Determinants of station-based round-trip bikesharing demand

Wilkesmann, Florian; Ton, Danique; Schakenbos, Rik; Cats, Oded

DOI
10.1016/j.jpubtr.2023.100048

Publication date
2023

Document Version
Final published version

Published in
Journal of Public Transportation

Journal of Public Transportation 25 (2023) 100048

Contents lists available at ScienceDirect

Journal of Public Transportation

journal homepage: www.journals.elsevier.com/journal-of-public-transportation

Determinants of station-based round-trip bikesharing demand

Florian Wilkesmann[a], Danique Ton[b], Rik Schakenbos[b], Oded Cats[a,*]

[a] Department of Transport and Planning, Delft University of Technology, Stevinweg 1, 2628 CN Delft, the Netherlands
[b] Research Department, NS Stations, Stationshal 17, 3500GM Utrecht, the Netherlands

ARTICLE INFO

Keywords:
Bikesharing
Demand analysis
Last mile
Public transport
Round-trip

ABSTRACT

First and last mile connectivity of public transport hubs is a key component in promoting multi-modal travel. The Dutch train station operator (NS Stations) promotes the combination of bike and train by offering a train station-based round-trip bikesharing (SBRT) scheme, known as 'OV-fiets', located at train stations throughout the country. This scheme allows users to rent a bike to travel between train stations and their destination and vice versa. The round-trip nature of the SBRT makes it unique in comparison to widely applied one-way bikesharing schemes. Little is known about the determinants of demand for round-trip bikesharing, especially when being integrated into an existing PT scheme. This paper aims to fill this gap by identifying potential temporal and weather-related determinants for SBRT-rentals of the Dutch SBRT-system using multiple linear regression (MLR) and an in-depth analysis for selected stations. The results are compared with the findings of one-way bikesharing schemes. The results show that for hourly rentals in an SBRT-system, the highest explanatory power is attributed to the number of train travelers leaving the corresponding train station, followed by temporal and weather-related determinants. Furthermore, the magnitude of the correlation between the determinants and the hourly demand varies considerably across stations, depending on the underlying demand patterns.

Introduction

Urban areas all around the world face the challenge of a growing population, leading to increased traffic demand resulting in negative external effects such as road congestion and greenhouse-gas emissions. One way to reduce the external effects caused by road traffic is by increasing the attractiveness of car-independent multimodal trips chains. These allow individuals to shift away from car usage towards alternative, resource-efficient modes of transportation. Multimodal trips often consist of one main mode and different modes used for the so-called first and last mile (sometimes referred to as access and egress leg, respectively) to connect the main mode with the travelers' origin and destination.

The global rise of one-way bikesharing throughout the last decade has also led to an increase in the accessibility of data for investigating their usage. Multiple studies conduct an analysis to identify potential determinants for the usage of bikesharing schemes, with these findings being summarized within multiple reviews (Gu et al., 2019; Médard de Chardon, 2019; Shui and Szeto, 2020; Todd et al., 2021; Eren and Uz, 2020). In contrast, little research has been conducted insofar for identifying the determinants of round-trip bikesharing usage (Nello-Deakin and Brömmelstroet, 2021). This might, arguably, be attributed to the limited availability of these schemes. To the best of our knowledge, only two station-based round-trip bikesharing (SBRT) systems exist, which allow users to do round-trip bookings to get around at their public transport (PT) trip destination: OV-fiets in the Netherlands and Bluebike in Belgium (Villwock-Witte and van Grol, 2015; de Visser, 2017). Unlike one-way bikesharing, round-trip systems provide users with the certainty of having a bike available for their return trip, since the bike rented remains available exclusively for the user who rented it until its return to the origin. Consequently, there is a lack of knowledge of the determinants of demand for these SBRT-systems and the underlying usage patterns. To highlight the difference between one-way and round-trip bikesharing, a visualization of the different systems' functionality is provided in Fig. 1 based on research conducted by van Waes et al. (2018).

We address the knowledge gap on round-trip bikesharing by providing insights into the usage patterns of the SBRT-system known as OV-fiets in the Netherlands, a system which pre-COVID experiences high annual growth, from 1.5 million rentals in 2014 to 5.3 million rentals in 2019, of which around 96 % were returned at the station where they were rented out. Arguably, the high fee for returning one's bike at another station, an additional fee of 10€, refrains users from doing so. Even though PT ridership did not yet return to pre-COVID

* Corresponding author.
E-mail address: o.cats@tudelft.nl (O. Cats).

https://doi.org/10.1016/j.jpubtr.2023.100048
Received 22 July 2022; Received in revised form 10 April 2023; Accepted 26 June 2023
Available online xxxx

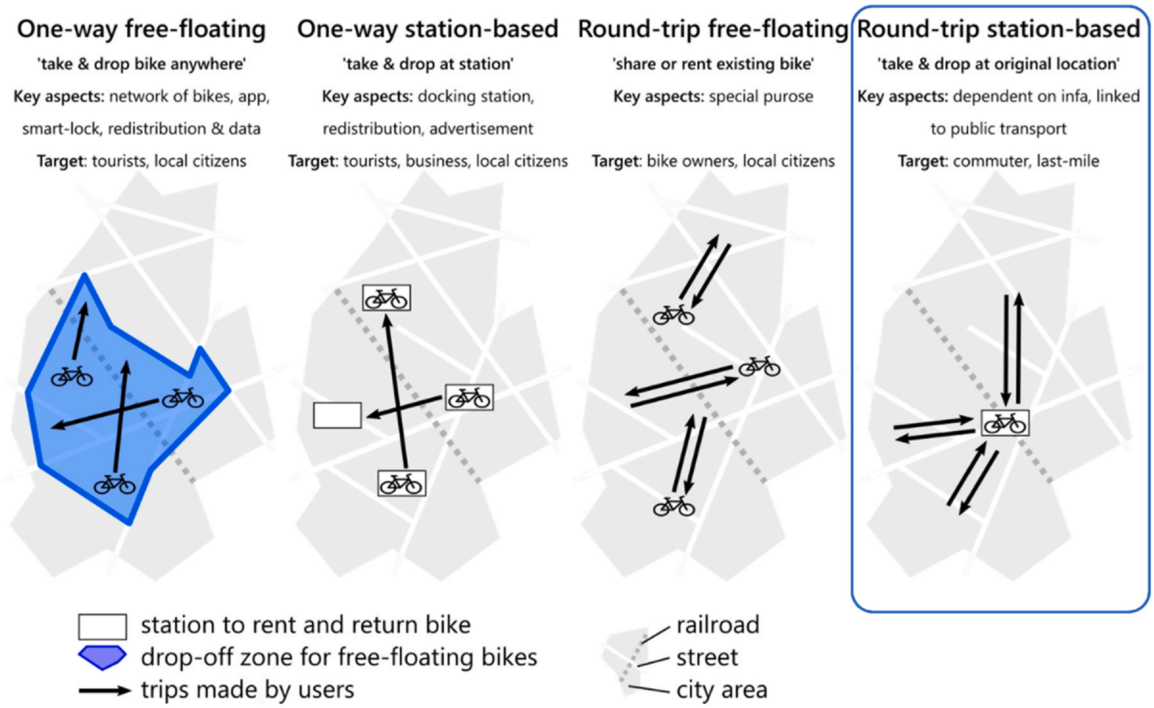

Fig. 1. Visualization based on the Bikesharing typology described by van Waes et al. (2018). (created by the authors).

levels, recently (April 2022) the rentals of the SBRT-system already exceeded the rentals in the same month in 2019 by 10 %. Different from conventional one-way bikesharing-schemes, OV-fiets charges usage on a 24 h flat-fee basis with a possible extension of up to 72 h. At the time this research was conducted, the price was 3.85€ for a 24-h-period. Within the system, The systems' booking data is complemented with additional data sources such as historical passenger flows leaving the train stations next to the considered SBRT-stations and historical weather data to identify potential determinants for the number of bikes rented per hour. We perform the identification of determinants of SBRT rentals on an hourly level of aggregation using multiple linear regression (MLR) and descriptive analytics. In our analysis, we focus on potential weather-related, train traveler-related and temporal determinants. We then compare our findings to previously reported results for one-way bikesharing schemes.

The remainder of this paper is organized as follows: the next subsection identifies potential weather-, time-, and train travel-related determinants for bikesharing demand and contemplates their relevance for SBRT. Next, we elaborate on the data collection and the methods used for the analysis of SBRT demand, followed by presentation and discussion of the results of the performed MLRs and the descriptive analysis. The last two sections discuss and conclude the findings of this study and provide recommendations for future research.

Related work on the determinants for bikesharing demand

Many weather-related determinants were found in several studies to have an impact on one-way bikesharing demand, see the review by Eren and Uz (2020). For one-way bikesharing, it is found that sunny weather results in a higher usage, while rain and wind have a negative impact on hourly rentals. Also, individuals tend to use one-way bikesharing more in moderate temperatures between 0 °C and 30 °C. The usage is identified to be highest between 20 °C and 30 °C, while scorching heat and temperatures below 0 °C are found to have a negative correlation with the number of rented bikes (Eren and Uz, 2020). Recent research also found that in cities with a higher share of young people and a high-quality cycling network, bike usage in general is more robust to unfavorable weather conditions (Goldmann and Wessel, 2021).

In terms of temporal differences, the usage of one-way bikesharing-schemes differs across seasons, with a higher usage in summer than in winter (Eren and Uz, 2020). Most studies investigating the usage throughout the week for specific systems see a clear difference in usage patterns between weekdays and the weekend (Gu et al., 2019; O'Brien et al., 2014). The findings are confirmed by a cluster-analysis performed including 322 station-based free-floating systems (Todd et al., 2021). According to the authors, the distribution throughout the day differs slightly between systems (e.g., different starting time of the morning peak), but recurring patterns can be identified such as distinct morning and evening peaks on weekdays and a moderate usage during afternoons on weekends. Furthermore, it is found that during peak hours bikesharing is more competitive to cars in terms of travel time due to congestion, making the modal shift towards cycling more attractive (Jensen et al., 2010). It is unclear to what extent the described determinants for one-way bikesharing also hold for SBRT systems. The temporal use might differ from one-way schemes, as users do not end a booking after reaching their destination. Instead, their booking continues until returning the rented bike at the station.

Regarding the integration of bikesharing into existing PT services, one-way bikesharing is found to be often used to substitute PT trips involving transfers (Leth et al., 2017). The proper integration of one-way bikesharing into the existing PT network is found to reduce the extent to which both modes compete and increase the added value of both modes for travelers (Böcker et al., 2020). This is the case especially for longer trips (Eren and Uz, 2020) and in times of reduced PT services, i.e. at night and on weekends (Fishman et al., 2013). Further, the added value of round-trip bikesharing lies in the egress leg after traveling by PT as it allows users to cover a longer distance compared to walking, while also allowing to reach destinations which might have limited accessibility by PT (Kager and Harms, 2017).

While determinants of one-way bikesharing demand have been thoroughly investigated in the literature, little is known about the determinants of SBRT-demand. Preliminary conclusions made for one-way schemes might be applicable for SBRT, but there is no scientific evidence supporting this to date. It is likely that the use case of SBRT differs from one-way schemes due to the requirement of ending a booking at the point where it started. This makes the SBRT especially suitable for the activity-end of multimodal trips (Kager and Harms, 2017). This is supported by the rising number of rentals in existing operational systems (de Visser, 2017; NS, 2021).

Method

Data preparation

Data is provided by the Dutch national train station operator, NS Stations. Weather data is extracted from the website of the Dutch Royal Meteorological Institute, KNMI, for multiple weather measurement stations across the country. The national holiday calendars are used to include the national and school holidays. The provided SBRT dataset contains all individual bookings having the same origin and destination. The rentals included in the dataset are cleaned and aggregated on an hourly level per station for the entire year of 2018. The resulting data is combined with a static dataset containing information per SBRT station on the related capacity, the corresponding train station, and the provided service type. For further analysis, only staffed stations with a capacity of more than ten bikes are included, resulting in the inclusion of 48 stations. Also, information on the hourly number of travelers finishing their train journey at the corresponding train station is included. This information is based on the nationwide smartcard check-in/-out system. Furthermore, the school holidays in all three independent holidays regions (North, Middle, South) are included for all stations as SBRT users might use the service for their last mile in regions different from the one they live in. Lastly, the SBRT stations are connected to the closest KNMI weather stations (total of 17 in the study area) to obtain weather-related information. A larger distance between weather stations and SBRT stations may result in lower accuracy on determinants such as rain duration per hour. The final filtered dataset contains 2,646,657 bookings across 48 SBRT stations. These are 75.5 % of all bookings performed in 2018 at 15 % of all stations, indicating that the vast majority of SBRT rentals are retained in the analysis.

Definition of determinants

In previous research, meteorological and temporal factors are found to explain most variance in hourly one-way bikesharing rentals (Du et al., 2020). Other factors, such as a SBRT station's accessibility, the surrounding cycling infrastructure, topography, and land use are defined by the circumstances in which a SBRT-station is located and thus are considered out of scope. The number of hourly travelers leaving the corresponding train station is included as the analyzed SBRT system is integrated into the national train system. This allows for assessing whether hourly SBRT rentals depend on the number of train travelers leaving the corresponding train station, as identified for one-way bikesharing (Zhang et al., 2018). The specific variables used to represent the determinants are summarized in Fig. 2.

For time-related variables, we decided to represent each characteristic with a separate dummy-variable to independently assess their explanatory power. To reduce the number of the resulting dummy-variables, an aggregated representation of the temporal determinants is added: Hours are aggregated into five time-of-day periods (namely Night, Morning peak (7:00–9:00), Daytime, Evening peak (17:00–19:00), Evening). Weekdays are represented using a single dummy-variable indicating whether it is weekend or a weekday, and the months are aggregated on a seasonal basis (Spring, Summer, Autumn, Winter). Holidays, national and school in the three holiday regions, are represented using one dummy-variable each. In total, 40 temporal variables on a non-aggregated level, and eight on an aggregated level are defined.

As for weather-related determinants, multiple determinants are selected for further analysis: Windspeed, Temperature, Sunshine Duration, and Rain Duration. These determinants are translated into variables using the interval scales defined by the KNMI: Windspeed and Temperature are assessed using averages for the last hour in 0.1 m/s and 0.1 °C, respectively. Rain and Sunshine Duration are indicated based on their occurrence, measured in tenths of an hour. Additionally, dummy variables are included indicating whether Rain, Fog, Snow, Thunder, or Ice occurred for any given hour. Together, this results in nine different weather-related variables. In the following we assume that the choice for a bike is based on the weather during that hour, leaving out the potential impact of weather forecasts for later hours.

Lastly, the number of train travelers leaving a train station is used to assess its explanatory power on the hourly rentals of a SBRT-system. This determinant is included as a nominal variable, i.e., the number of checkouts per hour. In addition, to capture the role of a station within it's surrounding environment and to represent it's function in the national train network as a whole, the internal station-category-system used by the national train operator is used, dividing train stations into six categories: A very large station in the middle of a big city (KIS 1), a large station in the middle of a medium-sized city (KIS 2), a suburban station with a transfer role (KIS 3), a station next to the middle of a small town/village (KIS 4), a suburban station without a transfer role (KIS 5), a station in peripheral location to a small town/village (KIS 6) (Mark van Hagen). This characteristic is represented in the dataset using dummy variables. Further, to test for regional differences throughout the country, a dummy variable is added to the dataset for the different regions the stations are located in.

In total there are 50 different independent variables to assess their explanatory power regarding hourly SBRT-rentals. This number is reduced to eighteen when using aggregated time-related variables. Both approaches are used to assess which variables can best capture the observed variance in the number of hourly rentals.

Identification of significant determinants

To assess the explanatory power of the different variables defined above, we estimate a Multiple Linear Regression (MLR) model using the hourly rentals across all stations as the dependent variable and the previously defined determinants as independent variables. The

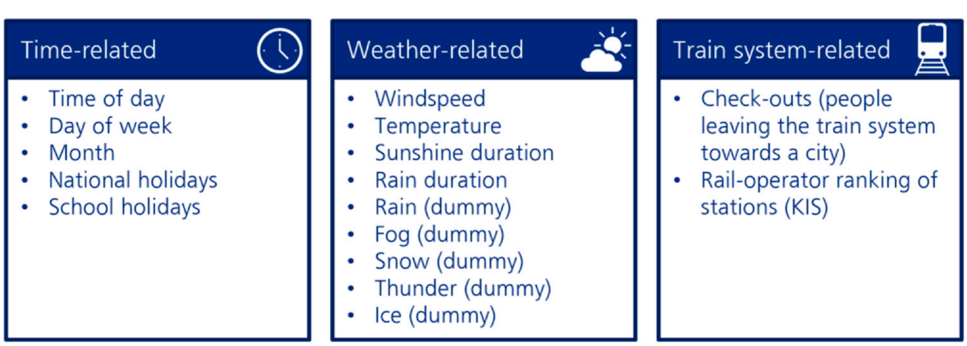

Fig. 2. Grouping of determinants considered for analysis (*exist on an aggregate and disaggregate level).

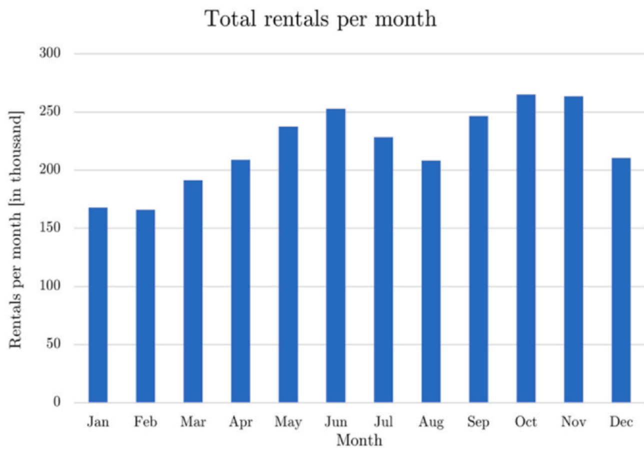

Fig. 3. Total of rentals per month.

mathematical formulation of an MLR is as follows:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n + \varepsilon$$

In this formula, Y represents the dependent variable - in this case the number of SBRT-rentals per hour- while X_1 to X_n represent the independent variables described in the previous section, and β_0 to β_n the weights of the corresponding independent variables, with β_0 being a constant. n is the total number of considered independent variables, while ε is the error-term capturing noise unexplainable by the explanatory variables. The implementation is roughly following the work on determinants for one-way bikesharing schemes performed by Feng and Wang (2017).

To identify the most significant determinants, multiple backward searches are applied using the loss of R^2 as a performance indicator to identify the most significant determinants. This estimation procedure starts with a regression model including all considered variables, and then iteratively removes the least significant variable one at a time based on a pre-defined criterion. The latter is defined here as the loss of R^2 caused by removing the variable. Since the data set consists of a large sample size, the risk of unstable variable selection is considered negligible (Steyerberg et al., 2001). The R^2-loss is used as an indicator as a high R^2-value suggests that the selected variables can explain most noise within the hourly bookings throughout the assessed dataset (Miles, 2005).

The variables contributing to a change of R^2 higher than 0.001 within the backward search application are selected for further analysis at the station level. Station-specific MLRs are performed to examine whether differences exist in terms of the identified determinants' ability to explain the variance in the dependent variables for different stations. This is done as stations are found to have distinctively different usage patterns when it comes to the bike-train combination (Nello-Deakin and Brömmelstroet, 2021; Schakenbos and Ton, 2021).

The results of the station specific MLRs are compared using the number of significant variables per station and the resulting R^2-value from each MLR-application is used to examine to what extent the noise in the data can be explained by using the identified variables.

Descriptive analysis

Based on the application of the general MLR-model for each individual station separately, eight exemplary stations are selected and investigated in greater detail to further investigate their demand determinants. The station selection is based on the distribution of the R^2-values of the general model's performance for each station. The selected exemplary stations include two stations with a low and a high remaining noise, selected using the highest and lowest R^2-value across all stations, respectively. Additionally, the stations being closest to the mean, the median, as well as the 25 % and 75 % quantile of the corresponding R^2-value are selected to provide a wide range of exemplary stations. The subsequent descriptive analysis is performed to understand why the general model is performing better for some and worse for other stations, and thereby investigate the underlying reasons for the large differences observed in terms of model performance. The analysis covers a small subset of the 48 stations included in the first analysis. However, it is deemed sufficient to provide first insights into the usage patterns and reduces complexity of the research. The selected stations are then compared using a visual representation of the average hourly rentals across days and weeks in combination with the identified determinants. The aim of this descriptive analysis is to assess whether the determinants have a similar impact across different stations, or whether the patterns differ to the extent that no overarching findings can be concluded.

Results

Aggregate demand analysis

In the following, an aggregate analysis of the dataset is conducted to provide an understanding of the data before being analyzed in greater detail in the upcoming sections. In Fig. 3 the overall number of rentals per month is provided. As can be observed, the lowest number of rentals was recorded in the beginning of the year 2018 in January and March, with two demand peaks in the months of June and October/November, and a slight drop between these peaks during the summer holiday months July and August. When splitting the total number of rentals on a weekday level as shown in Fig. 4, it becomes visible that rentals on weekdays (Monday–Friday) exceed the demand on weekends (Saturday & Sunday), with Sunday with 3588 average daily rentals having fewer than half the number of rentals of the high-demand weekdays Tuesday, Thursday and Friday with 8025, 8930, and 8860 average daily rentals, respectively. It should be noted that there is a high variation in the number of rentals per weekday throughout the year, with this variability being higher on weekdays than on weekends.

When investigating the data on an average hourly basis for an average day, grouped by the day being during the week or on a weekend, as shown in Fig. 5, two contradicting patterns can be observed: While the no-weekend pattern provides a distinct morning peak between 7 a.m. and 9 a.m. as well as a smaller peak around 6 p.m., the weekend pattern shows a less distinct peak around midday (12 a.m. to 1 p.m.), and a generally lower level of average hourly rentals. Additionally, to understand the potential impact of the defined determinants on the number of rentals, in Fig. 6 the average hourly rentals are shown in relation to the temperature measured at the weather station being closest to each OV-fiets station in degrees Celsius [°C]. It can be

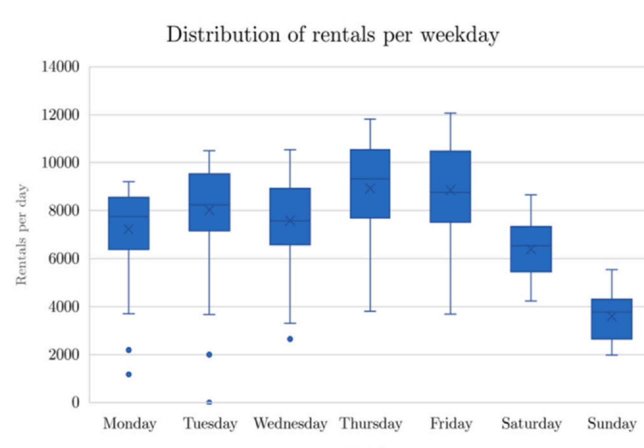

Fig. 4. Distribution of rentals per weekday.

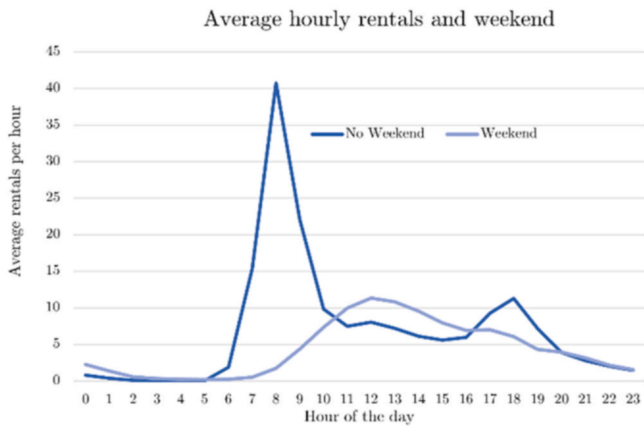

Fig. 5. Average hourly rentals, separated by weekend/non-weekend.

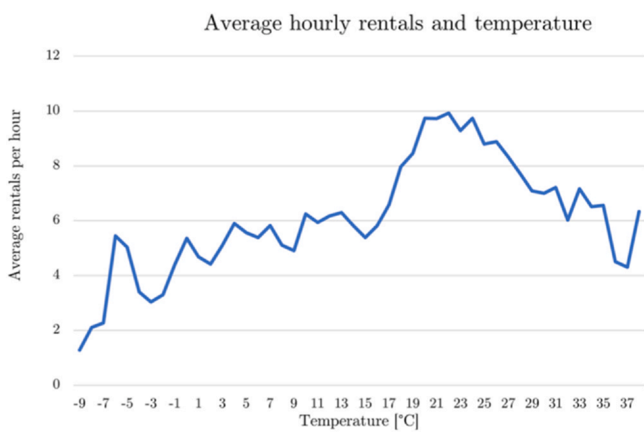

Fig. 6. Average hourly rentals related to the temperature in that hour [°C].

seen that the number of average hourly rentals is highest when the measured temperature was between + 19 °C and + 27 °C, with a moderate decline towards lower as well as higher temperatures.

While the provided aggregate analysis of the dataset offers some initial insights concerning potential determinants of the demand for this SBRT-system, in the following we systematically identify the underlying determinants by means of a statistics analysis.

Identification of significant determinants

Before conducting the backward searches, it is important to assess which interaction effects to include in the analysis. To assess this, a test for correlations among the different variables is conducted, with the results being shown in Fig. 7. As can be seen, the weather-related variables dewpoint and temperature are highly correlated. Other correlations are found between sunshine duration, relative humidity, cloud coverage, and rain duration. As these weather-related determinants are likely to have a different impact on the hourly rentals in different seasons of the year, these are included in the aggregated backward search with an interaction with the time-related determinant season. At the temporal level, interactions between the number of checkouts and hour/time of day, weekday/weekend, and the four different holiday types are included as it is expected that the number of travelers in the corresponding PT system differs across these different temporal variables, which is expected to have an impact on the number of rentals in the SBRT-system in line with past empirical findings (Zhang et al., 2018).

Then, the backward search algorithms, using the previously described selection of variables, are applied on the dataset to identify the most significant variables. Multiple backward searches are performed to include, in separate iterations, the aggregated and disaggregated variables. This is done to identify whether the more general, aggregated version might be sufficient to represent a determinant. The results for the different backward searches are visualized in Fig. 8, with each bar indicating the magnitude of drop in R^2 caused by the backward search removing the corresponding variable.

The number of hourly checkouts is found to have the highest explanatory power across all backward searches when determining the relative number of rentals per station, as 19 % of the variance in the hourly rentals across all stations can be explained using this variable. Furthermore, the multiple time-related variables are found to together explain 22 % of the variance. The explanatory power is mostly covered by the disaggregated Hours 8–18 and the aggregated variables Morning peak, Evening, Evening peak, and Night. Other considered variables are the time-related variables Saturday, Sunday, and weekends, respectively, as well as both national and school holidays. The weather-related variables are found to explain about 5 % of the variance in the data, while the only weather-related variable resulting in a change of R^2 of more than 0.001 is the sunshine duration.

All variables causing a drop of R^2 of more than 0.001 across all four different backward search applications are displayed in Fig. 9. Both the regional and the station-type related dummy-variables were not found to result in a drop of more than 0.001 R^2. The color of the variables indicates whether the correlation with the hourly bookings is positive or negative in relation to the reference variable. For example, when looking at the time-related variables at a disaggregate level, the hourly rentals in Hour 7–19 show a positive significant difference from the reference Hour 0. When aggregated, the reference variable is Daytime, to which hours in Morning peak show a significantly higher number of rentals compared to the reference. The other three, Evening, Evening peak, and Night are associated with significantly fewer hourly rentals. The interaction effects should be read as a combination of two variables. For example, the negative significant interaction between Morning peak and Weekend indicates that during morning peaks in the weekend fewer bikes are rented out per hour in comparison to morning peaks during the week.

Performance of variables across all station-specific MLRs

While the findings are identified on a dataset across all stations providing insights on the general tendencies of the determinants, they provide limited information on which determinants can explain the variance of the hourly rentals at the individual station level. To assess their performance per station, station specific MLRs are performed using the previously identified significant variables. 48 separate MLRs are performed using the defined variables, one per station. For each station, the reference variables for the dummy variables consist of the time-of-day Hour 0 (midnight), the season Autumn, and a day being not on a Weekend. The results are shown in Fig. 10, showing the number of station specific MLR models (out of the total of 48) for which a certain significance level has been attained for each of the explanatory variables considered. For example, the determinant Weekend has a significance level $p < 0.0001$ for 27 out of 48 stations, while for two stations it has a significance level $0.001 \le p < 0.001$. Moreover, for 10 of the 48 stations the significance level for the Weekend is $p > 0.1$. In the case of the interaction variables between checkouts and hour during night-time (1–6 a.m.), the total number of observed determinants being present in the MLR-results is lower than 48. This is a result of the fact that for some stations no check-ins and -outs are registered during these hours.

For train traveler-related determinants, the main effect of checkouts at the respective station has a high significance level only for few stations. However, many stations show interaction variables including checkouts being highly significant.

Regarding weather-related determinants, sunshine duration is the only weather-related variable which is significant on a 95 %-level (i.e.,

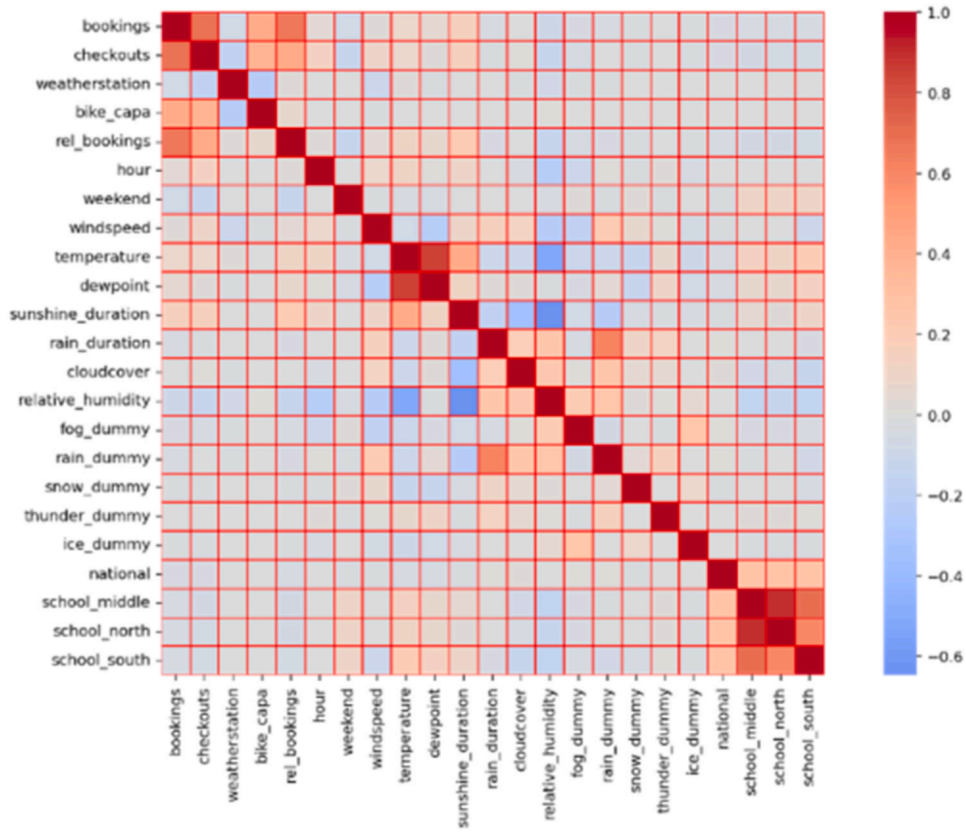

Fig. 7. Indication of correlations (red indicates positive, blue negative correlation).

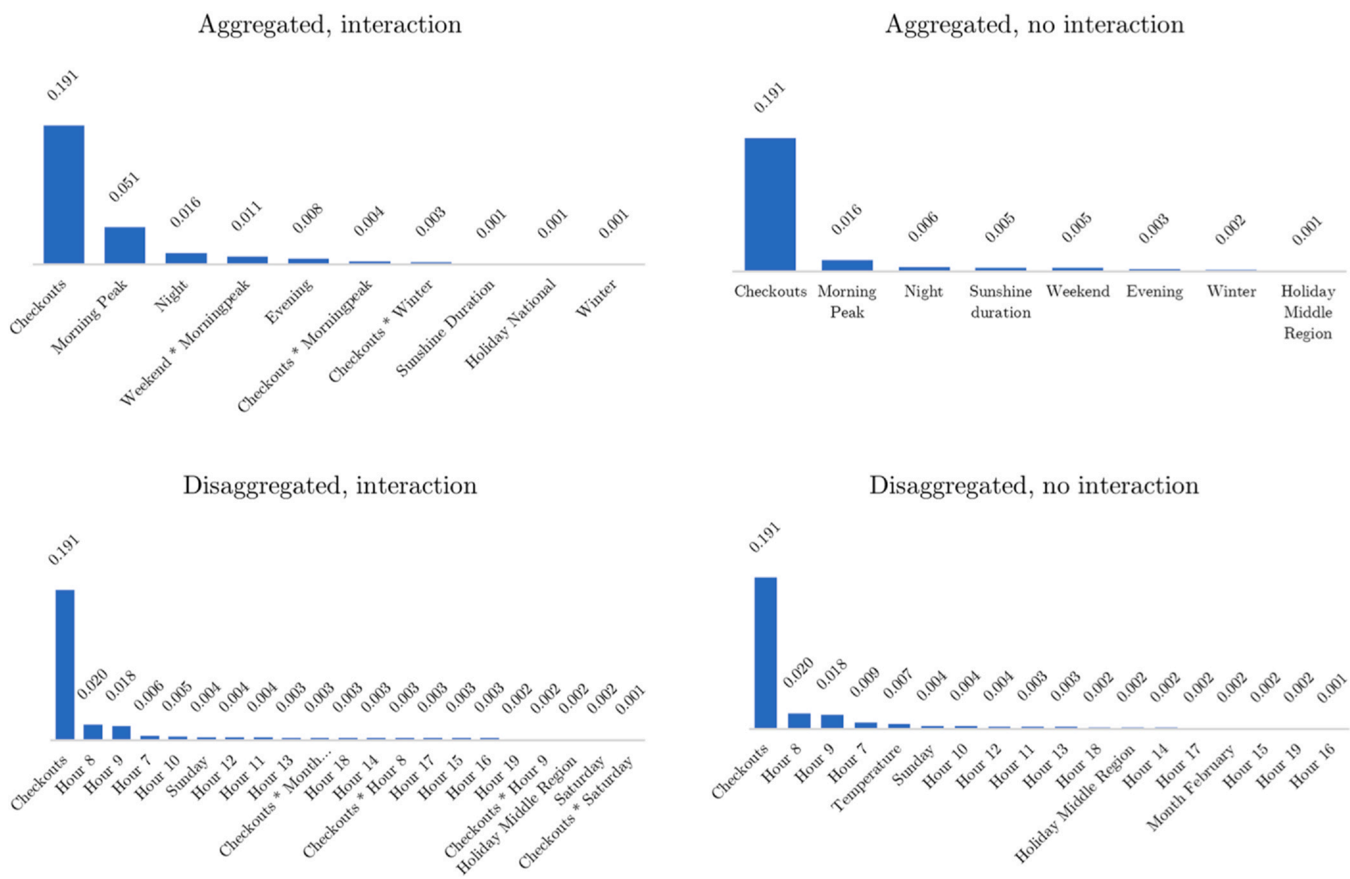

Fig. 8. Backward search – change in R^2 per removed variable for the four different searches applied on different sets of variables.

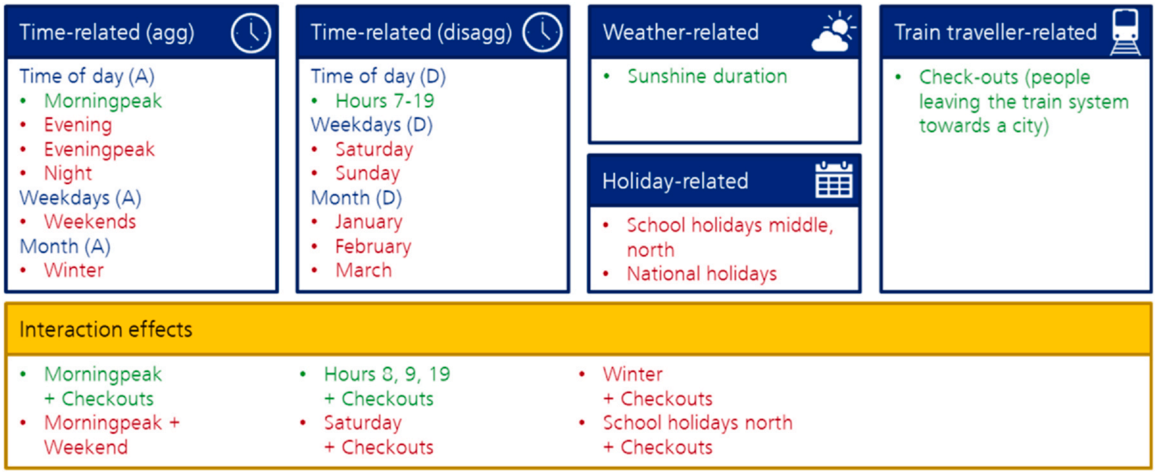

Fig. 9. Indication of correlations between the different variables (green indicates a positive correlation, red indicates a negative correlation).

a p-value below 0.05) across 44 % of the MLRs. A similar result is found for the interaction between sunshine duration and checkouts (50 % of MLRs on 95 %-level). The interaction variables representing sunshine duration and seasons are found to have a high significance across 36 % and 39 % of all stations for Spring and Summer at the 95 %-significance level, respectively. This implies that there are fewer differences in terms of hourly rentals when interacting with sunshine duration during the spring months compared to autumn, i.e., the reference level. During the Winter months, 54 % of all stations indicate the related interaction variable to be significant at the 95 % significance level compared to the reference level.

When investigating time-related determinants such as hour-of-day and the interaction between hour-of-day and checkouts, it can be observed that the timeslots of Hours 22, 23, and 1–5 are insignificant even at the 90 %-significance level. This is arguably because of the limited variability observed during these time periods due to the limited number of rentals in these hours. In addition, for some stations no interaction variables could be assessed for the timeslots between 1 and 6 a.m., as either no checkout and/or no SBRT-booking data is available for these timeslots. Many of the relevant facilities are closed during that time. The timeslots in the morning peak (7:00–9:00) show a high significance among most stations when interacting with checkouts in that timeslot compared to their independent counterparts. The opposite effect can be seen for Hours during Daytime (12:00–18:00), which are mostly significant when included as an independent variable and thus seem to be less explainable by the interaction with Checkouts. An exception can be seen for the evening timeslots (18:00–20:00), where up to 42 % of the stations exhibit a high significance of interaction variables with Checkouts. Weekends have a significant correlation with the hourly rentals for most stations, with 75 % of the stations having a significance level > 95 %, and 56 % even higher than 99.99 %.

The seasonal variables are found to be significant for a few stations when considered separately but especially when interacting with sunshine duration and checkouts (for the interaction with the sunshine duration, see above). The interaction with Checkouts is prominent for Winter, as for 92 % of the stations this variable is significant at a > 95 %-level. Lastly, the variables representing the national and school holidays are found to be significant at the 95 %-level for at least nine of the stations, but none of the variables is significant for more than 50 % of the stations. Also, the interaction variables combining National Holiday and Seasons are found to be insignificant at the 95 %-level for at least 83 % of the stations, suggesting that the presence of holidays is only relevant for a small number of stations. There is no interaction variable between summer and national holidays as no national holidays take place in the summer period.

Descriptive analysis

Next, we perform a descriptive analysis of different determinants using selected exemplary stations. This includes a discussion on potential causes when identifying recurring patterns among multiple stations. The exemplary stations are selected based on the performance of the station specific MLRs. Then, the determinants are descriptively analyzed to unravel their potential dependency with the rental patterns, which are aggregated or averaged on a monthly, daily, and hourly basis. The stations selected based on the goodness-of-fit of the station-

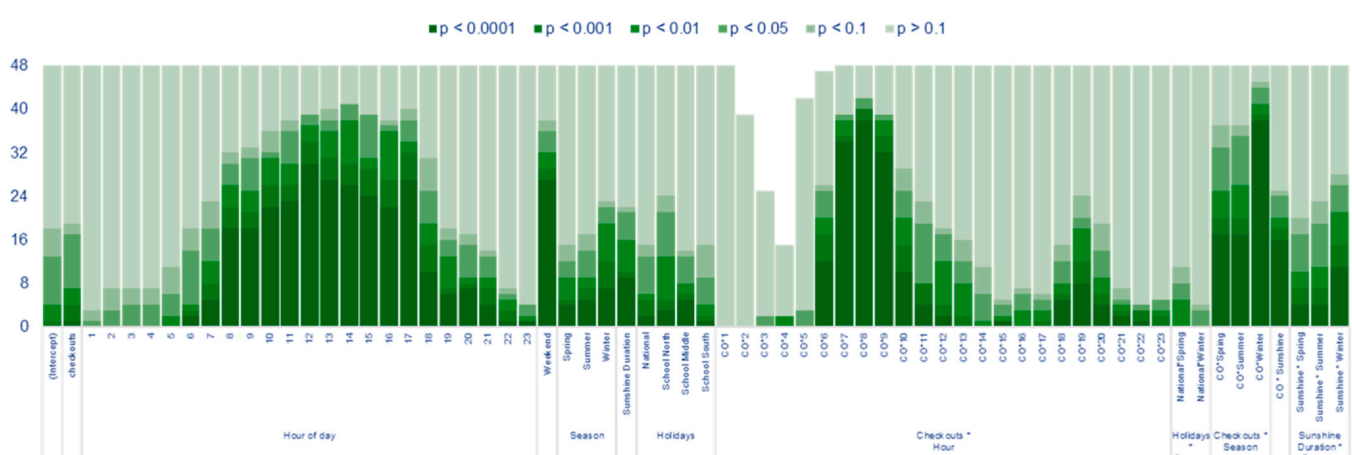

Fig. 10. Number of significant variables across significance levels and stations per variable (reference levels: Autumn, midnight, no weekend).

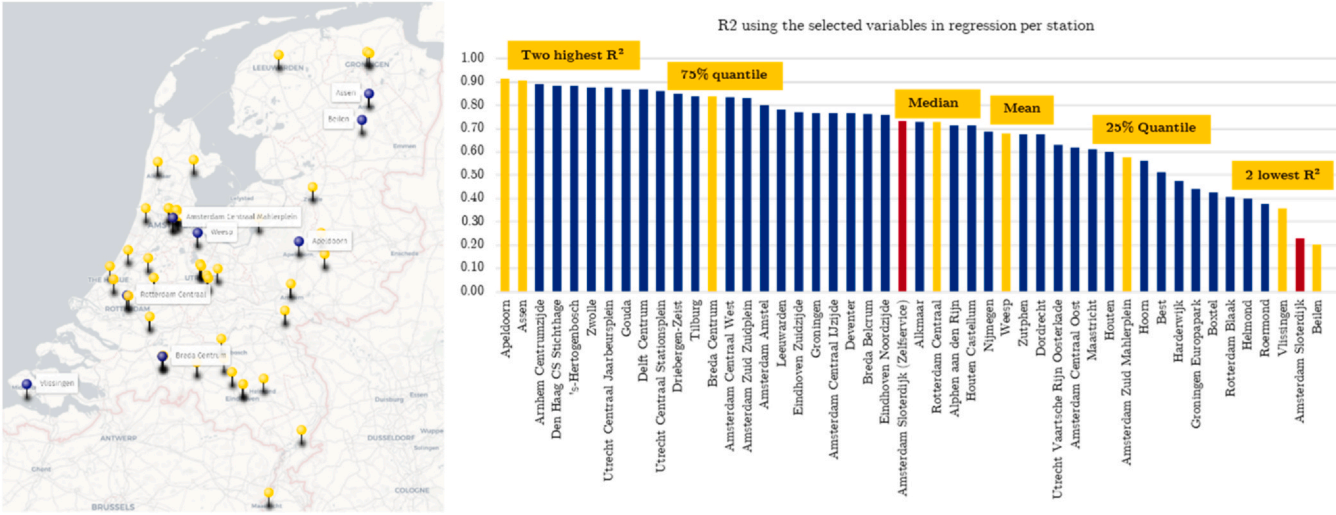

Fig. 11. Considered stations displayed on a map (left) and their R^2-performance using the general MLR-model (right).

specific general MLR-model with the previously identified determinants are *Beilen*, *Vlissingen*, *Weesp*, *Rotterdam Centraal*, *Amsterdam Zuid Mahlerplein*, *Breda Centrum*, *Assen*, and *Apeldoorn* (Fig. 11). The selection of determinants considered for comparison differs per level of aggregation: the aggregated rentals and checkouts are compared on a monthly and daily level, whereas further time- and weather-related variables are analyzed on an hourly level to prevent loss of information for the latter.

First, the monthly, daily, and hourly levels are compared to identify recurring patterns across multiple stations. The aim is to investigate whether usage patterns are similar enough to allow for a generalization of the previous findings across multiple stations. If it is found that the patterns are unique per station across multiple variables, then distinct models per station are required. The interpretation of the differences amongst stations were discussed with and confirmed by individuals working for the operational department of the SBRT-scheme within the Dutch Railways. The performed MLRs provide insights into these causations allowing for a visual high-level analysis yet should be interpreted with caution.

Monthly patterns

When comparing the distribution of rentals per month (see blue lines in Fig. 12), all selected stations apart from *Beilen* and *Vlissingen* show an increase in rentals throughout the first half of the year, followed by a decline in July and August in which the school summer holidays take place. This is in line with the decrease in the total number of checkouts at the corresponding train stations, visualized by the red lines in Fig. 12. In Autumn, the number of rentals rises again, in line with the increasing number of checkouts (with *Weesp* being an exception). This confirms the previous finding that checkouts can explain a high share of the variance in the hourly rentals.

The patterns of the other two stations, *Beilen* and *Vlissingen*, show limited similarity with the other stations. Due to the low amount of bike rentals (on average 3–4 a day) there is too much noise in the data in the case of *Beilen* to allow examining patterns. For *Vlissingen*, the station being located close to outdoor recreation areas might suggest a higher usage throughout summer compared to winter. To conclude, the patterns of the stations themselves and in combination with the monthly checkouts provide little potential for generalization.

Daily patterns

To compare the average number of rentals per weekday throughout the week, two different patterns occur at multiple stations, as visualized in Fig. 13 for three of the exemplary stations selected for illustration purposes. Among the exemplary stations, the first pattern (*Beilen*, *Weesp*, *Breda Centrum*, *Assen*, *Apeldoorn*) follows a stable level of rentals throughout working days Monday to Thursday, with a small drop on Wednesdays and a sharp decrease from Friday to Sunday. The checkouts of these stations follow a similar pattern, suggesting that the lower numbers of rentals on Wednesdays and Fridays might be caused by fewer commuters on these days, which use the SBRT to perform the trip between workplace and train station. *Vlissingen* follows a similar tendency, but the high width of the confidence interval does not allow for an interpretation of an explicit pattern.

The second pattern occurs at both *Rotterdam Centraal* and *Amsterdam Zuid Mahlerplein*, showing an increase in rentals from Monday to Friday followed by a sharp decrease towards the end of the week. When comparing these daily rentals with the daily checkouts, checkouts show a similar pattern across all selected stations, with a stable level throughout the week and a drop towards the weekend. The difference between the two patterns might thus be reasoned in location-specific characteristics of the stations. For example, *Amsterdam Zuid Mahlerplein* and *Rotterdam Centraal* are located in the two biggest Dutch cities which attract both tourists and nightlife visitors using bikes to reach their destination (Jonkeren et al., 2021). An investigation of hourly rentals and the comparison between weekends and weekdays may provide additional insights.

Hourly patterns

Two recurring patterns become visible when analyzing the average hourly rentals per day, with *Vlissingen* being an exception (see Fig. 14 for three examples): While *Weesp*, *Breda Centrum*, *Assen*, and *Apeldoorn* show a high number of rentals in the morning peak hours, they remain low throughout the rest of the day. This pattern is different from the hourly checkouts throughout the day, which exhibits an increase in the evening peak (16:00–19:00). These evening peak checkouts are train commuters on their way back home. Since the SBRT-system is typically used for the activity-end of a trip and not for the home-end, this explains the difference between rentals and checkouts in the evening.

The high number of SBRT-rentals in the morning peak could be attributed to commuters traveling by train, using the SBRT-system for their last mile to reach their workplace. The second pattern occurs at *Rotterdam Centraal* and *Amsterdam Zuid Mahlerplein* showing a less steep decrease after the morning peak compared to the first pattern. Instead, the number of hourly rentals remains at a high level before displaying a second increase in the evening peak. Remarkably, for these two stations the number of hourly SBRT-rentals closely follows the hourly checkouts pattern. This suggests that SBRT-bikes are rented out for different purposes throughout the day. For example, the evening peak in rentals might be reasoned in a higher attraction of the corresponding cities to

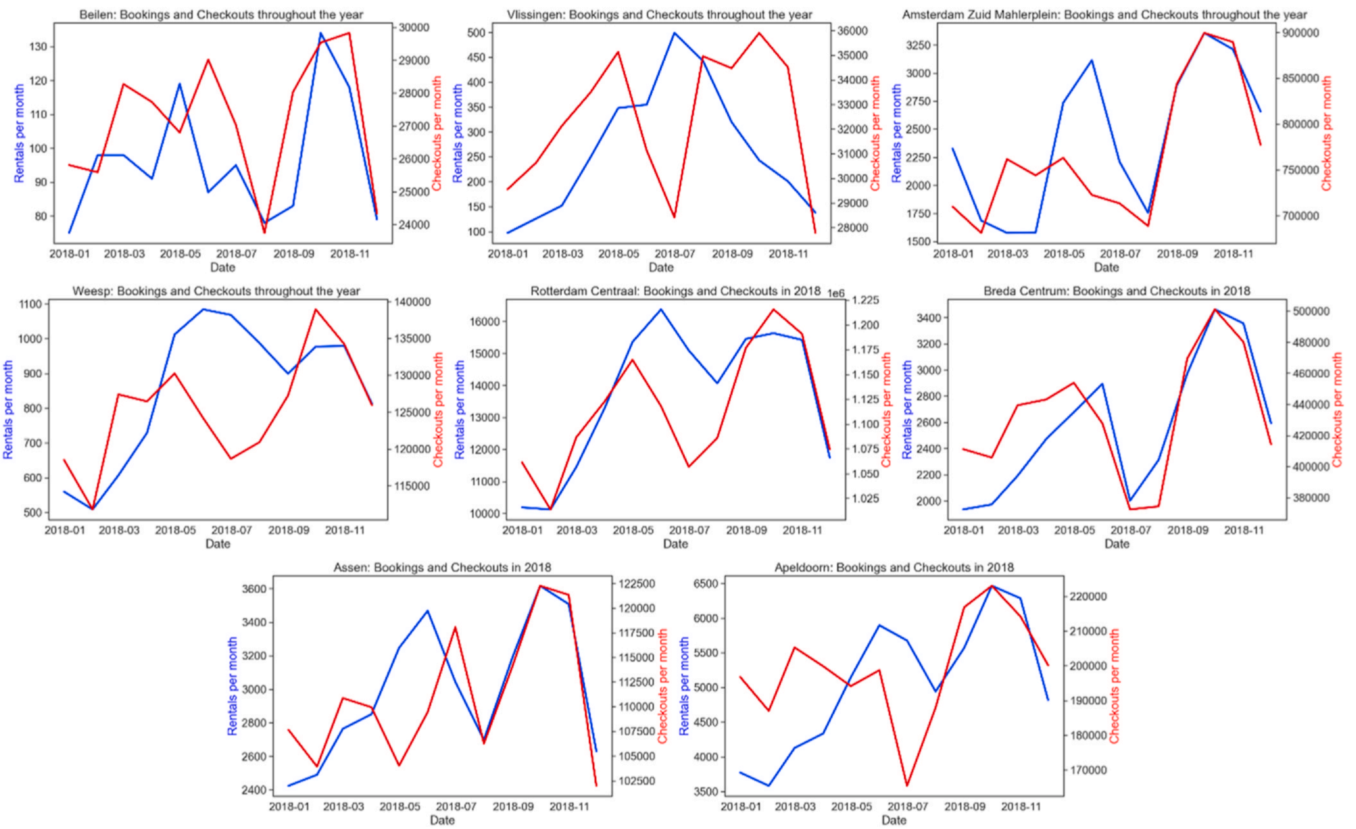

Fig. 12. Aggregated monthly rentals and checkouts in 2018 for the exemplary stations.

serve recreational purposes. This finding would be in line with previous findings on a daily level.

To assess the findings obtained from the daily and weekly patterns, the daily patterns are analyzed based on the time-related determinants 'weekend', 'national holiday', and 'school holiday'. Exemplary results are visualized in Fig. 15. It is found that weekends and national holidays have a similar effect on the daily pattern, with morning and evening peaks being replaced by an increase in rentals around noon and the early afternoon. While this new peak is more elevated for stations located in big cities (*Rotterdam Centraal* and *Amsterdam Zuid Mahlerplein*), for stations located in smaller cities such as *Breda Centrum*, *Apeldoorn*, and *Assen*, the peak is less distinct. *Amsterdam Zuid Mahlerplein*, *Weesp*, *Rotterdam Centraal*, *Breda Centrum*, *Assen*, and *Apeldoorn* show similar effects of the different seasons, with an overall higher level of rentals per hour in Summer and Autumn and a lower level in Winter. An exception, again, is *Vlissingen*. There, in Spring and Summer an increase of rentals can be seen around noon, supporting the interpretation that in warmer seasons people are likely to use the SBRT for recreational purposes. It is found that at *Amsterdam Zuid Mahlerplein*, *Rotterdam Centraal*, and *Vlissingen* the presence of rain has almost no effect on the number of rentals within the morning peak among the selected stations, while leading to a slight decrease in other hours of the day, possibly because of greater demand elasticity. The other exemplary stations show no significant impact of occurring rain at all. This might be attributed to the lack of other options to reach the destination.

In summary, while some similar patterns among some exemplary stations are observed, there are considerable variations in demand determinants across stations. It is therefore advised to apply models separately per station to allow for an appropriate representation of the local context.

Comparison with past results for one-way bikesharing systems

Regarding weather-related determinants, the lack of impact of the occurrence of rain on hourly rentals in the morning peak differs from the negative impact of rain for one-way bikesharing systems (Bean et al., 2021). This might be caused by commuters relying on the SBRT-system for the egress leg of their trip as there might be no or few (likely less attractive) alternatives to reach their destination, making them less sensitive to occurring rain. Additionally, when renting an SBRT-bike the users are assured of also having it available for their return trip to the station. This results in a certainty of availability which differs from one-

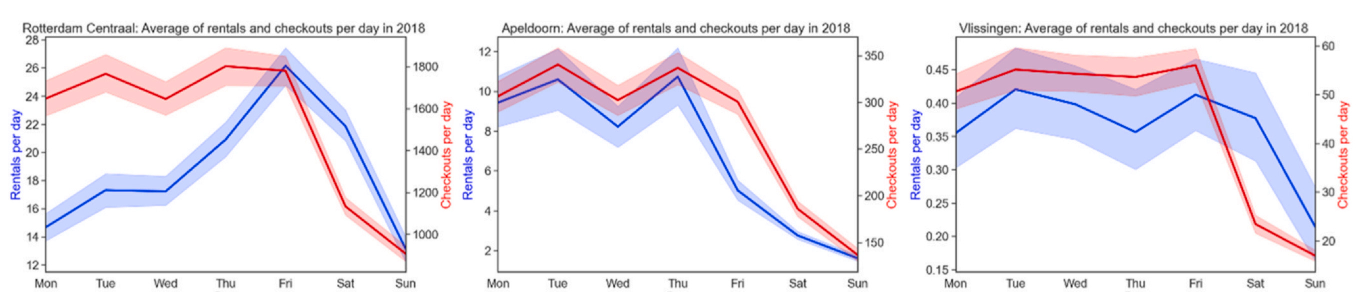

Fig. 13. Average hourly rentals and checkouts per day for Rotterdam Centraal, Apeldoorn, and Vlissingen.

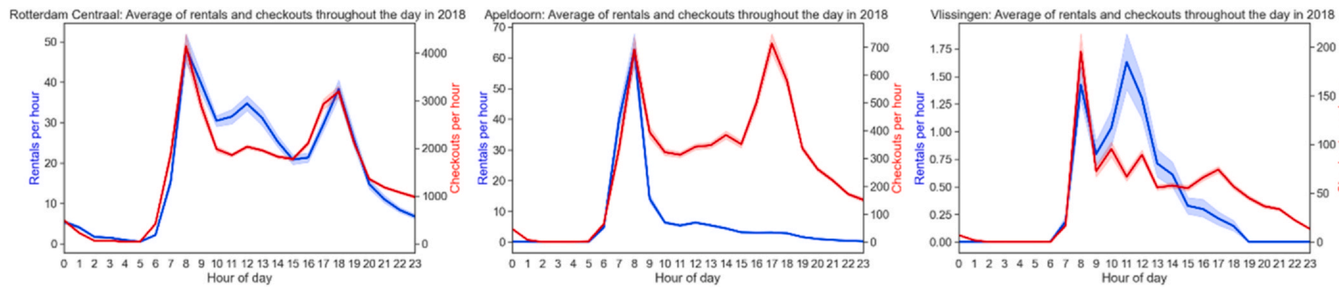

Fig. 14. Aggregated hourly rentals for Rotterdam Centraal, Apeldoorn, and Vlissingen.

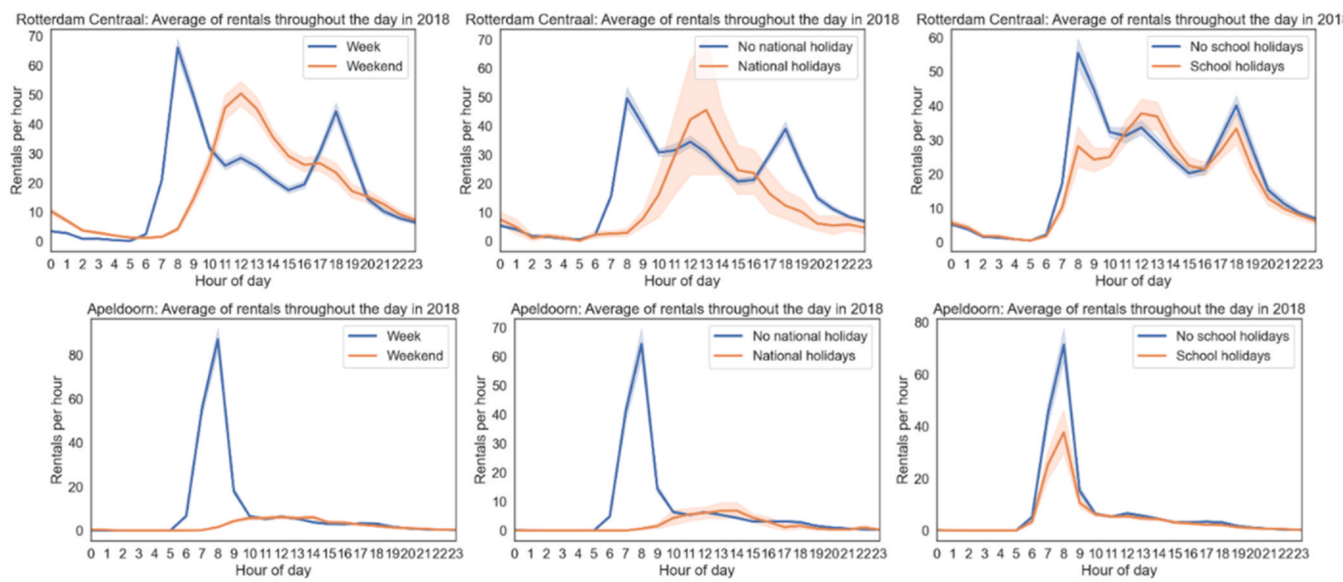

Fig. 15. Aggregated hourly rentals for Rotterdam Centraal and Apeldoorn, compared regarding the related days being on a weekend or a national holiday, or in school holidays.

way bikesharing systems in which users cannot be certain that a bike will be available at a certain time and location when they need it. Regarding the positive correlation between hourly rentals and sunshine duration, this is in line with the gathered literature findings for one-way bikesharing by Eren and Uz (2020). This finding is supported by the national train operator NS using a separate model to forecast train traveler demand for stations with a high recreational attraction in times of sunshine and elevated temperature, especially for destinations close to the beach.

Regarding the time-related determinants, the findings on a monthly aggregation are overall in agreement with previously reported findings (Eren and Uz, 2020). Notwithstanding, the present research identifies the highest number of monthly rentals during autumn, while preliminary one-way bikesharing-related research identified a peak of rentals during summer. This difference can be reasoned in lower numbers of commuters and train travelers during holidays, leading to a lower SBRT-usage. A special case, again, are SBRT-stations located at destinations with a high attraction for recreational trips. For example, in our analysis *Vlissingen*, a coastal retreat destination, has the highest number of rentals in summer, which is more in line with findings for one-way bikesharing.

Regarding rental patterns aggregated per day throughout the week, some of the SBRT-patterns selected for the in-depth analysis are in line with preliminary findings gathered by Todd et al. (2021). Both SBRT- and one-way bikesharing show a higher number of rentals on weekdays compared to weekends. However, one pattern identified among the selected stations has no similar counterpart in one-way bikesharing literature: the peak of rentals at bigger stations such as *Rotterdam Centraal* and *Amsterdam Zuid Mahlerplein* between Thursday and Saturday. This might be caused by the 24-h pricing scheme of OV-fiets making long-term bookings comparatively cheap. Another reason might be the round-trip nature of the system, making it more attractive to book a bike overnight and/or for an entire weekend in comparison to one-way bikesharing.

When comparing the literature findings for the distribution of rentals throughout the day, the distinct morning peak is in line with one-way schemes while the evening peak is less distinct. A potential reason is that individuals renting bikes in the morning have their rented bikes available to return to the station in the evening. In the one-way-schemes discussed in literature, these return-trips are separately booked, leading to distinct evening peaks (Todd et al., 2021). Still, evening peaks exist in the SBRT-system for some stations, but they occur later compared to one-way schemes and only at stations located in bigger cities such as *Rotterdam Centraal* and *Amsterdam Zuid Mahlerplein*. On weekends, the hourly patterns throughout the day show peaks in the early afternoon, which is in line with findings for one-way schemes.

Conclusions

This study analyzed bike rentals at SBRT-stations of the OV-fiets scheme operated by the Dutch Railways for the year of 2018 to identify key determinants of scheme usage. We combine several data sources to perform multiple linear regressions (MLR) across the entire dataset consisting of 48 stations as well as per individual SBRT-station to identify significant weather- and time-related determinants. For some stations using few variables can already achieve a high explanatory power, while for others achieve a lower explanatory power is obtained even when including a larger number of significant variables. Thus, there is no clear set of variables being able to explain variance across the entire set of stations.

To further investigate whether the available data can be used to identify temporal usage similarities and differences among SBRT-stations, a descriptive analysis is conducted for eight selected stations. The hourly rentals per station are then aggregated on a monthly, daily, and hourly level and compared with the previously identified determinants. When comparing the patterns of the different stations, it is found that while the patterns mostly differ across the stations, several general trends can be identified. For example, on average all stations have their highest number of hourly rentals during the morning peak between 7 and 9 a.m., and the two selected SBRT-stations located in bigger cities also experience a second peak in the afternoon between 5 and 7 p.m. The latter suggests a different use case of the SBRT-system in the evening peak compared to the morning peak. Another identified difference becomes visible between the patterns of hourly rentals on weekends and weekdays, as on weekends neither morning nor evening peaks appear. Instead, the rentals either remain at a low level throughout the day or experience a peak during the early afternoon between 12 and 2 p.m. Another finding is that the occurrence of rain is unlikely to impact the number of rentals in the morning peak, while the number of rentals throughout the rest of the day slightly drops when rain occurs.

We compared our results with those reported for one-way bikesharing schemes and found that, while there are determinants which exercise similar effects in both SBRT- and one-way bikesharing like sunshine duration, temperature and time of the year, other determinants show noteworthy differences, such as the higher number of rentals on Fridays or the lack of evening peaks at many SBRT-stations.

Further research might investigate whether the given findings hold during and post-COVID, as changes in mobility usage patterns are reported across the Netherlands and beyond (Ton et al., 2022). Also, given the known limitations of the application of MLR, it will be valuable to conduct a similar study using more advanced methods to identify determinants for SBRT, which might overcome the MLR-limitation of being capable to capture linear dependencies only. Moreover, it can be interesting to capture additional determinants such as events and service disruptions or conduct the in-depth analysis across all stations in the system, as this research provides first insights only.

For operators, the provided information on the determinants of a SBRT-system in combination with suitable demand forecasting methods can potentially allow for an increase in efficiency in terms of staff scheduling, maintenance of bikes, and a potentially higher user satisfaction due to an improved match of supply and demand. Knowing the projected demand per station can help operators relocating bicycles and staff to stations where the projected demand exceeds the current supply or can schedule bicycle maintenance in times of predicted low demand. These improvements in operations are likely to increase the attractiveness of the system for existing and potential users, making it a valuable tool to support bike-train multimodal trip chains.

CRediT authorship contribution statement

The authors confirm contribution to the paper as follows: study conception and design: F. Wilkesmann, D. Ton, R. Schakenbos, O. Cats; data collection: F. Wilkesmann, R. Schakenbos, D. Ton; analysis and interpretation of results: F. Wilkesmann, D. Ton, R. Schakenbos, O. Cats; draft manuscript preparation: F. Wilkesmann, O. Cats, R. Schakenbos, D. Ton. All authors reviewed the results and approved the final version of the manuscript.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to thank the Dutch train station operator NS Stations for providing usage data on their SBRT-System OV-fiets as well as additional information used in this research. The work of the last author was supported by the CriticalMaaS project (No. 804469), which is financed by the European Research Council and Amsterdam Institute for Advanced Metropolitan Solutions.

References

Bean, R., Pojani, D., Corcoran, J., 2021. How does weather affect bikeshare use? A comparative analysis of forty cities across climate zones. J. Transp. Geogr. 95, 103155. https://doi.org/10.1016/j.jtrangeo.2021.103155

Böcker, L., Anderson, E., Uteng, T.P., Throndsen, T., 2020. Bike sharing use in conjunction to public transport: exploring spatiotemporal, age and gender dimensions in Oslo, Norway. Transp. Res. Part A: Policy Pract. 138, 389–401. https://doi.org/10.1016/j.tra.2020.06.009

Du, M., Cao, D., Chen, X., Fan, S., Li, Z., 2020. Short-term demand forecasting of shared bicycles based on long short-term memory neural network model. In: Sun, X., Wang, J., Bertino, E. (Eds.), Artificial Intelligence and Security. Lecture Notes in Computer Science. Springer International Publishing, Cham, pp. 350–361. https://doi.org/10.1007/978-3-030-57884-8_31

Eren, E., Uz, V.E., 2020. A review on bike-sharing: the factors affecting bike-sharing demand. Sustain. Cities Soc. 54, 101882. https://doi.org/10.1016/j.scs.2019.101882

Feng, Y., Wang, S., 2017. A forecast for bicycle rental demand based on random forests and multiple linear regression. In: Proceedings of the 2017 IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS), pp. 101–5. 〈https://doi.org/10.1109/ICIS.2017.7959977〉.

Fishman, E., Washington, S., Haworth, N., 2013. Bike share: a synthesis of the literature. Transp. Rev. 33 (2), 148–165. https://doi.org/10.1080/01441647.2013.775612

Goldmann, K., Wessel, J., 2021. Some people feel the rain, others just get wet: an analysis of regional differences in the effects of weather on cycling. Res. Transp. Bus. Manag. 40, 100541. https://doi.org/10.1016/j.rtbm.2020.100541

Gu, T., Kim, I., Currie, G., 2019. To be or not to be dockless: empirical analysis of dockless bikeshare development in China. Transp. Res. Part A: Policy Pract. 119, 122–147. https://doi.org/10.1016/j.tra.2018.11.007

Jensen, P., Rouquier, J.-B., Ovtracht, N., Robardet, C., 2010. Characterizing the speed and paths of shared bicycle use in Lyon. Transp. Res. Part D: Transp. Environ. 15 (8), 522–524. https://doi.org/10.1016/j.trd.2010.07.002

Jonkeren, O., Kager, R., Harms, L., Te Brömmelstroet, M., 2021. The bicycle-train travellers in the Netherlands: personal profiles and travel choices. Transportation 48 (1), 455–476.

Kager, R., Harms, L., 2017. Synergies from improved cycling-transit integration: towards an integrated urban mobility system. Int. Transp. Forum Discuss. Pap.(2017–2023). https://doi.org/10.1787/ce404b2e-en

Leth, U., Shibayama, T., Brezina, T., 2017. Competition or supplement? Tracing the relationship of public transport and bike-sharing in Vienna. J. Geogr. Inf. Sci. 137 (2), 137–151. https://doi.org/10.1553/giscience2017_02_s137

Mark van Hagen, N., Menno de Bruyn, N. Typisch NS.

Médard de Chardon, C., 2019. The contradictions of bike-share benefits, purposes and outcomes. Transp. Res. Part A: Policy Pract. 121, 401–419. https://doi.org/10.1016/j.tra.2019.01.031

Miles, J., 2005. R-squared, adjusted R-squared. Encycl. Stat. Behav. Sci.

Nello-Deakin, S., Brömmelstroet, M., 2021. Scaling up cycling or replacing driving? Triggers and trajectories of bike–train uptake in the Randstad area. Transportation. https://doi.org/10.1007/s11116-021-10165-9

NS, 2021. Gebruik OV-fiets – NS Jaarverslag 2020. Jaarverslag 2020. 〈https://www.nsjaarverslag.nl/grafieken/grafieken/gebruik-ovfiets〉.

O'Brien, O., Cheshire, J., Batty, M., 2014. Mining bicycle sharing data for generating insights into sustainable transport systems. J. Transp. Geogr. 34, 262–273. https://doi.org/10.1016/j.jtrangeo.2013.06.007

Schakenbos, R., Ton, D., 2021. De Fietsende Treinreiziger: Spits of Dal Reiziger?. Presented at the Colloquium Vervoersplanologisch Speurwerk, Utrecht, Utrecht.

Shui, C.S., Szeto, W.Y., 2020. A review of bicycle-sharing service planning problems. Transp. Res. Part C: Emerg. Technol. 117. https://doi.org/10.1016/j.trc.2020.102648

Steyerberg, E.W., Eijkemans, M.J., Harrell, F.E., Habbema, J.D., 2001. Prognostic modeling with logistic regression analysis: in search of a sensible strategy in small data sets. Med. Decis. Mak. 21 (1), 45–56. https://doi.org/10.1177/0272989X0102100106

Todd, J., O'Brien, O., Cheshire, J., 2021. A global comparison of bicycle sharing systems. J. Transp. Geogr. 94. https://doi.org/10.1016/j.jtrangeo.2021.103119

Ton, D., et al., 2022. Teleworking during COVID-19 in the Netherlands: understanding behaviour, attitudes, and future intentions of train travellers. Transp. Res. Part A: Policy Pract.

van Waes, A., Farla, J., Frenken, K., de Jong, J.P.J., Raven, R., 2018. Business model innovation and socio-technical transitions. A new prospective framework with an application to bike sharing. J. Clean. Prod. 195, 1300–1312. https://doi.org/10.1016/j.jclepro.2018.05.223

Villwock-Witte, N., van Grol, L., 2015. Case study of transit–bicycle integration: openbaar Vervoer-fiets (public transport–bike) (OV-Fiets). Transp. Res. Rec. 2534 (1), 10–15. https://doi.org/10.3141/2534-02

de Visser, J., 2017. Succesfactoren Blue-bike. Breda University of Applied Sciences, Antwerp.

Zhang, C., Zhang, L., Liu, Y., Yang, X., 2018. Short-term prediction of bike-sharing usage considering public transport: a LSTM approach. In: Proceedings of the 2018 21st International Conference on Intelligent Transportation Systems (ITSC), pp. 1564–71. 〈https://doi.org/10.1109/ITSC.2018.8569726〉.