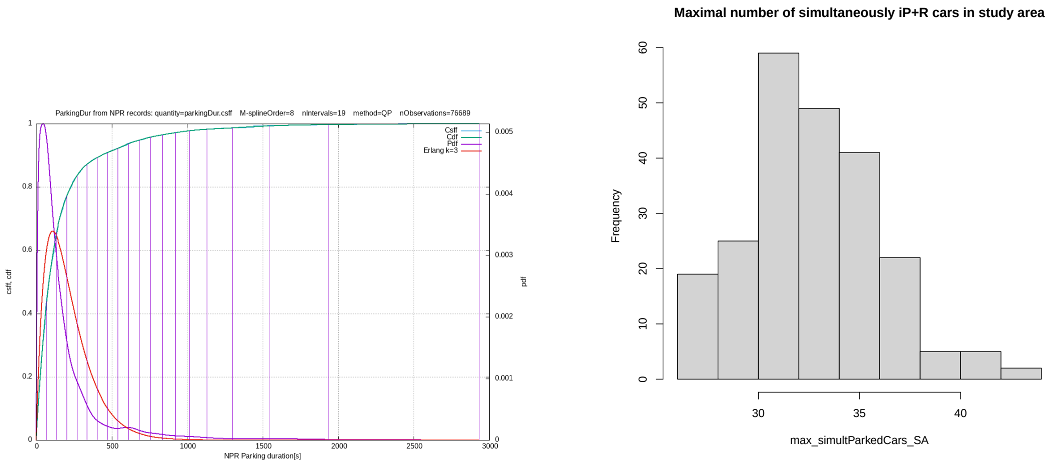
2. analytic derivation and integration $I_i(x|k,t) = \int_L^x M_i(u|k,t)du$
3. both the expected value and variance for the empirical distribution can be computed analytically from the coefficients of the approximation.

We approximate the observed cumulative relative frequency by a linear combination of members of an I-spline family. Computing an I-spline family requires two parameters: (1) the order k of the splines and (2) the number N of knots. For a given pair of parameters $\langle k, N\rangle$, the approximation is determined by minimizing the square of the deviation between the observed cumulative relative frequency and the approximation. The parameters have been determined by a greedy optimizer. From a given set of k and N values we determine the pair $\langle k, N\rangle$ for which the expected value of the resulting distribution which can be computed analytically (see [9]) is nearest to the sample mean. The resulting basis functions are shown in Figure 3. The resulting PDF and CDF are shown in Figure 4a. Parking duration is determined by inverse transform sampling using this CDF.

5. Tools Used

GraphHopper[4] is used to compute travel distance an duration (i) in advance to create impedance matrices for walk segments between parking spots, metro stations and activity locations and (ii) during the simulation to evaluate car

[4] https://www.graphhopper.com/



(a) Empirical distribution for parking duration. Empirical (pink) and estimated Erlang ($k = 3$) (brown) densities.

(b) Histogram for predicted maximum number of simultaneous iP+R cars in stu-area on 2019-Sep-21 (based on 231 simulation runs).

Fig. 4: Left: Observed parking duration in NPRrecords - Right: Predicted parking pressure induced by iP+R in the stu-area.

trips as soon as the origin is known and PT trips as soon as the start time is known (PT service depends on time-of-day). All data are collected in a postgresql-postgis database. Spatial results are shown by QGIS. The simulator is written in Python3.

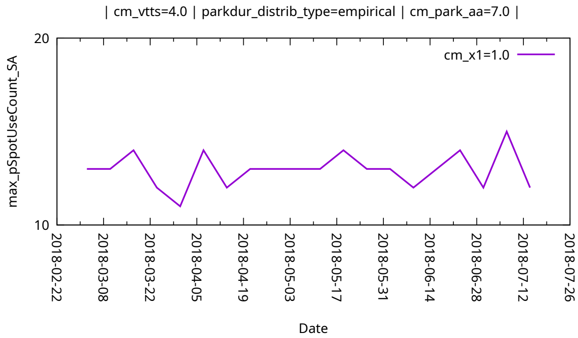
6. Scenarios Simulated - Results

For scenario analysis the parameters (i) VOT (value of time), (ii) parking tariff, (iii) type of parking duration distribution, (iv) behavioural model and (v) initial parking pressure are varied using values specified in the Table 1. No estimated choice model to select between *direct* and iP+R trips was available yet. Hence, the generalized cost difference $x = C_{iP+R} - C_{direct}$ is used as independent variable for the sigmoid function $P_{iP+R} = 1/(1 + e^{-\alpha\cdot(x-(x_1-x_0)/2)})$ specifying the sampling probability to choose iP+R, where x_1 corresponds to the cost difference for which iP+R is chosen with a given probability p and x_0 corresponds to the cost difference for which the direct trip is chosen with probability p. The value for α is computed from x_0, x_1 and p. These synthetic models are used to show the sensitivity of the results to the choice model that is still to estimated. Forty-eight (48) scenarios have been evaluated.

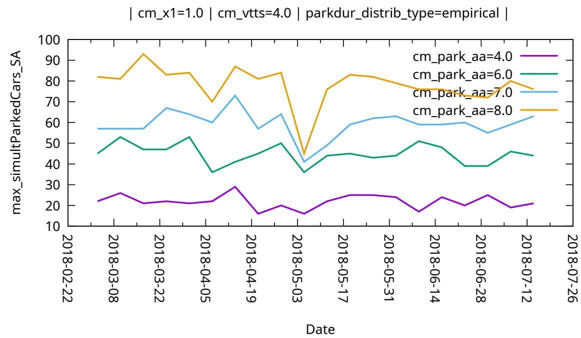
Table 1: Parameter values for scenario analysis.

Name	Values set
VOT [€/h]	$\{2, 4, 6\}$
Parking tariff [€/h]	$\{4, 6, 7, 8\}$
Parking duration distribution	{`Erlang`($k = 3$), `empirical`}
Behavioural (choice) model	$x_0 = -c, x_1 = c, p = 0.95$ for the cases $c \in \{1, 2\}$
Initial parking occupancy (pressure)	$\{0.8\}$

1. Parking occupation by iP+R during simulated periods: Figure 5 shows the parking spot occupation in the stu-area due to iP+R for a 5-month period of time. A particular parking spot can be occupied by multiple cars consecutively during a day. Figure 5a shows the maximum number of different cars *consecutively* occupying a particular parking spot during a simulated day. Figure 5b shows the maximum number of cars *simultaneously* parked in the stu-area for several values of the parking tariff in the act-area.

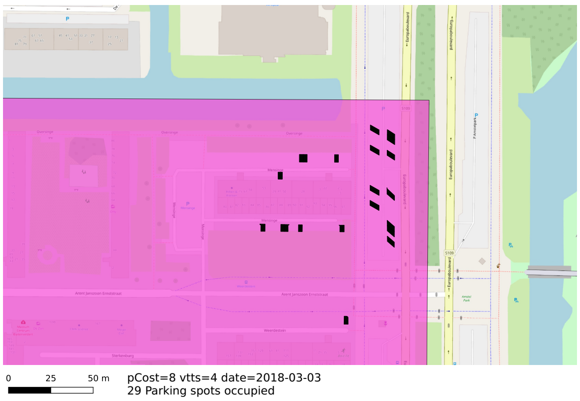


(a) Maximum number of times a parking spot is used stu-area.

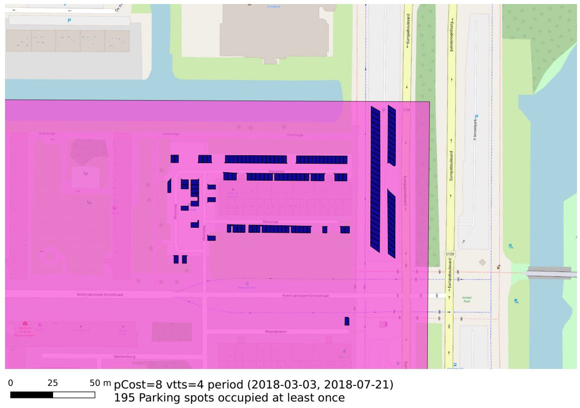


(b) Max number of simultaneously parked cars in stu-area for empirical, vtts=4, choiceModel=1 for a range of pCost values. Top curves correspond to higher parking fares.

Fig. 5: Parking spot occupation for a series of simulated days.

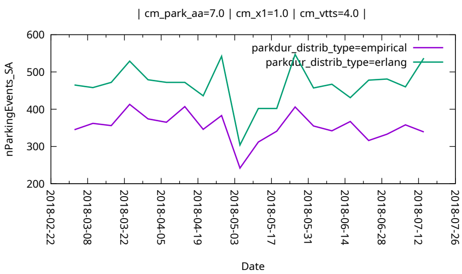


(a) Parking spots occupied at least once during a single day.

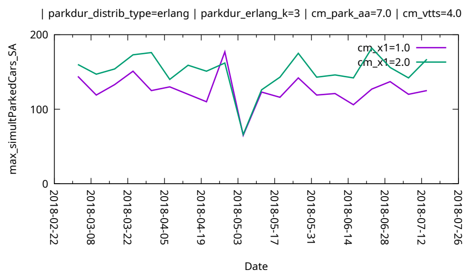


(b) Parking spots occupied at least once during a 5-month period.

Fig. 6: Parking occupation due to iP+R: spatial distribution. There is a metro station in the upper right corner of the pink area.



(a) Number of stu-area parking events for pCost=7, vtts=4, choiceModel=1 for two *parking duration distributions*. Top: erlangK=3, Bottom: empirical



(b) Max number of simultaneously parked cars in stu-area for ErlangK=3, pCost=7, vtts=4 for two *choice models*. The top curve corresponds to model 2 (less strict).

Fig. 7: Max number of simultaneously parked cars: scenarios comparison

Figure 6 shows the spatial distribution of parking spot occupation by iP+R related to activities in the act-area. Figure 6a shows the parking spots occupied by iP+R for a single simulated day. Figure 6b shows the parking spots occupied at least once by iP+R during a 5-month period.

2. Few cars having three or more passengers (carpooling) are involved in iP+R because the PT ticket cost starts to outweigh the parking cost in the act-area.
3. Figure 4b shows the histogram (231 simulation runs) for the predicted maximum number of simultaneously parked iP+R cars in the stu-area under the parking demand for 2019-Sep-21 (71964 parking events) for *duration=empirical, pCost=7, vtts=5*. It illustrates the combined effect of the stochastic models that constitute the simulation.

7. Discussion - Conclusion

The effect of iP+R is observed only in a small region near the metro stations in the stu-area. Detailed simulation at city level delivers interesting insights for policy evaluation but requires a large variety of datasets.

I-splines provide a good approximation for experimental data using a fairly small set of basis functions. This is important because inaccurate stochastic models may heavily affect distributions of the simulation results. Stochastic microsimulation delivers probability distributions for results.

Acknowledgements

The authors are grateful to ETHZ for the post-parking walking trip leg data extracted from the MOBIS dataset.

References

[1] Guner, S., Ozdemir, A., Serbes, G., 2016. Impact of car arrival/departure patterns on EV parking lot energy storage capacity, in: 2016 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS), pp. 1–5. doi:10.1109/PMAPS.2016.7764130.

[2] Magsino, E.R., Arada, G.P., Ramos, C.M.L., 2022. An Evaluation of Temporal- and Spatial-Based Dynamic Parking Pricing for Commercial Establishments. IEEE Access 10, 102724–102736. doi:10.1109/ACCESS.2022.3209806.

[3] Mahmoud, M.S., Habib, K.N., Shalaby, A., 2014. Park-and-Ride Access Station Choice Model for Cross-Regional Commuting: Case Study of Greater Toronto and Hamilton Area, Canada. Transportation Research Record 2419, 92–100. URL: https://doi.org/10.3141/2419-09, doi:10.3141/2419-09. _eprint: https://doi.org/10.3141/2419-09.

[4] Miller, E., Vaughan, J., King, D.M., Austin, M., 2015. Implementation of a "Next Generation" Activity-Based Travel Demand Model: The Toronto Case, in: TAC 2015: Getting You There Safely - 2015 Conference and Exhibition of the Transportation Association of Canada, Transportation Association of Canada (TAC), Ottawa, Ontario Canada. URL: https://api.semanticscholar.org/CorpusID:130580522.

[5] Molloy, J., Castro, A., Götschi, T., Schoeman, B., Tchervenkov, C., Tomic, U., Hintermann, B., Axhausen, K.W., 2022. The MOBIS dataset: a large GPS dataset of mobility behaviour in Switzerland. Transportation URL: https://doi.org/10.1007/s11116-022-10299-4, doi:10.1007/s11116-022-10299-4.

[6] Pang, H., Khani, A., 2018. Modeling park-and-ride location choice of heterogeneous commuters. Transportation 45, 71–87. URL: https://doi.org/10.1007/s11116-016-9723-5, doi:10.1007/s11116-016-9723-5.

[7] Parmar, J., Das, P., Dave, S.M., 2021. A machine learning approach for modelling parking duration in urban land-use. Physica A: Statistical Mechanics and its Applications 572, 125873. URL: https://www.sciencedirect.com/science/article/pii/S037843712100145X, doi:10.1016/j.physa.2021.125873.

[8] Piovesan, N., Turi, L., Toigo, E., Martinez, B., Rossi, M., 2016. Data Analytics for Smart Parking Applications. Sensors 16. URL: https://www.mdpi.com/1424-8220/16/10/1575, doi:10.3390/s16101575.

[9] Ramsay, J.O., 1988. Monotone Regression Splines in Action. Statistical Science 3, 425 – 441. URL: https://doi.org/10.1214/ss/1177012761, doi:10.1214/ss/1177012761. publisher: Institute of Mathematical Statistics.

[10] Sharma, B., Hickman, M., Nassir, N., 2019. Park-and-ride lot choice model using random utility maximization and random regret minimization. Transportation 46, 217–232. URL: https://doi.org/10.1007/s11116-017-9804-0, doi:10.1007/s11116-017-9804-0.

[11] Zhao, X., Li, Y., Xia, H., 2017. Behavior decision model for park-and-ride facilities utilization. Advances in Mechanical Engineering 9, 1687814017708907. URL: https://doi.org/10.1177/1687814017708907, doi:10.1177/1687814017708907. _eprint: https://doi.org/10.1177/1687814017708907.

[12] Zhao, Z., Zhang, Y., Zhang, Y., 2020. A Comparative Study of Parking Occupancy Prediction Methods considering Parking Type and Parking Scale. Journal of Advanced Transportation 2020, 5624586. URL: https://doi.org/10.1155/2020/5624586, doi:10.1155/2020/5624586. publisher: Hindawi.