

# Chapter 1: Introduction

Draft version

## 1.1 Context

Over the past decades, the world has been urbanizing at a rapid pace. Where in 1950, 30% of the world's population lived in urban areas, this grew to 55% by 2018. It is a development of which the end is not yet in sight. By 2050, the urban population is projected to have increased to 68% (United Nations 2018). Managing the urban growth in a sustainable manner, balancing economical, social and environmental factors, forms a tremendous challenge. One of the cores of this challenge relates to transportation. Nowadays, urban transport policies still have a strong focus on the private car as leading transportation mode. However, neither from the economical, social nor environmental perspective, this is sustainable. Air pollution, resource depletion, fuel costs, congestion, noise, accidents and space requirements are among the elements that set limits to a car-oriented urban environment (Hickman and Banister 2014). Therefore, on our way towards more sustainable cities, with a high quality of life, a drastic shift of focus is required, which includes the integration of transport modes that provide feasible alternatives to car usage. Although changing travel behaviours is a complex and slow process, more and more cities across the world are acknowledging this need, and start acting upon it.

As argued by Pucher and Buehler (2017), the most sustainable and functional mode of transport in urban areas, is probably the bicycle. It can be used for short and medium distance trips, that form a large part of everyday city travel. Since it causes no air pollution, is healthy, has low usage and infrastructure costs, and requires little space, cycling is sustainable in both the economical, social and environmental sense. Not for nothing, it receives increasing attention in cities all over the world. The share of trips done by bicycle, has risen sharply in recent years, also in cities without a cycling tradition, and investments are made to improve the bicycle infrastructure. Furthermore, urban cycling is becoming a hot topic in academia as well, with a strong growth in the number of publications related to this topic over the last few years (Pucher and Buehler 2017; Fishman 2016).

Public bike sharing systems (PBSS) form an important part of the shift towards more cycling-oriented cities. They are build on the concept of a shared utilization of bicycle fleets, in which individuals can use bicycles whenever they need them, eliminating the costs and responsibilities that come with private bike ownership (Shaheen, Guzman, and Zhang 2010). Especially in cities without a cycling tradition, PBSS can normalize the image of cycling as a legitimate mode of everyday transport (Goodman, Green, and Woodcock 2014). Furthermore, the flexibility of PBSS make them a suitable way to improve the first and last mile connection to other transit modes (X.-H. Yang et al. 2018).

The number of PBSS worldwide grew from 5 in 2005 to 1571 in 2018 (Schmidt 2018). Although this extraordinary growth is a relatively recent development, PBSS have been around for much longer. Already in 1965, the first one started in Amsterdam. This system, known as *Witte Fietsen* (White Bikes), consisted of fifty permanently unlocked, white painted bikes that were scattered throughout the city, could be freely used, and left in any location. However, due to theft and vandalism, the experiment became a failure, and the system was shut down within days. It took 25 years before a new generation of PBSS entered the stage, in Denmark, 1991. In these systems, bikes had to be picked up and dropped off at designated docking stations, and payments were done with coin deposits. Because of the anonymity of the users, theft remained an issue. This lead to the third generation PBSS, which really took hold when Lyon's *Velo'v* system was introduced in 2005. Third generation PBSS used more advanced technologies, including electronic docking stations, information tracking, and electronic, dynamic payments with smartcards (DeMaio 2009; Shaheen, Guzman, and Zhang 2010). Over time, this evolved into the station-based bike sharing systems as we know them now, with the smartcard being replaced by mobile applications.

The modern, station-based systems are organized and well manageable. However, the accessibility of docking stations forms a large barrier for its usage: either there are no docking stations close enough to the desired trip origin, either there are no docking stations close enough to the desired trip destination (Fishman et al. 2014). The possibilities of increasing the number of docking stations are generally limited, due to high costs and space requirements. As an answer to this inconveniences, so-called *dockless bike sharing systems*, in some literature also referred to as *free floating bike sharing systems* or *flexible bike sharing systems*, have rapidly gained popularity in the last few years. In 2018, an estimated 16 to 18 million dockless shared bikes were in use, compared to 3.7 million station-based shared bikes (Schmidt 2018).

The dockless systems build on the philosophy behind the first generation PBSS, with bikes that can be picked up and dropped off at any location in the city. They are, however, supported by new technologies that limit the risk of theft and vandalism. The bicycles generally have integrated GPS tracking and an electronic lock. Usage is simple and convenient. Users download a mobile application and create a profile, which they link to their credit card, or charge with money from their bank account. On the app, they find an available bike, which they can unlock with their smartphone, usually by scanning a QR code. The ride can be ended anywhere, and a mobile payment is made, with the price depending on the travel time (Shen, Zhang, and Zhao 2018).

The revival of dockless bike sharing systems started in 2015, when two Chinese start-up companies, Mobike and Ofo, introduced them to several cities in China (Sun 2018). Since that moment, the expansion evolved extremely fast. Within one year, Mobike was present in eighteen Chinese cities with more than a million bikes in total, and other bike sharing start-ups emerged to follow their example (Van Mead 2017). From 2017 on, the systems spread across the world. In most cases, they were not welcomed with open arms by the local authorities and inhabitants, since regulations did not exist, and the companies simply dumped bikes throughout the cities, fighting for the highest market share. *BBC* talked about the dockless bikes that are ‘flooding the world’ (Belton 2018), *The New York Times* named it ‘the bike sharing explosion’ (Larmer 2017), and *The Guardian* even used the term ‘bike war’ (Collinson 2018).

Currently, the peak of the storm seems to have passed, and more and more cities have defined rules to control the dockless bike sharing expansion. These include requirements for operators to obtain a permit before entering the market, designated ‘system areas’ in which the bikes can be used, limits to the fleet size and restrictions to allowed parking spots (Deighton-Smith 2018). Thus, cities embrace dockless bike sharing in a well balanced way, take advantage of its convenience, and make it an important part of the strive to a more sustainable urban environment.

## 1.2 Objective

In order to become a serious alternative for motorized transport, PBSS need to be reliable. For station-based systems, this means that situations in which a user either can not start a trip at the desired time and location because of an empty docking station, or can not end a trip at the desired time and location because of a full docking station, should be avoided. In dockless bike sharing systems, the latter is not an issue, since bikes can be left anywhere and anytime. The first, however, remains a challenge. A user does not have the certainty of finding an available bike near the desired trip origin, at the desired time of departure.

Accurate, short-term forecasts of the distribution of available bikes both over space and time, enable system operators to anticipate on imbalances between supply and demand, such that the occurrence of situations as described above is limited to a minimum. Furthermore, users can take advantage of these forecasts also in a direct way, to plan their trips effectively, and reduce the time spend on searching for a bike. Hence, forecasting bike availability is of great importance when turning the shared bike into a reliable, pleasant and uncomplicated mode of transport.

Several approaches have been developed to forecast the bike availability in station-based bike sharing systems. However, dockless bike sharing systems remain fairly unexplored in that sense, since their rapid expansion only started very recently, and in contradiction to the station-based system, large historical data sets are not widely and openly available.

Furthermore, the spatio-temporal nature of dockless bike sharing data brings additional challenges to the forecasting task. In station-based systems, forecasts are only required at fixed spatial locations, and although some works include spatio-temporal relationships between stations, they can mostly be treated as distinct entities, with different forecasting models for each station. In dockless bike sharing systems, however, available bikes can be at any spatial location inside the system area. Besides, the bike availability not only has a spatial dependence, with more bikes being available in certain areas, and a temporal dependence, with more bikes being available at certain times, but those dependencies are also linked to each other. That is, temporal patterns may differ over space, and vice versa.

The objective of this thesis is to deal with those challenges, and develop a generally applicable methodology for bike availability forecasting in dockless bike sharing systems.

## 1.3 Related work

Pal, Zhang, and Kwon (2018) define two main categories in which publications related to PBSS can be classified. The first category involves the studies whose primary objective is to forecast the system’s future bike availability, bike demand, or similar. The second category involves the studies whose primary objective is to understand and describe the system, so that either its service level can be improved or the system can be expanded. Since this thesis aims to forecast bike availability, publications of the first category are the ones of interest. In essence, they all deal with time series analysis, and a wide range of forecasting methods has already been applied, primarily focused on station-based PBSS. These are covered in section 1.3.1, while the first attempts targeting dockless PBSS are reviewed in section 1.3.2.

### 1.3.1 Forecasting in station-based systems

One of the first works on bike availability forecasting was published ten years ago, by Froehlich, Neumann, and Oliver (2009). The study had an exploratory approach, focused on gaining a better understanding of the spatio-temporal patterns in station-based bike sharing systems and the factors that influence forecasts of those patterns. Some forecasting models were selected, ranging from very simple (e.g. a Last Value Predictor) to slightly more complex (e.g. a Bayesian Network). The most accurate forecasts were made at stations with a low usage, where the variability in the data was small. The Bayesian Network started outperforming the simpler methods at stations with a higher usage, and a corresponding large variability in the data. Usage intensity did not only vary over space, but also over time. The same conclusions were drawn in this sense, with the most accurate forecasts during the night, and the highest difference between simple and advanced methods during peak hours. Furthermore, the length of the forecasting window influenced forecasts, with higher forecast errors at larger forecasting windows. Using more historical data generally increased forecast accuracies, but recent observations were found to have a considerably higher influence than distant ones. External variables such as weather and special events were not considered.

Although the conclusions of Froehlich, Neumann, and Oliver (2009) seem obvious now, they formed an important basis for further research, with more advanced forecasting methods. Kaltenbrunner et al. (2010) fitted Auto Regressive Integrated Moving Average (ARIMA) models to the historical bike availability data of each docking station in Barcelona’s *Bicing* system, incorporated information from neighbouring stations into these models, and forecasted the amount of available bicycles up to one hour ahead. Won Yoon, Pinelli, and Calabrese (2012) used a similar approach, but included recurring seasonal patterns in the ARIMA models. Borgnat et al. (2011), instead, split the forecasting task in two: linear regression, with weather, usage intensity and specific conditions like holidays as explanatory variables, was used to forecast a non-stationary amplitude for a given day, while an autoregressive (AR) model forecasted the hourly fluctuations in the data.

Rixey (2013) created linear regression models to forecast the number of monthly rentals per docking station. Explanatory variables such as demographic and built environment factors were extracted from a circular buffer with a radius of 400 meters around each station. D. Singhvi et al. (2015) did a comparable study,

but on a macroscopic level, arguing that aggregating stations in neighborhoods can substantially improve predictions.

In 2014, Kaggle, a well-known online community for data scientists, hosted a machine learning competition, in which participants were asked to combine historical usage patterns with weather data, and forecast bike demand in the station-based PBSS of Washington D.C. (Kaggle 2014). The competition increased the academic interest in bike availability forecasting with machine learning methods, and new publications on this topic followed rapidly. Giot and Cherrier (2014) compared Ridge Regression, Adaboost Regression, Random Forest Regression, Gradient Tree Boosting Regression and Support Vector Regression by predicting Washington’s city-wide bike sharing usage up to one day ahead, inputting both recent observations, time features (e.g. season, day of the week, hour of the day) and weather features (e.g. temperature, wind speed). Ridge Regression and Adaboost Regression turned out to be the best performing methods. Y. Li et al. (2015) used Gradient Tree Boosting Regression with time and weather features. However, they first grouped the docking stations into distinct, spatially contiguous clusters, and forecasted the bike usage per cluster. Dias, Bellalta, and Oechsner (2015) attempted to simplify the task, by using a Random Forest Regressor to forecast only if a docking station will be either completely full, completely empty, or anything in between.

Later, spatial relations between individual docking stations were incorporated into the machine learning approaches, which resulted in more sophisticated forecasting methods. Z. Yang et al. (2016) proposed a probabilistic spatio-temporal mobility model that considered previous check-out records at other stations and expected trip durations to estimate the future check-ins at each station. Subsequently, check-outs at each station were estimated with a Random Forest Regressor, that also took weather forecasts into account. Lin, He, and Peeta (2018) modelled station-based PBSS as a graph, with each station being a node. Then, they applied a Convolutional Neural Network in a form generalized for graph-structured data, to forecast hourly bike demand. Lozano et al. (2018) created an automated multi-agent bike demand forecasting system for the Salamanca’s *SalenBici* system, with digital ‘agents’ that collect different types of data, such as calendar events, weather forecasts and historical usage. These data are then combined, and used to forecast check-ins and check-outs separately for each station, with a Random Forest Regressor.

The Random Forest and Convolutional Neural Network approaches were compared by Ruffieux et al. (2017). Furthermore, they differentiated themselves by forecasting bike availability in six different cities, instead of only one. The outcomes showed that the Random Forest Regressor works better for short-term forecasts, while Convolutional Neural Networks are more suitable for long-term forecasts. B. Wang and Kim (2018) compared the Random Forest approach to two of the latest machine learning innovations, Long Short Term Memory Neural Networks and Gated Recurrent Units, but found no substantial improvements in the results.

All discussed academic publications related to bike availability forecasting in station-based PBSS are listed in Table 1.1, in chronological order. It can be seen that machine learning methods have received a lot of attention in recent years, compared to traditional statistical forecasting methods such as ARIMA. The two serve the same goal, but have a different philosophy and model development process. Traditional statistical methods assume a functional form in advance, concern inference and estimation, and aim at providing models that offer insights on the data. Machine learning methods approximate the functional form via learning inside a ‘black box’, do not require specifications of the model and the error distribution, and aim solely at providing an efficient forecast (Ermagun and Levinson 2018).

Studies in the field of bike availability forecasting that make valuable comparisons between traditional time series forecasting and machine learning methods, are scarce. Y. Li et al. (2015) and Z. Yang et al. (2016) included ARIMA in their baseline methods, but only in a form that does not take any seasonal component into account, while Dias, Bellalta, and Oechsner (2015) compared a seasonal ARIMA model with their proposed Random Forest Regressor, but without using the same amount of data for both methods. In spatio-temporal transportation forecasting in general, there is no clear consensus on which method is ‘better’. Despite the strong increase in publications that use machine learning, Ermagun and Levinson (2018) state that “there is not a certain superiority when machine learning methods are compared with advanced statistical methods such as autoregressive integrated moving average.”

Table 1: Publications regarding forecasts in station-based bike sharing systems, known to the author

Authors	Year	Main forecasting method	Spatial unit of analysis	Case study
Froehlich, Neumann, and Oliver	2009	Bayesian Network	Docking station	Barcelona
Kaltenbrunner et al.	2010	Spatial ARIMA	Docking station	Barcelona
Borgnat et al.	2011	Linear regression and AR	Docking station	Lyon
Won Yoon, Pinelli, and Calabrese	2012	Spatial, seasonal ARIMA	Docking station	Dublin
Rixey	2013	Linear regression	Docking station	Washington DC, Minneapolis and Denver
Giot and Cherrier	2014	Ridge regression, Adaboost regression, Random Forest regression, Gradient Tree Boosting regression and Support Vector regression	City	Washington DC
D. Singhvi et al.	2015	Linear regression	Neighbourhood	New York
Y. Li et al.	2015	Gradient Tree Boosting regression	Cluster	Washington DC
Dias, Bellalta, and Oechsner	2015	Random Forest regression	Docking station	Barcelona
Z. Yang et al.	2016	Probabilistic mobility model and Random Forest regression	Docking station	Hangzhou
Ruffieux et al.	2017	Random Forest regression and Convolutional Neural Network	Docking station	Namur, Essen, Glasgow, Budapest, Vienna and Nice
Lin, He, and Peeta	2018	Graph Convolutional Neural Network	Docking station	New York
Lozano et al.	2018	Random Forest regression	Docking station	Salamanca
B. Wang and Kim	2018	Long Short Term Memory Neural Networks and Gated Recurrent Units	Docking station	Suzhou

### 1.3.2 Forecasting in dockless systems

Publications regarding bike availability forecasting in dockless PBSS can be counted on the fingers of one hand. In addition to them, a small number of similar studies have been done in the field of dockless car sharing, which is based on the same principles as dockless bike sharing.

Those few attempts can be divided in two groups of approaches, which are labeled here as the *grid based approach* and the *distance based approach*. In the grid based approach, the system’s operational area, i.e. the ‘system area’, is divided into distinct grid cells, which may be regularly shaped, but do not have to be. For example, the cell boundaries may also correspond with administrative boundaries. Subsequently, each cell is treated as being a docking station. That is, from the historical data on locations of available vehicles, the number of vehicles in each cell is counted at several timestamps in the past, creating a time series of counts.

Hence, for a given geographical point, the forecasted value will be the expected number of vehicles inside the cell that contains the point.

The distance based approach uses the historical data to calculate the distance from a given geographical location to the nearest available vehicle, for several timestamps in the past, creating a time series of real-valued distances. Hence, the forecasted value will be the expected distance from the given point to the nearest available vehicle, at a given timestamp in the future.

The grid based approach was taken by Caggiani et al. (2017), who were among of the first to forecast bike availability in dockless PBSS. They did not use a regular grid, but defined the outline of the cells with a spatio-temporal clustering procedure, that worked as follows. Historical time series containing daily counts of available bikes were obtained for each zone, and clustered based on temporal patterns. Geographically connected zones that belonged to the same temporal cluster, were aggregated. The resulting clusters formed the spatial units of analysis, and for each of them, the number of available bikes was forecasted one day ahead with a Nonlinear Autoregressive Neural Network. However, at that time, they lacked data to test their methodology on. Instead, they used data from the station-based PBSS in London, pretending each docking station to be a cluster centroid.

Yi et al. (2018) did have access to a dataset of a dockless PBDSS, in the Chinese city of Chengu, and laid a regular grid with square cells over its systems area. Besides historical count data of available bikes in each grid cell, they included a categorical variable representing the time of day in their system. Forecasting was done with a recent machine learning method, named Convolutional Long Short Term Memory Network. Müller and Bogenberger (2015), in contradiction, used traditional statistical methods, and focused on forecasting the number of car bookings in a dockless car sharing system, for each ZIP-code area in Berlin. They compared a seasonal ARIMA model with a Holt Winters exponential smoothing model, and found acceptable and similar forecasting errors for both of them.

The only example of the application of the distance based approach, known to the author, is the work of Formentin, Bianchessi, and Savaresi (2015). Their methodology, focused on dockless car sharing systems, can be summarized as follows. For a set of given locations, evenly distributed over the city, historical location data is pre-processed into a time series of historical distance data. A strictly linear trend is subtracted from these time series, and if present, the seasonal component as well, using the classical time series decomposition method. AR models are fitted to the data that are left. Then, a model structure that is acceptable for all time series, is identified, and serves as the general forecasting model for the complete city. That is, for individual forecast requests, this general model structure will be used, no matter what the location of the request is. Only when a forecast request is located inside the city center, the corresponding historical distance data will be decomposed before forecasting, assuming a daily seasonal pattern. In that case, the seasonal component will afterwards be added to the forecast produced by the general model. The general model is updated every month.

Table 2: Publications regarding forecasts in dockless vehicle sharing systems, known to the author

Authors	Year	Main forecasting method	Approach	Case study	Type
Formentin, Bianchessi, and Savares	2015	AR	Distance based	Milan	Car
Müller and Bogenberger	2015	ARIMA and Holt Winters ETS	Grid based	Berlin	Car
Caggiani et al.	2017	Nonlinear Autoregressive Neural Network	Grid based	London	Bike
Yi et al.	2018	Convolutional Long Short Term Memory Network	Grid based	Chengu	Bike

## 1.4 Approach

The bike availability forecasting system as proposed in this thesis extends the methodology proposed by Formentin, Bianchessi, and Savaresi (2015), and uses the distance based approach. That is, the distance to the nearest available bike will be the variable of interest. The distance based approach is chosen, because it has the following advantages over the grid based approach.

- Forecasts will not be dependent on the chosen spatial resolution of the grid.
- Forecasts will not be made with count data, which would have limited the choice of suitable forecasting models.
- Forecasts can be interpreted in the same way at every location in space. This in contradiction to the grid based approach, where for a location at the edge of a grid cell, there might as well be closer bikes within the neighboring cell.
- Forecasts give more freedom to the user, who can decide for himself if he considers the forecasted distance acceptable or not. In the grid based approach, the cell size is fixed, and does not take into account that some people are willing to walk further than others.

Throughout the thesis, the proposed forecasting system will be referred to as the **Dockless Bike Availability Forecasting System (DBAFS)**. The focus will lay on the design of a general methodology for it, rather than on the creation of a fully configured, practical application. A forecast can be requested for any location within the system area of a dockless bike sharing system, and, theoretically, any timestamp in the future. However, short-term forecasts, up to one day ahead, form the target, and longer forecasting horizons will not be handled.

## 1.5 Outline

The rest of the document is structured as follows. Chapter 2 provides a theoretical background of the concepts used in this thesis, which are all linked to the field of time series analysis. In Chapter 3, the methodology of the proposed forecasting system is described into detail. Chapter 4 presents the data on which the system is tested, and describes the experimental setup. In Chapter 5, the results of the experiments are shown, interpreted, and discussed. Finally, Chapter 6 lists the conclusions of this thesis.

## References

- Belton, Pdraig. 2018. “How cheap dockless hire bikes are flooding the world.” <https://www.bbc.com/news/business-44066083>.
- Borgnat, Pierre, Celine Robardet, Jean-Baptiste Rouquier, Patrice Abry, Patrick Flandrin, and Eric Fleury. 2011. “Shared Bicycles in a City: A Signal Processing and Data Analysis Perspective.” *Advances in Complex Systems* 14. doi:10.1142/S0219525911002950.
- Caggiani, Leonardo, Michele Ottomanelli, Rosalia Camporeale, and Mario Binetti. 2017. “Spatio-temporal clustering and forecasting method for free-floating bike sharing systems.” In *Advances in Intelligent Systems and Computing*, 539:244–54. Springer Verlag. [http://dx.doi.org/10.1007/978-3-319-48944-5\\_{\\\_}23](http://dx.doi.org/10.1007/978-3-319-48944-5_{\_}23).
- Collinson, Patrick. 2018. “Bike wars: Chinese bike-share giants wheel out UK expansion plans.” <https://www.theguardian.com/uk-news/2018/mar/12/bike-wars-chinese-bike-share-giants-wheel-out-expansion-plans-in-uk>.
- Deighton-Smith, Rex. 2018. “The Economics of Regulating Ride-Hailing and Dockless Bike Share.” International Transport Forum. <https://www.itf-oecd.org/economics-regulating-ride-hailing-and-dockless-bike-share>.
- DeMaio, Paul. 2009. “Bike-sharing: History, Impacts, Models of Provision, and Future.” *Journal of Public Transportation* 12 (4). doi:10.5038/2375-0901.12.4.3.
- Dias, Gabriel Martins, Boris Bellalta, and Simon Oechsner. 2015. “Predicting occupancy trends in Barcelona’s bicycle service stations using open data.” In *2015 Sai Intelligent Systems Conference (Intellisys)*, 439–45. doi:10.1109/IntelliSys.2015.7361177.
- Ermagun, Alireza, and David Levinson. 2018. “Spatiotemporal traffic forecasting: review and proposed directions.” *Transport Reviews* 38 (6). Routledge: 786–814. doi:10.1080/01441647.2018.1442887.
- Fishman, Elliot. 2016. “Cycling as transport.” *Transport Reviews* 36 (1). Routledge: 1–8. doi:10.1080/01441647.2015.1114271.
- Fishman, Elliot, Simon Washington, Narelle Haworth, and Armando Mazzei. 2014. “Barriers to bikesharing: an analysis from Melbourne and Brisbane.” *Journal of Transport Geography* 41: 325–37. doi:<https://doi.org/10.1016/j.jtrangeo.2014.08.005>.
- Formentin, Simone, Andrea G. Bianchessi, and Sergio M. Savaresi. 2015. “On the prediction of future vehicle locations in free-floating car sharing systems.” In *2015 Ieee Intelligent Vehicles Symposium (Iv)*, 1006–11. doi:10.1109/IVS.2015.7225816.
- Froehlich, Jon, Joachim Neumann, and Nuria Oliver. 2009. “Sensing and predicting the pulse of the city through shared bicycling.” *IJCAI International Joint Conference on Artificial Intelligence*, no. 3: 1420–6. doi:10.1.1.150.4370.
- Giot, Romain, and Raphael Cherrier. 2014. “Predicting bikeshare system usage up to one day ahead.” In *2014 Ieee Symposium on Computational Intelligence in Vehicles and Transportation Systems (Civts)*, 22–29. doi:10.1109/CIVTS.2014.7009473.
- Goodman, Anna, Judith Green, and James Woodcock. 2014. “The role of bicycle sharing systems in normalising the image of cycling: An observational study of London cyclists.” *Journal of Transport & Health* 1 (1): 5–8. doi:<https://doi.org/10.1016/j.jth.2013.07.001>.
- Hickman, Robin, and David Banister. 2014. *Transport, Climate Change and the City*. Routledge Advances in Climate Change Research. Taylor & Francis.
- Kaggle. 2014. “Bike Sharing Demand: Forecast use of a city bikeshare system.” <https://www.kaggle.com/c/bike-sharing-demand>.
- Kaltenbrunner, Andreas, Rodrigo Meza, Jens Grivolla, Joan Codina, and Rafael Banchs. 2010. “Urban cycles and mobility patterns: Exploring and predicting trends in a bicycle-based public transport system.”



- Pervasive and Mobile Computing* 6 (4). Elsevier B.V.: 455–66. doi:10.1016/j.pmcj.2010.07.002.
- Larmer, Brook. 2017. “China’s Revealing Spin on the ‘Sharing Economy.’” <https://www.nytimes.com/2017/11/20/magazine/chinas-revealing-spin-on-the-sharing-economy.html>.
- Li, Yexin, Yu Zheng, Huichu Zhang, and Lei Chen. 2015. “Traffic prediction in a bike-sharing system.” In *23rd Sigspatial International Conference*, 1–10. doi:10.1145/2820783.2820837.
- Lin, Lei, Zhengbing He, and Srinivas Peeta. 2018. “Predicting station-level hourly demand in a large-scale bike-sharing network: A graph convolutional neural network approach.” *Transportation Research Part C: Emerging Technologies* 97: 258–76. doi:<https://doi.org/10.1016/j.trc.2018.10.011>.
- Lozano, Álvaro, Juan F De Paz, Gabriel Villarrubia, Daniel H De La Iglesia, and Javier Bajo. 2018. “Multi-Agent System for Demand Prediction and Trip Visualization in Bike Sharing Systems.” *Applied Sciences* 8 (1): 67. doi:10.3390/app8010067.
- Müller, Johannes, and Klaus Bogenberger. 2015. “Time Series Analysis of Booking Data of a Free-Floating Carsharing System in Berlin.” *Transportation Research Procedia* 10: 345–54. doi:<https://doi.org/10.1016/j.trpro.2015.09.084>.
- Pal, Aritra, Yu Zhang, and Changhyun Kwon. 2018. “Analysis of Free-floating Bike Sharing and Insights on System Operations.” Civil; Environmental Engineering University of South Florida. <https://hdl.handle.net/1813/56569>.
- Pucher, John, and Ralph Buehler. 2017. “Cycling towards a more sustainable transport future.” *Transport Reviews* 37 (6). Routledge: 689–94. doi:10.1080/01441647.2017.1340234.
- Rixey, Alexander R. 2013. “Station-Level Forecasting of Bikesharing Ridership.” *Transportation Research Record: Journal of the Transportation Research Board* 2387: 46–55. doi:10.3141/2387-06.
- Ruffieux, Simon, Nicolas Spycher, Elena Mugellini, and Omar Abdou Khaled. 2017. “Real-time usage forecasting for bike-sharing systems: A study on random forest and convolutional neural network applicability.” In *2017 Intelligent Systems Conference (Intellisys)*, 622–31. doi:10.1109/IntelliSys.2017.8324359.
- Schmidt, Charles. 2018. “Active Travel for All? The Surge in Public Bike-Sharing Programs.” *Environmental Health Perspectives* 126: 82001. doi:10.1289/EHP3754.
- Shaheen, Susan, Stacey Guzman, and Hua Zhang. 2010. “Bikesharing in Europe, the Americas, and Asia: Past, Present, and Future.” *Transportation Research Record Journal of the Transportation Research Board* 2143 (1316350). doi:10.3141/2143-20.
- Shen, Yu, Xiaohu Zhang, and Jinhua Zhao. 2018. “Understanding the usage of dockless bike sharing in Singapore.” *International Journal of Sustainable Transportation* 12 (9). Taylor & Francis: 686–700. doi:10.1080/15568318.2018.1429696.
- Singhvi, Divya, Somya Singhvi, Peter I. Frazier, Shane G. Henderson, Eoin O’Mahony, David B. Shmoys, and Dawn B. Woodard. 2015. “Predicting Bike Usage for New York City’s Bike Sharing System.” In *AAAI Workshop: Computational Sustainability*.
- Sun, Yiyun. 2018. “Sharing and Riding: How the Dockless Bike Sharing Scheme in China Shapes the City.” *Urban Science* 2 (3). doi:10.3390/urbansci2030068.
- United Nations. 2018. “World Urbanization Prospects 2018.” <https://population.un.org/wup/>.
- Van Mead, Nick. 2017. “Uber for bikes: how ‘dockless’ cycles flooded China – and are heading overseas.” <https://www.theguardian.com/cities/2017/mar/22/bike-wars-dockless-china-millions-bicycles-hangzhou>.
- Wang, Bo, and Inhi Kim. 2018. “Short-term prediction for bike-sharing service using machine learning.” *Transportation Research Procedia* 34: 171–78. doi:<https://doi.org/10.1016/j.trpro.2018.11.029>.
- Won Yoon, Ji, Fabio Pinelli, and Francesco Calabrese. 2012. “Cityride: A Predictive Bike Sharing Journey Advisor.” In *IEEE 13th International Conference on Mobile Data Management*, 306–11. doi:10.1109/MDM.2012.16.
- Yang, Xu-Hua, Zhi Cheng, Guang Chen, Lei Wang, Zhong-Yuan Ruan, and Yu-Jun Zheng. 2018. “The

impact of a public bicycle-sharing system on urban public transport networks.” *Transportation Research Part A: Policy and Practice* 107: 246–56. doi:<https://doi.org/10.1016/j.tra.2017.10.017>.

Yang, Zidong, Ji Hu, Yuanchao Shu, Peng Cheng, Jiming Chen, and Thomas Moscibroda. 2016. “Mobility Modeling and Prediction in Bike-Sharing Systems.” *Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services - MobiSys '16*, 165–78. doi:10.1145/2906388.2906408.

Yi, Ai, Zongping Li, Mi Gan, Yunpeng Zhang, Daben Yu, Wei Chen, and Yanni Ju. 2018. “A deep learning approach on short-term spatiotemporal distribution forecasting of dockless bike-sharing system.” *Neural Computing and Applications*, 1–13. doi:10.1007/s00521-018-3470-9.