Front. Eng. Manag. 2023, 10(4): 551–565
https://doi.org/10.1007/s42524-023-0279-8

RESEARCH ARTICLE

Zhuoli YIN, Kendrick HARDAWAY, Yu FENG, Zhaoyu KOU, Hua CAI

Understanding the demand predictability of bike share systems: A station-level analysis

Abstract Predicting demand for bike share systems (BSSs) is critical for both the management of an existing BSS and the planning for a new BSS. While researchers have mainly focused on improving prediction accuracy and analysing demand-influencing factors, there are few studies examining the inherent randomness of stations' observed demands and to what degree the demands at individual stations are predictable. Using Divvy bike-share one-year data from Chicago, USA, we measured demand entropy and quantified the station-level predictability. Additionally, to verify that these predictability measures could represent the performance of prediction models, we implemented two commonly used demand prediction models to compare the empirical prediction accuracy with the calculated entropy and predictability. Furthermore, we explored how city- and system-specific temporally-constant features would impact entropy and predictability to inform estimating these measures when historical demand data are unavailable. Our results show that entropy and predictability of demands across stations are polarized as some stations exhibit high uncertainty (a low predictability of 0.65) and others have almost no check-out demand uncertainty (a high predictability of around 1.0). We also validated that the entropy and predictability are *a priori* model-free indicators for prediction error, given a sequence of bike usage demands. Lastly, we identified that key factors contributing to station-level entropy and predictability include per capita income, spatial eccentricity, and the number of parking lots near the station. Findings from this study provide more fundamental understanding of BSS demand prediction, which can help decision makers and system operators anticipate diverse station-level prediction errors from their prediction models both for existing stations and for new ones.

Keywords bike share systems, demand prediction, prediction errors, machine learning, entropy

Received Apr. 4, 2023; revised Aug. 22, 2023; accepted Sep. 11, 2023

Zhuoli YIN, Zhaoyu KOU
School of Industrial Engineering, Purdue University, West Lafayette, IN 47907, USA

Kendrick HARDAWAY
Environmental and Ecological Engineering, Purdue University, West Lafayette, IN 47907, USA

Yu FENG
School for Environment and Sustainability, University of Michigan, Ann Arbor, MI 48109, USA

Hua CAI (✉)
School of Industrial Engineering, Purdue University, West Lafayette, IN 47907, USA; Environmental and Ecological Engineering, Purdue University, West Lafayette, IN 47907, USA
E-mail: huacai@purdue.edu

1 Introduction

The bike share system (BSS) market has rapidly expanded in recent years and is expected to triple by 2030 (Fishman and Allan, 2019; Straits Research, 2021). The traditional BSS is station-based, allowing travellers to pick up and return bikes at designated locations (Kou and Cai, 2019). BSSs allow travellers to use bikes on a need-basis either for a fee or for free, providing a convenient and accessible mobility option, especially for their first/last-mile trips (Bachand-Marleau et al., 2012; Fishman, 2016). As part of the sharing economy and as a viable substitute for short private car-based trips, BSS also has the potential to reduce greenhouse gas emissions (Shaheen et al., 2010; Kou et al., 2020; Zhou et al., 2023). However, these benefits may remain unattained if the BSSs are not well planned or managed. Unplanned or unmanaged BSS can lead to over- and under-supply, inconvenient parking, lower service level, street safety issues, and suboptimal business operations (Regue and Recker, 2014; Chen et al., 2016). The rapid growth in BSSs and the necessity for intelligent forward-looking design has led researchers to study the implementation and improvement of these systems (Luo et al., 2020; Kou and Cai, 2021a).

Demand prediction for BSS stations is the foundation

of planning and managing BSSs. Due to spatial and temporal imbalance in BSS demand, it is common for BSSs to suffer from disparity of either undersupply or oversupply of bikes among different stations (Li et al., 2015; Chen et al., 2016). Inaccurate demand prediction can cascade to improper bike rebalancing, increased operational costs, reduced user satisfaction, and more greenhouse gas emissions. Therefore, system operators and researchers have devoted significant attention to predicting the demands of BSS stations (El-Assi et al., 2017; Hyland et al., 2018; Zhou et al., 2018; Kou and Cai, 2021b).

Existing demand prediction studies have used traditional regression, machine learning, or deep learning models to predict BSS demand in different ways. For instance, Médard de Chardon and Caruso (2015) used linear regression-based method for estimating demands for daily BSS trips using station-level data in six cities. Hulot et al. (2018) used linear regression and machine learning models to predict hourly demand and recommended intervals for how regularly to rebalance the bikes. Convolutional neural network based models have been used to predict bike inflow and outflow at stations and station-level hourly demand (Chai et al., 2018; Lin et al., 2018; Yu et al., 2018). Cluster-based regressions have been used to predict pickups and returns of bikes to stations with similar characteristics and to predict citywide bike usage (Chen et al., 2016; Jia et al., 2019; Li and Zheng, 2020). Additionally, a few studies have investigated residual correction to reveal hidden stochasticity in time series data to improve prediction (Kim et al., 2022; Zheng et al., 2023). Among existing models, in order to improve the prediction accuracy, previous studies also integrated spatiotemporal variables to analyse their role in impacting demand. Bao et al. (2017) investigated bike share travel patterns and trip purposes by combining smart card data and point of interests (POIs). Zhou (2015) and Lin et al. (2020) revealed the spatiotemporal patterns of bike sharing behavior and identified influential factors such as the built environment on bike share trip demands.

Various features have been incorporated in demand prediction as well, in addition to spatiotemporal information. Yang et al. (2016; 2019) created a probabilistic spatiotemporal model using dynamic networks, time factors, and weather. Hulot et al. (2018) predicted demand with temporal and weather variables using linear regression and machine learning models. Singhvi et al. (2015) used temporal, demographic, and weather factors in their pairwise model. Chen et al. (2016) used the temporal and weather factors but added social events (such as city festivals, parades, or traffic accidents) in their cluster-based prediction. However, despite the advancement in demand forecasting techniques, existing studies mainly focused on how to improve the prediction of demands. The accuracy of demand prediction is only known posteriorly after the prediction model has been developed and it requires extensive domain-specific feature engineering from the researchers. No existing studies have addressed the fundamental predictability of demands at stations, and there is a gap in understanding how the intrinsic randomness (i.e., the inherent variation of BSS demands due to unpredictable factors such as weather conditions and human behaviors) governs the limit of future demand prediction. By understanding the governing randomness in demand levels, system operators and city administrators can better manage and maintain stations. Meanwhile, little research has evaluated how temporal invariant determinants shape the random nature of demand patterns at individual stations. These temporal invariant determinants intrinsically characterize the functionalities of different zones in cities as well as shape the heterogeneities among bike share stations. They are relatively stable over time compared to time dependent determinants, including seasonality and unique events. There exists a research gap in studying variables accounting for such randomness, where we could gain knowledge about heterogeneous station-level demand predictability even prior to the launch of a BSS in the city.

To evaluate the accuracy of the predictive models, however, current demand prediction studies have primarily applied one single model for the entire system and evaluated prediction model performance at the system-level. These studies rely on station-level data and employ evaluation metrics such as root mean squared error (RMSE) and mean absolute percentage error (MAPE), but only inspecting the aggregated outcome at the system-level may fail to capture the variation and anomalies among stations and dilute the local understanding of predictive performance on station-level (Li et al., 2015; Médard de Chardon and Caruso, 2015; Singhvi et al., 2015; El Sibai et al., 2018; Liu et al., 2022). For instance, station-level understanding is critical for rebalancing at the station-level and maintaining efficient system operation. Consequently, going beyond focusing on global average performance, there exists a knowledge gap on how prediction errors deviate at high resolution and how the inherent predictability of demands differs at individual stations (He and Shin, 2020).

To address the aforementioned research gaps, this study quantified the randomness and predictability rooted in time series bike check-out demands at individual stations based on entropy and predictability from information theory. The calculation of these two metrics was based on one-year's data, which can capture relatively long-term variations and seasonal changes. To test the validity of our measurements, we compared the station-level entropy and predictability with the empirical prediction performance from two benchmark prediction models — Auto Regressive Moving Average (ARMA) and XGBoost. This establishes a viable mapping from model-based accuracy/error to model-free measurements, allowing for anticipation of demand prediction performance without the need of prior feature engineering and prediction

algorithms. Additionally, to further examine how temporal invariant factors of a city impact such intrinsic randomness of BSS demands, this work used random forest regression model to identify the most significant temporal invariant factors (i.e., features that remain consistent over the study period, without significant variation, such as infrastructure) in contributing to station-level entropy and predictability. In the context of the existing literature, this study makes three primary contributions:

(1) We adopted entropy and predictability as model-free measurements for better investigating station-level inherent demand predictability.

(2) Through empirical experiments, we further established these two measurements as representative of the practical prediction algorithm performance, without the need of feature engineering and building prediction models.

(3) We offered managerial insights for system operators and city authorities who are considering launching new BSS or expanding their existing BSS by identifying the key factors impacting the entropy and predictability of individual stations.

The subsequent sections of the paper are organized as follows. Section 2 introduces the data, data processing methods, demand randomness measurements, benchmark demand prediction algorithms, and evaluation metrics. In Section 3, we present the overview of the computed entropy and predictability. We also show the association between the entropy/predictability and empirical performances achieved by demand prediction models at the individual station-level. In addition, we determine the most notable factors influencing station demands' entropy and predictability. Last, Section 4 draws inferences and implications about demand predictability from the results, discusses the limitations, and suggests future research directions.

2 Data and method

2.1 Data and data processing

In this study, our goal was to quantify the inherent randomness and predictability of bike check-out demands at individual bike share stations as well as identify what city factors would contribute to such randomness. To achieve this, we collected one year of historical Divvy trip records (366 days, Aug. 1, 2015 – Aug. 1, 2016) as demand data, where the record for each single trip includes the trip ID, start and end time of the trip, the check-out and check-in station IDs & names, and trip duration. There are 534 stations in total with complete trip records within our study period. Moreover, we partitioned the pre-processed historical trip data into two sets: One for training/analysis, spanning 361 days, and the other for testing, spanning 5 days. We selected Chicago as our case study city because of Chicago's long history of operating station-based BSS (active since 2013) and Divvy's wide service coverage, spanning throughout both downtown and suburban areas of the city. The start time and check-out station ID for each ride were extracted from the trip records for this study. Subsequently, we aggregated check-out demands at individual stations based on a certain time interval, Δt, which could range from hourly (Δt = 1 h) to daily (Δt = 24 h) segments, based on what were most used in previous works (Hulot et al., 2018; Kou and Cai, 2021a). In this study, we selected the time interval of four hours (Δt = 4 h) as the base scenario and this resulted in a series of six demand data points for each station per day (we also conducted sensitivity analysis to examine how using different time intervals may impact the results). We therefore formulated a series of observed check-out demand level at station i over the one-year timespan T as $T_i = \{X_{i,1}, X_{i,2}, X_{i,3}, ..., X_{i,t}, X_{i,t+1}, ..., X_{i,T}\}$ where $X_{i,t}$ indicates the aggregated check-out demand level at time step t. For each of the 534 Divvy stations, we also acquired data consisting of station ID, name, latitude, longitude, and capacity. We used such station information as source data to construct variables for the spatial network listed in Table 1. Additionally, we collected the point of interests (POIs) within a 1000 ft (304.8 m) radius around each station based on Google Maps Places API (application program interface). POIs encompass a series of specific functional location, such as bus station, school and parking lot. Furthermore, we sourced socio-demographic data, specifically per capita income and population density, from the American Community Survey (US Census Bureau, 2012) at the census-tract level. The spatial network, POIs, and socio-demographic data prepared above were utilized in Section 2.3 to analyze the key factors that contribute to the entropy and predictability.

2.2 Measures of demand randomness and predictability

In the context of BSS demand prediction, prediction models have been developed to statistically fit the relationship between the spatiotemporal variables and the bike usage demands, via either linear or nonlinear approaches (Senter, 2008). How much of the demands can be predicted depend on the degree of uncertainty rooted in the demand pattern at a station. Therefore, we harnessed the concept of entropy in information theory (Shannon, 1948) to measure the degree of randomness in a sequence of demands at individual stations (Section 2.2.1). In addition to the entropy, we computed each station's predictability to capture the upper limit to which these demands can be correctly forecasted (Section 2.2.2). These two measurements thereby provide a quantitative assessment of the inherent predictability of demands at each installed bike share station, based on a stable and relative long-term (one-year) trend of demands rather

Table 1 Definitions of temporally-constant feature variables at the station-level used in this study

Category	Variable	Definition	Unit
Land use	Bus station	The counts of POIs within a specified distance buffer [a)]	N/A
	Subway station		
	Restaurant		
	Park		
	Parking lots		
	Museum		
	School		
Spatial network	Spatial eccentricity	The average distance between station i and all other stations in the network [b)] (Zhang et al., 2016; Cazabet et al., 2017)	mile
	Station density	The number of other stations located within each distance band of a station i, where distance bands are the concentric distance bands around each station with a certain increment [c)] (Hyland et al., 2018)	N/A
	Station capacity	The maximum number of bike docks in a station	N/A
Socio-demographic [d)]	Per capita income	Mean income earned of each person in a given area	US dollar
	Population density	The population divided by the area of a census tract	People/mile^2

Notes: a) The POIs were collected using Google Maps Places API. We set the distance buffer as 1000 ft (304.8 m), representing a common walkable range for a person between a bike share station and a POI; b) As BSS is under continuous expansion by installing new stations over time, the spatial eccentricity of a station could vary. Therefore, we calculated the average spatial eccentricity within our studied time span; c) In this study, we set the increment as 0.1 miles (161 m) and limited the calculation of station density to be within the range of 0.05 miles to 4.5 miles, as suggested by Kou and Cai (2021b); d) Socio-demographic data were collected from the 2017 American Community Survey at the census-tract level. The values of one census tract are assigned to a station if the station is located within the census tract.

than various short-term shifts (Lu et al., 2013). Additionally, entropy and predictability imply the station-wise theoretical upper bound of unexplained uncertainty and potential margin of error associated with a given demand prediction model (Song et al., 2010). These two metrics can serve as *a priori* model-free evaluation for fundamental station demand patterns.

2.2.1 Entropy

Given a sequence of observed check-out demand level T_i for each station $i \in \{1, 2, 3, ..., 534\}$, we utilized three entropy measurements to reflect the randomness of station's demands using different amount of information, as proposed by Song et al. (2010):

(1) The random entropy $S_i^{\text{rand}} = \log_2 N_i$, which indicates the disorder of a station's demand level, with the assumption that each demand level is observed with equal probability among N_i unique demand levels (with the resolution of one bike check-out). As the random entropy increases, more diverse levels of demands are likely to be observed at a specific station.

(2) The temporal-uncorrelated entropy $S_i^{\text{unc}} = -\sum_{j=1}^{N_i} p_i(j) \log_2 p_i(j)$, which characterizes the heterogeneity of observed demand levels, where $p_i(j)$ is the probability that the jth demand level is observed among N_i unique demand levels. The temporal-uncorrelated entropy considers both the unique number of demand levels and their frequency of being observed over the timespan.

(3) The actual entropy $S_i^{\text{actual}} = -\sum_{T' \subset T_i} p(T_i') \cdot \log_2 [p(T_i')]$, which incorporates the probability of an observed demand level as well as the order in which demand levels are observed and the persistence of an observed demand level (whether the observed demand stays at a certain level for multiple time windows), where $p(T_i')$ is the probability of finding a subsequent T_i' in the observed full sequence T_i. Due to the computational complexity, the actual entropy could be estimated based on Lempel-Ziv data compression (Kontoyiannis et al., 1998).

2.2.2 Predictability

Naturally, a demand sequence with greater entropy would have more randomness in its demand pattern, which in turn decreases the predictability of future demands at this station. Considering the entropy (S) that represents the disorder of a series of demands, the upper bounds of predictability (Π) that could be attained by a predictive algorithm for correctly predicting future demands at a station is subject to Fano's inequality (Fano and Hawkins, 1961), when a station's demands with entropy S range between N distinct levels (Song et al., 2010):

$$\Pi \leqslant \Pi^{\max}(S, N), \tag{2.1}$$

where $\Pi^{\max}(S, N)$ has a relationship with S as outlined in Eq. (2.2):

$$S = H(\Pi^{\max}) + (1 - \Pi^{\max}) \log_2 (N - 1), \tag{2.2}$$

and Eq. (2.3) describes the binary entropy function:

$$H(\Pi^{\max}) = -\Pi^{\max} \log_2 (\Pi^{\max}) - (1 - \Pi^{\max}) \log_2 (1 - \Pi^{\max}). \tag{2.3}$$

By solving Eqs. (2.2) and (2.3) based on the specified N and S calculated from a station's demand series, we could compute the predictability Π. Based on three types of entropy defined in Section 2.2.1, for each station i, we can thereby denote random predictability $\Pi_i^{\text{rand}} = \Pi_i^{\text{Fano}}(S_i^{\text{rand}}, N_i)$, temporal-uncorrelated predictability $\Pi_i^{\text{unc}} = \Pi_i^{\text{Fano}}(S_i^{\text{unc}}, N_i)$, and actual predictability $\Pi_i^{\text{actual}} = \Pi_i^{\text{Fano}}(S_i^{\text{actual}}, N_i)$. Higher values of Π indicate that the demands could be better predicted. Comparing these three predictability metrics enables us to examine how temporal correlations within an individual station's demand sequence enhance potential predictive accuracy (Lu et al., 2013). This also aligns with existing works presented in Zhou (2015) and Lin et al. (2020) that incorporated different spatial and temporal variables into the prediction model with the aim to reduce prediction errors.

Since the entropy and predictability are fully subject to the time series of demands itself, we also performed a sensitivity analysis to examine how entropy and predictability of each station vary based on having different demand observation intervals (Δt). Specifically, we computed the entropy and predictability corresponding to six distinct demand observation intervals (Δt): 1 h, 2 h, 4 h, 6 h, 12 h, and 24 h. These observation intervals embody the potential granularity of performing demand monitoring and analysis by operators for each station.

2.3 Station-level temporally-constant features

The entropy and predictability of a station's demand can be measured using historical demand data. However, for planning new systems or expanding an existing system to build new stations, such historical demand data would be unavailable. To examine whether the intrinsic demand randomness is related to temporally-constant and city-specific features that are available prior to launching new stations, we adopted the temporally-constant features of individual stations in Chicago as proposed in Kou and Cai (2021b) and analyzed how such features contribute to the entropy and predictability across stations. The features are listed in Table 1 and they can be classified into three categories: Land use characterization, spatial network information of stations, and socio-demographics. Furthermore, we applied a random forest regression model to explore the importance of temporally-constant variables to the entropy and predictability of demands at each station. A random forest regression model is a machine learning model that trains multiple decision trees on different subsets of the dataset, aggregating their results to improve overall predictive performance (Biau and Scornet, 2016). We selected such a model because of its powerful regression capability as well as interpretability to the factor significance (Belgiu and Drăguţ, 2016). Specifically, the variables listed in Table 1 are treated as input variables in the model while the entropy (S^{rand}, S^{unc}, S^{actual}) and predictability (Π^{rand}, Π^{unc}, Π^{actual}) computed in Section 2.2 are dependent variables, respectively. Once a random forest regression model is trained, we can extract the feature importance according to each feature's contribution in making correct regression and reducing loss. This enables the interpretability in understanding the contribution of temporally-constant featuring in impacting the entropy and predictability of demands.

2.4 Benchmark prediction algorithms and evaluation metrics

The entropy and predictability measurements discussed in Section 2.2 could establish the theoretical upper limit of the predictive power for demand prediction algorithms. To verify the relationship between these theoretical upper limits and the practical predictive performance at the station level, we implemented two widely used benchmark algorithms for BSS demand prediction: ARMA and XGBoost. For each predictive performance (detailed in Section 2.4.3) pertaining to the predictive algorithms, we modelled their relationship and calculated R^2 scores, which shows how well such relationship holds. We also did a sensitivity analysis to examine whether such relationship sustains when using different time intervals (i.e., six time-intervals as described in Section 2.2 were analyzed), examining their fitted curves and associated R^2 scores. Please note that we rounded the output of prediction to the nearest integer as the demand level (i.e., the number of bikes being checked-out) can only be integers.

2.4.1 ARMA

The ARMA algorithm is a widely recognized statistical method for time series forecasting, which has been applied to conduct demand prediction in various applications, such as stock price, inventory and disease infection (Saboia, 1977; Nochai and Nochai, 2006; Chen et al., 2008; Benvenuto et al., 2020). The ARMA model consists of two main parameters: p (order of autoregression (AR)) and q (order of moving average (MA)), denoted as $ARMA(p, q)$, which is equivalent to an ARIMA model without differencing a time series. The future demand $X_{i,t}^{\text{pred}}$ to be forecasted at a specific station i at time step t is a linear combination of past demands and past errors, formulated as follows:

$$X_{i,t}^{\text{pred}} = \varphi_{i,1}X_{i,t-1} + \varphi_{i,2}X_{i,t-2} + \cdots + \varphi_{i,m}X_{i,t-m} + \cdots + \varphi_{i,p}X_{i,t-p} + e_{i,t} - \theta_{i,1}e_{i,t-1} - \cdots - \theta_{i,n}e_{i,t-n} - \cdots - \theta_{i,q}e_{i,t-q}, \quad (2.4)$$

where $\varphi_{i,m}$ and $\theta_{i,n}$ are the coefficients, $e_{i,t-n}$ is the random noise with $E[e_{i,t-n}] = 0$ and variance σ^2, and p and q are the orders of AR and MA polynomials, respectively. In our study, individual stations had their own ARMA models and models were trained on stations' historical check-out demand data (i.e., the training dataset introduced in Section 2.1) separately, so the ARMA model is station

specific. In addition, the parameters of each $ARMA(p, q)$ model were determined through automatic ARMA algorithm (Hyndman and Khandakar, 2008).

2.4.2 XGBoost

XGBoost is a tree boosting machine learning algorithm which is able to handle high-dimension data with selected features (Chen and Guestrin, 2016). Within a XGBoost model, the output forecasted demand $X_{i,t}^{\text{pred}}$ at station i at time step t is predicted by using K additive functions in a tree-ensemble model:

$$X_{i,t}^{\text{pred}} = \Phi(x_{i,t}) = \sum_{k=1}^{K} f_k(x_{i,t}), \quad f_k \in F, \tag{2.5}$$

where F is the space of regression trees; each f_k represents an independent tree structure within K additive functions; and $x_{i,t}$ corresponds to the set of input variables for station i at time step t. To be more specific, the input variables featuring each station include both time-invariant variables described in Table 1 and weather variables (average temperature, average humidity, average wind speed, average precipitation, and average pressure).

2.4.3 Evaluation metrics for prediction algorithms

The models introduced in Sections 2.4.1 and 2.4.2 were trained on the training dataset and then their prediction performances were evaluated based on the testing dataset. We selected Root Mean Square Error (RMSE) and Cumulative Scores (CS) as the performance metrics, which represent the absolute and relative errors, respectively. Specifically, RMSE is defined as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(X_{i,t}^{\text{pred}} - X_{i,t}^{\text{true}}\right)^2}. \tag{2.6}$$

The value of RMSE ranges from 0 to $+\infty$. A lower RMSE score indicates a better fitted regression model while a higher RMSE indicates a large prediction error.

Because RMSE could be highly influenced by the stations' capacity (larger stations tend to have higher RMSE), we further employed CS adopted from Kocer (2013) and Niu et al. (2016) as a measurement of relative prediction accuracy. The value of CS ranges from 0 to 1 and is defined as:

$$CS(\eta) = D_\eta / D, \tag{2.7}$$

where D is the total number of tested time steps for each station, and D_η refers to the number of predicted demand values whose predicted demands do not exceed the lower and upper percentage bound η (%) of the ground truth demand. In this study, we set the percentage bound η as 10%. A CS value closer to 1 signifies a higher prediction accuracy.

3 Results and discussions

3.1 Overview of station-level demands, entropy, and predictability

Figure 1(a) shows the distribution of demand levels X_i observed in each 4-hour time interval at individual stations during the study period of 361 days. Aligning with previous studies, e.g., Kou and Cai (2019), the distribution of demand is heavy tailed. This indicates that within each 4-hour window, the majority of stations have low check-out demand while the number of stations with high-volume demands is much smaller.

However, low demands do not necessarily mean low entropy or high predictability. We calculated the entropy and predictability across all stations based on their demand sequences in the study period. The resulting distribution for S^{rand}, S^{unc}, and S^{actual}, as well as Π^{rand}, Π^{unc}, and Π^{actual} are illustrated in Fig. 1(b) and 1(c), respectively. The S^{unc} encodes the additional frequency information of demands compared to the S^{rand} which only considers the number of unique demand levels N_i. Such additional information explains partial randomness rooted in the time series of demands, thus $P(S^{\text{unc}})$ exhibits left shift (i.e., lower entropy in general) comparing with $P(S^{\text{rand}})$. Likely, S^{actual} encodes the additional temporal order information on top of the S^{unc}, so $P(S^{\text{actual}})$ is more left shifted than $P(S^{\text{unc}})$. With $P(S^{\text{rand}})$ peaking at $S^{\text{rand}} \approx 4$, one can expect that, if users come to rent bikes at stations randomly, $2^{S^{\text{rand}}} \approx 16$ different demand levels can be observed in a station on average. In contrast, the actual entropy reveals the real uncertainty considering the sequence of demands that could be observed at a station. Surprisingly, the actual entropy S^{actual} has two major peaks at 0.1 and 2.5, respectively. For stations with $S^{\text{actual}} = 0.1$, they almost have no demand uncertainty. The next demand that can be forecasted at those stations is always at $2^{0.1} \approx 1$ level, suggesting that these stations are likely to have no check-out demand but could randomly witness one check-out demand. While for stations with $S^{\text{actual}} = 2.5$, their future demands are likely to be at $2^{2.5} \approx 5.66$ (i.e., less than six) random levels. A consistent conclusion is observed in terms of predictability. We find that $\Pi^{\text{actual}} > \Pi^{\text{unc}} > \Pi^{\text{rand}}$ for demands at individual stations on average as more information is incorporated. With only the information of number of unique demand levels, the majority of stations exhibit very low predictability. As shown in Fig. 1(c), the distribution of actual predictability has two peaks. One group of stations has nearly perfect predictability for demands, which aligns with our observation that these stations almost have no check-out demands. The other group of stations has predictability around 0.65. In other words, no matter how good the predictive algorithm is, the future demand levels of stations with $\Pi^{\text{actual}} = 0.65$ has maximum prediction

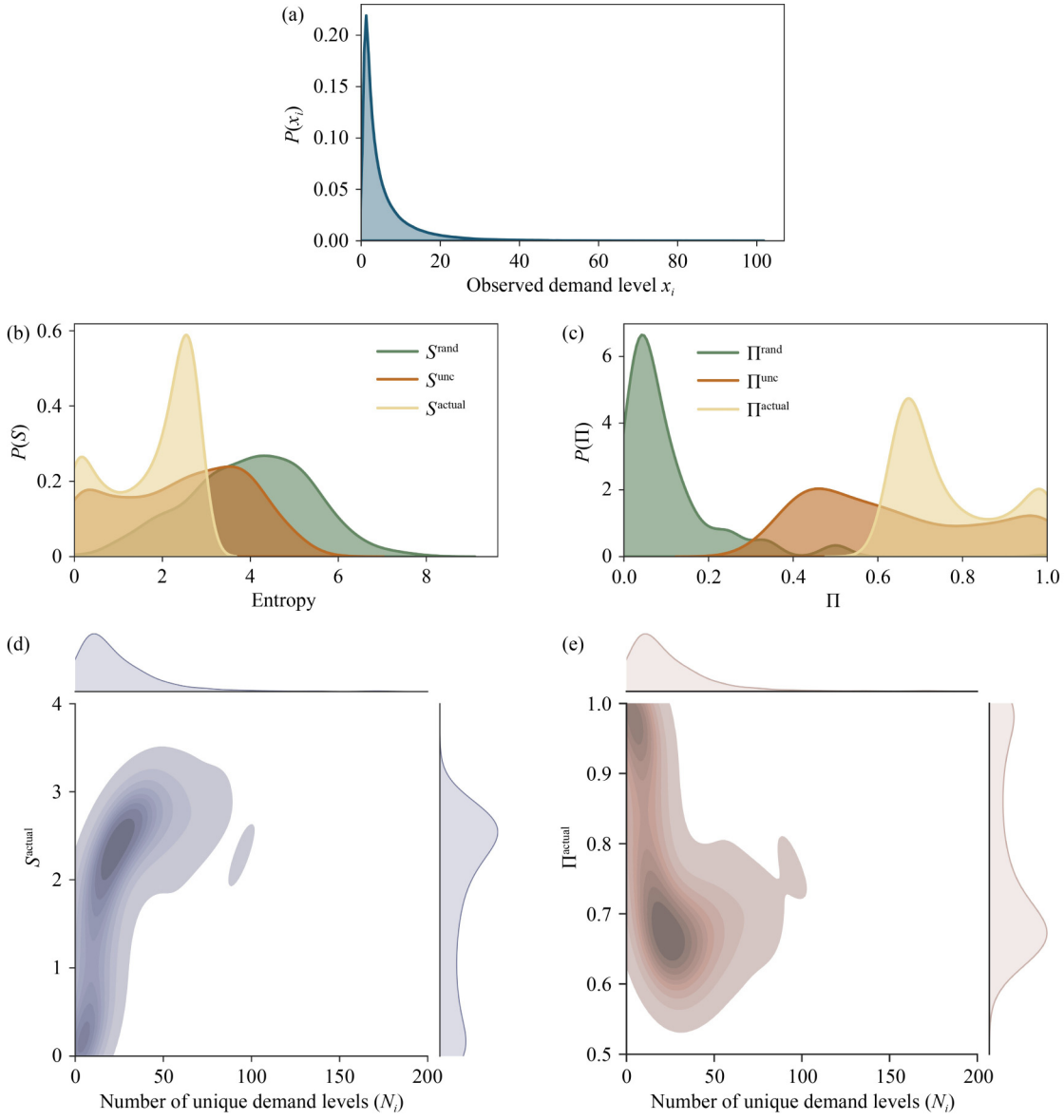

Fig. 1 Overview of a station's demands and its entropy and predictability. (a) The distribution of demand levels observed across all stations during the study period. (b) The distributions of random entropy, temporal-uncorrelated entropy, and actual entropy. (c) The distributions of random predictability, temporal-uncorrelated predictability, and actual predictability. (d) The joint distribution of the number of unique demand levels observed at stations (N_i) and the actual entropy (S^{actual}), showing the correlation between N_i and S^{actual}. (e) The joint distribution of the number of unique demand levels observed at stations (N_i) and the actual predictability (Π^{actual}), showing the correlation between N_i and Π^{actual}.

accuracy of 65%. Therefore, Π^{actual} serves as the intrinsic limit for predictability of individual station. Notably, the two-peaked distribution of S^{actual} and Π^{actual} indicates a polarized demand predictability rooted in stations. Despite that the overall station demands are very low across stations, a significant number of stations exhibit high entropy and reduced predictability. This implies that system operators should pay particular attention to these stations.

We depicted Figs. 1(d) and 1(e) to further reconcile the contradiction that, overall, stations have low observed demands but with two peaks in the distributions of S^{actual} and Π^{actual}. For the group of stations with low predictability (high entropy), they span a broad spectrum of unique levels of historically observed demands from 0 to approximately 100. In particular, the peak corresponds to an average of 30 unique demand levels. Contrastingly, the stations associated with high predictability (low entropy) span a narrower range of unique levels from 0 to around 25, with an average of 1. Therefore, while both groups of stations could experience low level of demands, the stations within the high entropy group can be anticipated to present a broader spectrum of demand variation. The increased variation of demand spectrum at individual

stations introduces more randomness into the time series of observed demands, which inherently leads to higher entropy and lower predictability. In addition, the contrast of polarized stations is also illustrated in Fig. 2, which displays the overall spatial distribution of stations with varying S^{actual} and Π^{actual}. Stations with high S^{actual} but low Π^{actual} (smaller circles in dark red) mainly locate within the downtown area whereas stations with low S^{actual} but high Π^{actual} (larger circles in dark blue) mostly distribute in the suburban regions.

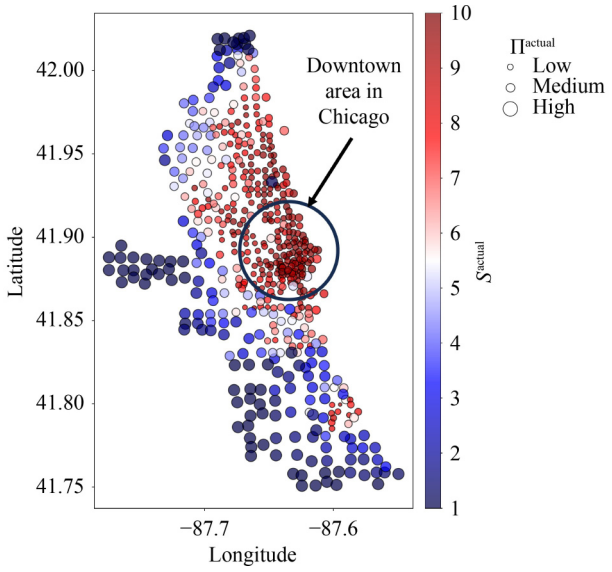

Fig. 2 The distribution of stations with different actual entropy S^{actual} and actual predictability Π^{actual} levels in Chicago (each circle represents a station).

3.2 Relationship between entropy/predictability and prediction performance

In this section, we confirm that the entropy and predictability, which measures the randomness rooted in a station's demands, are feasible and model-free estimators for empirical prediction errors and accuracy of demand prediction models, respectively.

In Figs. 3(a)–3(f), the RMSE of each data point represents the average RMSE over the testing horizon of a single station derived from a single predictive algorithm. The results reveal the presence of an exponential association between the entropy over one year's historical demands at individual stations and the resulting RMSE values from future demand prediction. This relationship shows an $R^2 = 0.94$ for ARMA and an $R^2 = 0.88$ for XGBoost when correlating S^{rand} and RMSE. Likewise, Figs. 3(g)–3(l) demonstrate a logistic relationship established between the predictability of demands and the CS values, where R^2 ranges from 0.81 to 0.87 for ARMA between CS and Π while R^2 ranges from 0.59 to 0.74 for XGBoost between CS and Π. Note that we paired RMSE with entropy and CS with predictability because each of the paired values share the same unit of measurement. RMSE and entropy are expressed in terms of absolute magnitude of errors, while CS and predictability are expressed as the percentage scale of accuracy.

Figures 3(i) and 3(l) validate predictability as a theoretical limit for prediction accuracy, which is consistent with the upper limit identified in Lu et al. (2013). Entropy and predictability capture the theoretical limits for the predictive analysis of BSS demand prediction, offering an approachable upper bound of predictive power for such BSS demand data. For instance, for the group of stations with average $\Pi^{\text{actual}} = 0.65$, their CS values stay below 0.5 in both ARMA and XGBoost models. Even though some stations are expected to reach maximum $\Pi^{\text{actual}} = 0.8$, they still exhibit practical CS values under 0.5, verifying that Π^{actual} bounds the empirical prediction accuracy. This is in line with data point distribution shown in Figs. 3(i) and 3(l) that corresponding practical CS from predictive models is always smaller than the theoretical predictability. That being said, by applying the same demand prediction model across all stations and regressing models on certain variables to forecast station-level demands, it can be anticipated that various prediction errors across different stations are bounded by their inherent predictability. Furthermore, it is worth noting that entropy and predictability are model-free estimators, as they are derived exclusively from historical demands at individual stations. Their calculations are independent of any predictive models and serves as *a priori* measurements in lieu of extensive domain-specific feature engineering for the input variables of the prediction models.

3.3 The impact of features on entropy and predictability

After verifying that entropy and predictability were feasible and model-free representations of practical prediction effects, we then identified which temporally-invariant factors are most significant contributors to entropy and predictability of station demands. In this work, we ran the random forest regression model whose inputs are temporally-constant variables listed in Table 1 and whose outputs are either the set of entropy (S^{rand}, S^{unc}, S^{actual}) or the set of predictability (Π^{rand}, Π^{unc}, Π^{actual}). Based on five-fold cross validation, our results show that the model achieved an RMSE of 0.66 for entropy and 0.83 for predictability. This highlights that a statistically significant relationship holds between the temporally-constant factors and the entropy/predictability of a station's demands.

In Table 2, we present three most significant temporally-constant attributes affecting entropy and predictability, respectively (see the full list of ranked feature importance in Supplementary Information Tables A1 and A2). Our findings indicate that per capita income, spatial eccentricity, and the number of parking lots hold significant

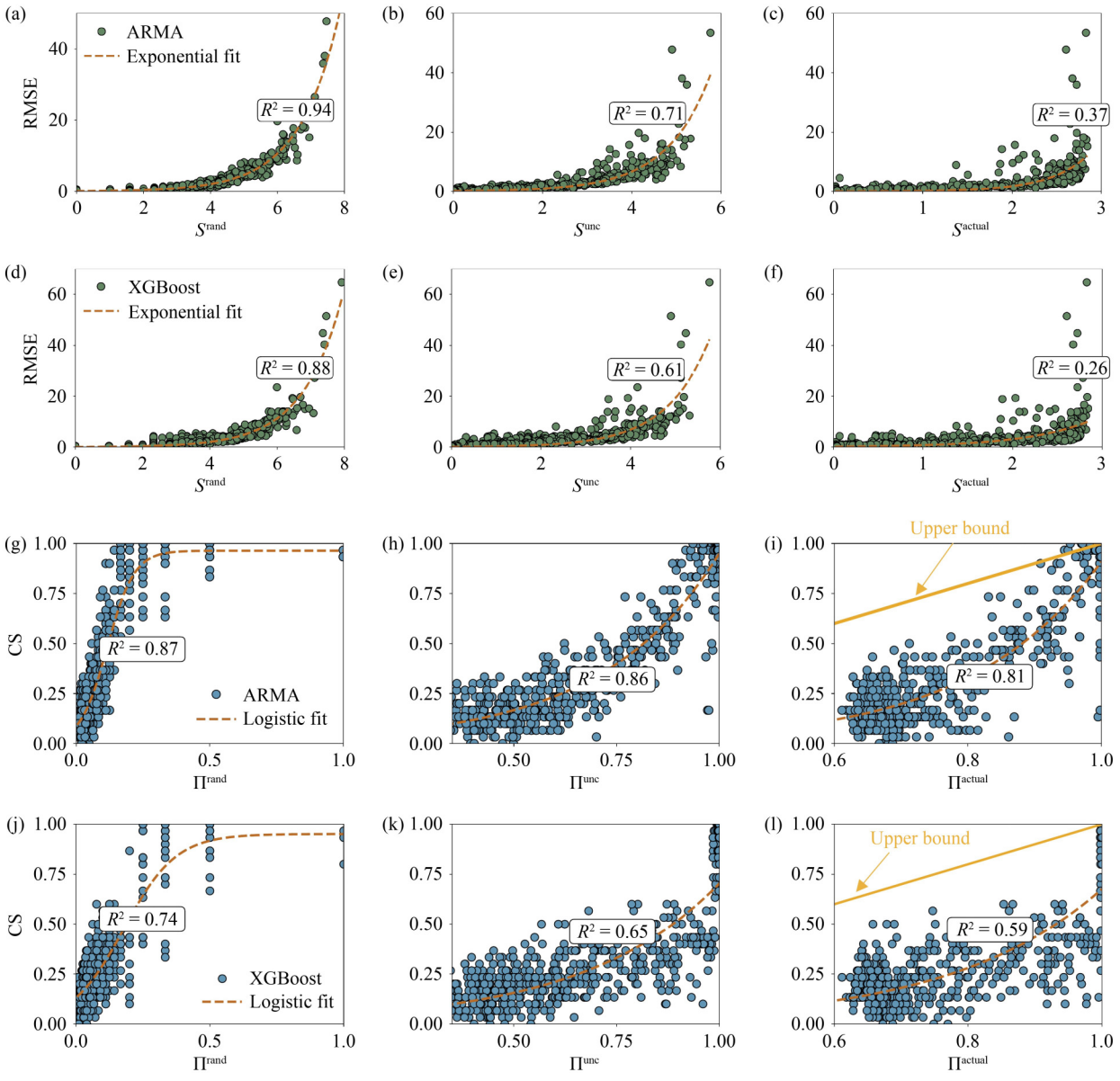

Fig. 3 (a)–(f) present correlation between entropy and RMSE; (g)–(l) present correlation between predictability and CS. The orange bound lines in (i) and (l) show that the practical CS from predictive models for each station stayed below its theoretical predictability.

Table 2 Significant temporally-constant variables and their feature importance in the models for entropy/predictability

Variable	Feature importance	
	Entropy	Predictability
Per capita income	0.542	0.494
Spatial eccentricity	0.113	0.170
Parking lot	0.107	0.0916

importance for both entropy and predictability. Notably, each of these top three variables comes from a different variable category: Per capita income is a socio-demographic variable; spatial eccentricity is a spatial network variable; and the number of parking lots is a land use variable.

In Fig. 4, we showcase the specific quantities of these three temporally-constant variables identified above for the top ten stations with the highest and lowest actual predictability Π^{actual}, respectively, in order to examine the positive/negative impact of variables on entropy and predictability. One can observe that per capita income negatively impacts predictability since stations with lower predictability exhibit higher per capita income. Similarly, the number of parking lots negatively influences the predictability. On the other hand, spatial eccentricity positively influences predictability, with more eccentric stations exhibiting higher predictability. Furthermore, in Fig. 5, we depict the spatial distribution of stations in the entire systems and their associated three temporally-constant variables, showing the spatial interplay between

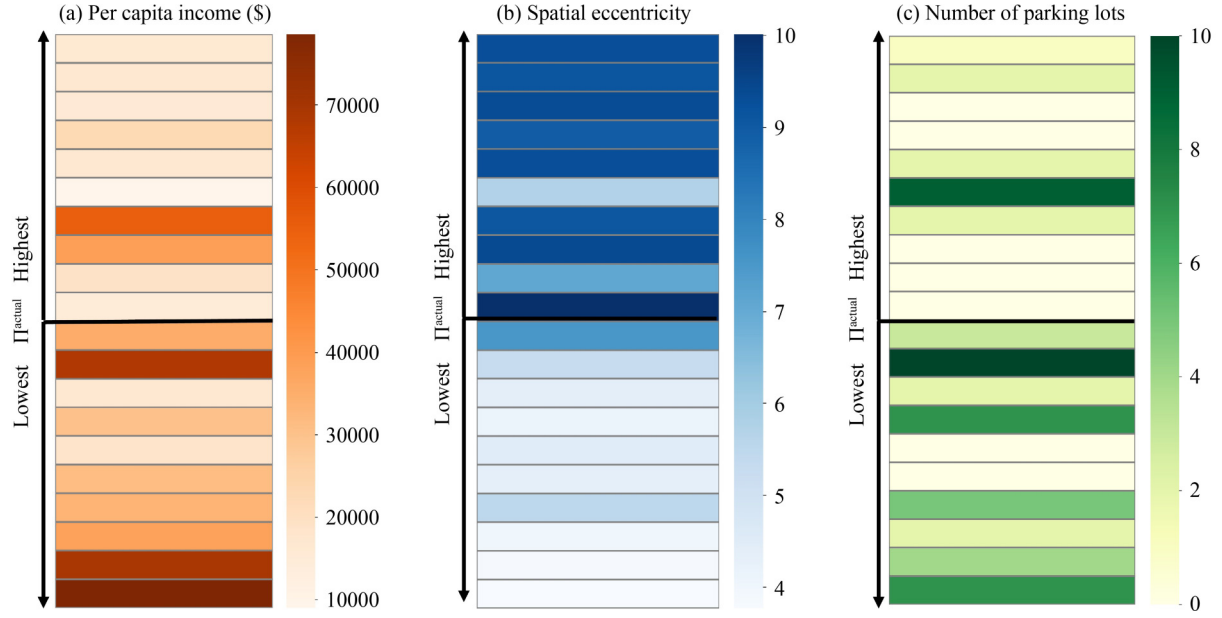

Fig. 4 The (a) per capita income, (b) spatial eccentricity, and (c) parking lot attributes of stations with top 10 highest and lowest actual predictability (Π^{actual}).

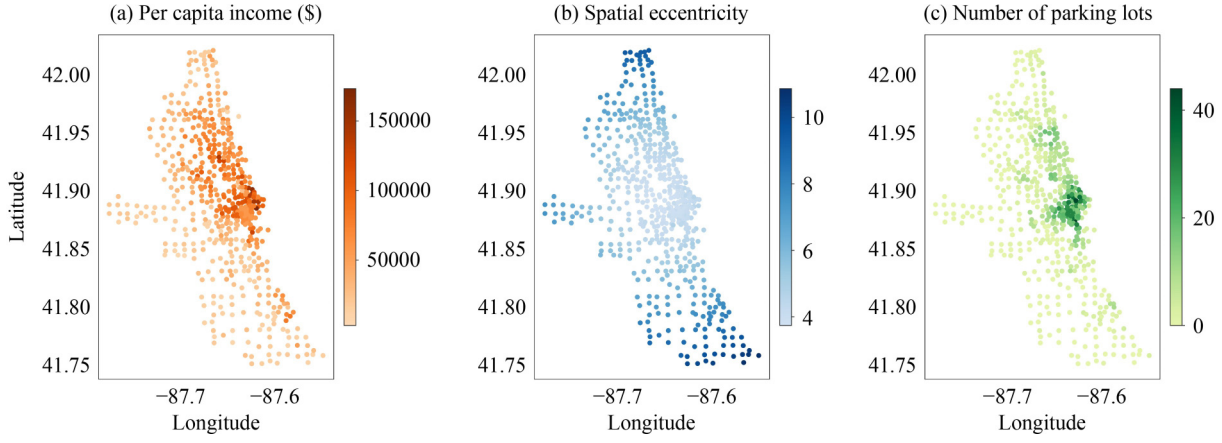

Fig. 5 The spatial distribution of stations in terms of the attributes of (a) per capita income, (b) spatial eccentricity, and (c) parking lots.

entropy and predictability and three variables across the service region. For per capita income (Fig. 5(a)), we observed that stations exhibiting higher predictability and lower entropy as shown in Fig. 2 are typically located in suburban areas with lower per capita income. Conversely, stations with lower predictability are often found in downtown area with higher per capita income. In terms of spatial eccentricity, one can observe that in suburban areas, stations with higher predictable demands are sparsely located from their neighbouring stations. In contrast, stations with less predictable demands tend to be concentrated in the urban center, where stations are denser. Additionally, for parking lot, stations exhibiting lower predictability have more parking lots surrounding them in the downtown compared with their more predictable, lower entropy counterparts.

The significant city- or system-specific, temporally-constant variables impacting entropy and predictability can be interpreted in several ways. From the significance of per capita income, we can infer that in census tracts with higher wealth, more stations could serve leisure-based or recreational trips (Stromberg, 2015), which would induce more entropy compared to stations that serve regular commute- or errand-based trips. On the other hand, high entropy in high-income areas could be a symptom of the accessibility to various transportation options in high per capita income (Smith et al., 2020), inducing a higher variety of trips and options in those places. The spatial eccentricity implies that those stations located farther from other stations are more predictable. That is, demands of suburban stations are generally more predictable. This could be the result of suburban places

having lower accessibility to diverse trip purposes, or it may be representative of low overall use in suburban stations. The quantity of parking spaces may signal an area's capacity to accommodate various trip purposes. Figure 5(c) reveals that parking lots mainly concentrate in the downtown region. Considering the functionality of BSS in first/last mile trips, customers are able to park their private cars in the parking spaces (e.g., public parking lot and street parking) and continue the subsequent trips by using shared bikes in the downtown area. Such a capacity of hosting various trips increases the variation of number of trips that can be anticipated in the downtown area, which leads to high entropy and in turn reduces the predictability of demands.

3.4 Sensitivity analysis

3.4.1 Entropy and predictability under different observation intervals

Entropy and predictability distribute differently under different demand observation intervals. Figure 6 suggests the distinct effects of demand observation frequencies: Increasing the frequency of demand observations enhances the predictability and decreases the entropy of those stations initially characterized by low predictability. However, this increase in observational frequency appears to have no impact on the stations that already exhibit high levels of predictability. More specifically, in Figs. 6(a) and 6(d), as the frequency of observation lowers, stations tend to have higher S^{rand} and lower Π^{rand}, respectively. This is because demands cumulate to higher levels when less frequently monitored. For instance, the demand level of one unit over two consecutive hours could aggregate to the demand level recorded for two units within a two-hour observation window. This is also in line with the distribution of $P(X_i)$ displayed in Fig. 6(g), where the peak slightly shifts towards zero and the variance of distribution decreases when implementing more frequent monitoring. Moreover, in Figs. 6(b), 6(c), 6(e) and 6(f), one can observe that the peaks of high predictability stations (i.e., low entropy stations) stays aligned regardless of variations in the monitoring frequency. This is because these stations are typically associated with very low demand levels (typically zero or one), resulting in unvarying time series of demands even though the monitoring changes drastically from hourly to daily. In contrast, it is also noticeable that the peaks

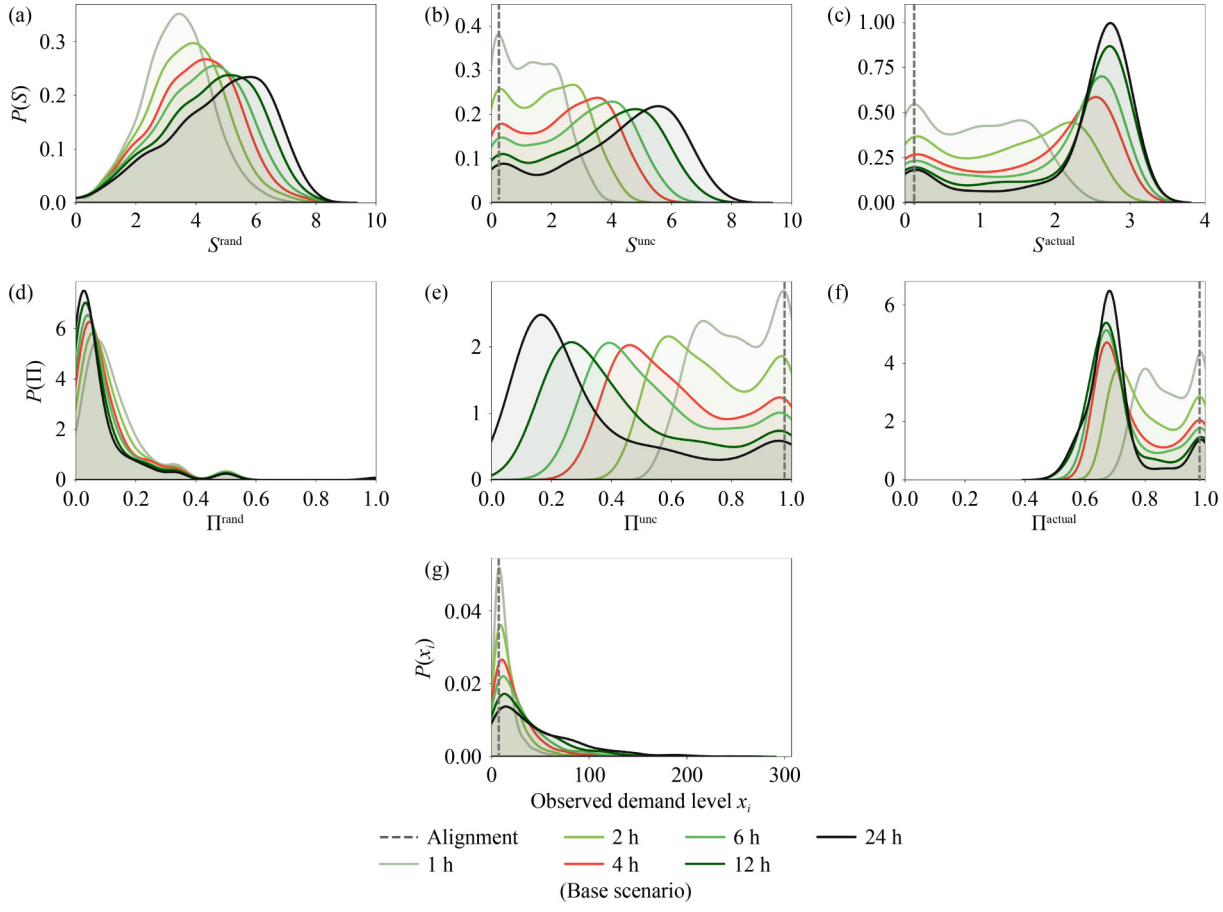

Fig. 6 The (a)–(c) entropies, (d)–(f) predictabilities, as well as (g) demand level distribution under different demand observation intervals.

of low predictability stations shift towards higher predictability when de-escalating the monitoring interval from 24 hours to 1 hour.

This sensitivity analysis also delivers implication of potential operational strategies to mitigate the low predictability associated with certain stations. Operators could heterogeneously monitor the demands across stations situated in various locations: The downtown area may necessitate more frequent monitoring (such as every half an hour or every hour); conversely, suburban areas may require less frequent observations (such as every 12 hours or every day). By shortening the time interval between demand observations, the time series of demands tend to be less random so operators can anticipate more accurate predictions with less errors in their predictive models. This strategy also suggests a practical way to balance operational resources for different service regions while keeping service reliability and predictability.

3.4.2 Relationship between theoretical upper limits and practical predictive performances under different observation intervals

Our sensitivity analysis solidifies the relationship between entropy/predictability (theoretical upper limits) and RMSE/CS (practical predictive performances) identified in Section 3.2. Figure 7 presents the fitted curves under different observation intervals for both predictive algorithms with both performance metrics and Table 3 lists the associated R^2 score of each fitted curve in Fig. 7. Same as Fig. 3, RMSE and entropy follows an exponential relationship while the relationship between CS and

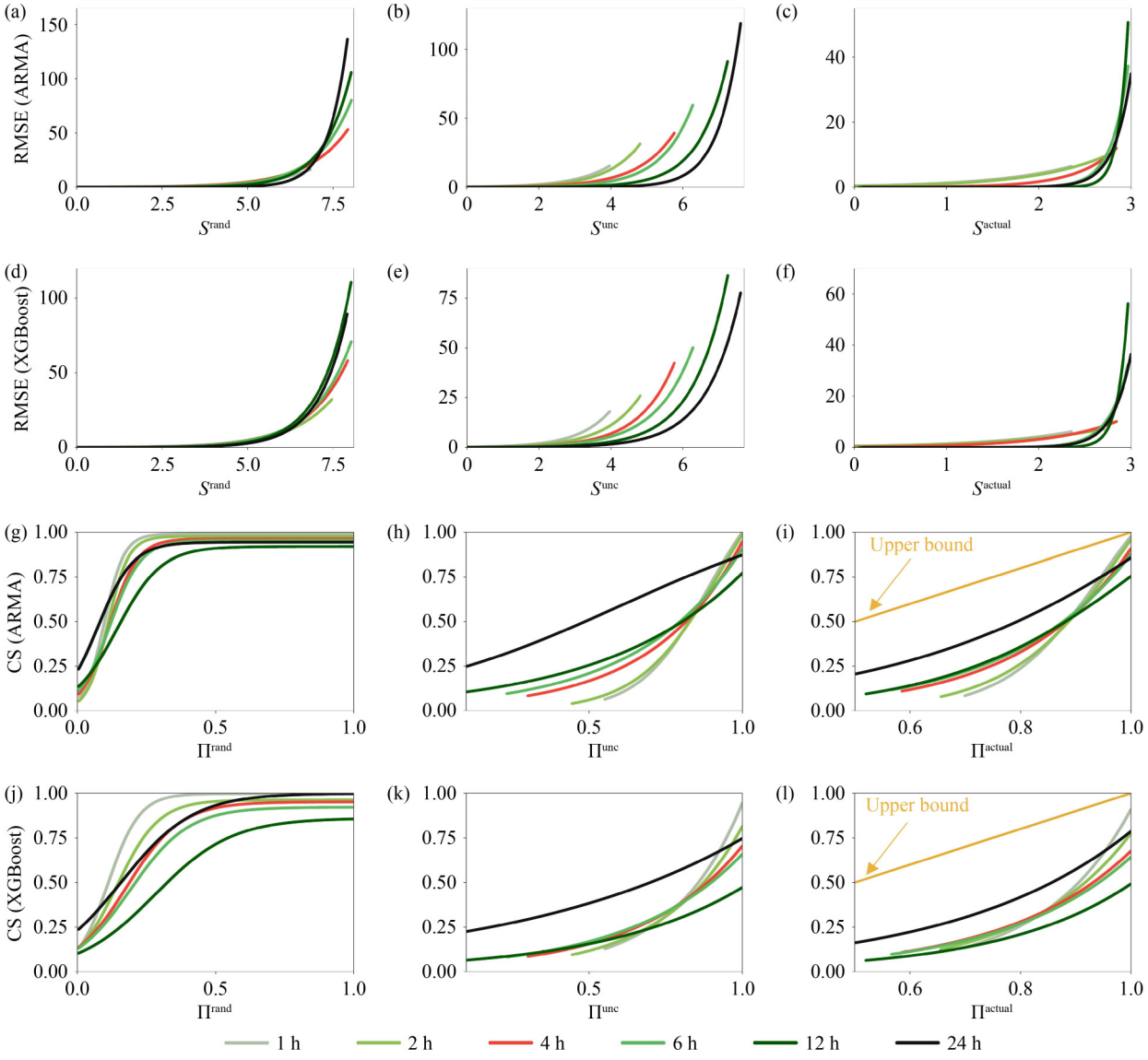

Fig. 7 Relationship between (a)–(c) RMSE of ARMA and entropy, (d)–(f) RMSE of XGBoost and entropy, (g)–(i) CS of ARMA and predictability, and (j)–(l) CS of XGBoost and predictability under different demand observation intervals.

Table 3 R^2 scores of fitted curves for Fig. 7

Observation interval	RMSE of ARMA for Figs. 7(a–c)			RMSE of XGBoost for Figs. 7(d–f)			CS of ARMA for Figs. 7(g–i)			CS of XGBoost for Figs. 7(j–l)		
	S^{rand}	S^{unc}	S^{actual}	S^{rand}	S^{unc}	S^{actual}	S^{rand}	S^{unc}	S^{actual}	S^{rand}	S^{unc}	S^{actual}
1 h	0.93	0.86	0.52	0.90	0.76	0.38	0.90	0.94	0.87	0.77	0.78	0.71
2 h	0.95	0.79	0.42	0.87	0.66	0.31	0.89	0.92	0.87	0.82	0.78	0.71
4 h	0.94	0.71	0.37	0.88	0.61	0.26	0.87	0.86	0.81	0.74	0.65	0.59
6 h	0.91	0.63	0.38	0.81	0.56	0.31	0.83	0.80	0.75	0.67	0.58	0.55
12 h	0.81	0.65	0.42	0.75	0.57	0.40	0.61	0.55	0.55	0.41	0.27	0.32
24 h	0.70	0.60	0.26	0.67	0.58	0.34	0.47	0.42	0.29	0.32	0.29	0.31

predictability is logistic. When monitoring the demands more frequently (the time interval is smaller), the R^2 score becomes higher for the corresponding fitted curve. This is because for a given fixed period (i.e., 366 days in this work), the observation with higher frequency results in more data points for both training and testing the model and the model gets better trained.

4 Conclusions

In summary, this study measured the inherent randomness of station demands via entropy and predictability based on information theory. We validated those metrics against empirical performance metrics to show that entropy and predictability are applicable to function as model-free estimators to anticipate the error and accuracy in prediction models. Furthermore, we identified three most significant temporally-constant factors — per capita income, spatial eccentricity, and the number of parking lots — that influence station demand entropy and predictability. The results we obtained from this study can serve as *a priori* implication for BSS management, operation, expansion, and launch, without the need of domain-specific feature engineering and modelling demand prediction algorithms.

For cities with active BSSs, our findings imply that there are inherent errors and randomness that existing predictive models might fail to capture, so the operators could apply customized monitoring strategies in different service regions to escalate the demand predictability and service reliability. For instance, for stations with high entropy and low predictability, monitoring and maintenance (e.g., rebalancing) can be employed more frequently so as to improve operational functionality. On the other hand, for cities considering an expansion or investment in a BSS, the identified temporally-invariant key factors associated with high entropy stations can be qualitatively taken into account when deciding where to add new stations to the BSS. Moreover, measuring the entropy and predictability at station-level allows for a finer resolution of operation among stations and offers the possibility to enhance the specificity of management tailored for each station. Thus, these results have direct implications for operators to make decisions about station monitoring, rebalancing, capacity design, siting, or decommissioning.

In a broader context, for cities without BSSs, the municipal officials and potential system operators can utilize entropy and predictability to link temporally-constant city and system characteristics to potential demand prediction efficacy. These *a priori* measurements can help inform future planning and designing of the BSSs as well as corresponding infrastructure such as parking lots and bike lanes in a city. Our model is also adaptable so when there are significant changes occurred for temporally-constant city characteristics, operators can re-implement it to update their planning decisions for the next stage.

Suggested next stages of this research are 1) to perform a large-scale analysis to generalize the entropy and predictability in other cities and 2) to compare entropy and predictability to other evaluation metrics (which could emphasize prediction error by how often rather than how much) with more predictive algorithms. In the first case, BSS demand prediction performance could be heavily influenced by the number of high-entropy stations in a specific system, thus case study selection and generalization of results must consider this aspect of BSS. A large-scale analysis could help determine generalization possibilities of these methods and provide more insight on the dynamics of individual BSS throughout the world. In the second case, it could be just as valuable to know how often demand prediction is wrong as by how much the demand prediction was wrong, especially for dynamic rebalancing of bikes. Therefore, an evaluation metric that captures the temporal prediction error of demand could be useful in relating to the different forms of entropy and predictability calculated herein. Furthermore, the exponential and logistic association found in our results in Section 3.2 were slightly different from the linear association identified in the literature on entropy. Investigating the reason for this is a direction for future research.

One challenge in BSS research regarding demand prediction has been the absence of trip purpose data. Without knowing the purpose of people's trips at bike sharing stations, it has been more difficult to predict the demand. High entropy at stations could suggest a diversity

and heterogeneity of trip purposes, whereas stations with low entropy could mean consistent travellers using BSS for similar trip purposes. Future research could explore how entropy is related to trip purpose for predicting demand, when trip purpose data become available. Furthermore, future research can explore in more detail the operational costs of high- versus low-entropy stations. If a relationship between entropy measurement and maintenance costs of a station exists, our method could be employed to anticipate the operational costs of adding new stations or, in the case of cities without active BSS yet, investing in a BSS.

Electronic Supplementary Material Supplementary material is available in the online version of this article at https://doi.org/10.1007/s42524-023-0279-8 and is accessible for authorized users.

Competing Interests The authors declare that they have no competing interests.

References

Bachand-Marleau J, Lee B, El-Geneidy A (2012). Better understanding of factors influencing likelihood of using shared bicycle systems and frequency of use. Transportation Research Record: Journal of the Transportation Research Board, 2314(1): 66–71

Bao J, Xu C, Liu P, Wang W (2017). Exploring bikesharing travel patterns and trip purposes using smart card data and online point of interests. Networks and Spatial Economics, 17(4): 1231–1253

Belgiu M, Drăguţ L (2016). Random forest in remote sensing: A review of applications and future directions. ISPRS Journal of Photogrammetry and Remote Sensing, 114: 24–31

Benvenuto D, Giovanetti M, Vassallo L, Angeletti S, Ciccozzi M (2020). Application of the ARIMA model on the COVID-2019 epidemic dataset. Data in Brief, 29: 105340

Biau G, Scornet E (2016). A random forest guided tour. Test, 25(2): 197–227

Cazabet R, Borgnat P, Jensen P (2017). Using degree constrained gravity null-models to understand the structure of journeys' networks in bicycle sharing systems. In: 25th European Symposium on Artificial Neural Networks, Computational Intelligence, and Machine Learning. Bruges: ENSANN, 25

Chai D, Wang L, Yang Q (2018). Bike flow prediction with multi-graph convolutional networks. In: Proceedings of the 26th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. Seattle, WA: ACM, 397–400

Chen L, Zhang D, Wang L, Yang D, Ma X, Li S, Wu Z, Pan G, Nguyen T, Jakubowicz J (2016). Dynamic cluster-based over-demand prediction in bike sharing systems. In: Proceedings of the International Joint Conference on Pervasive and Ubiquitous Computing. Heidelberg: ACM, 841–852

Chen P, Yuan H, Shu X (2008). Forecasting crime using the ARIMA model. In: 5th International Conference on Fuzzy Systems and Knowledge Discovery. Jinan: IEEE, 627–630

Chen T, Guestrin C (2016). Xgboost: A scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco, CA: ACM, 785–794

El-Assi W, Salah Mahmoud M, Nurul Habib K (2017). Effects of built environment and weather on bike sharing demand: A station level analysis of commercial bike sharing in Toronto. Transportation, 44(3): 589–613

El Sibai R, Chabchoub Y, Fricker C (2018). Using spatial outliers detection to assess balancing mechanisms in bike sharing systems. In: 32nd International Conference on Advanced Information Networking and Applications (AINA). Krakow: IEEE, 988–995

Fano R, Hawkins D (1961). Transmission of information: A statistical theory of communication. American Journal of Physics, 29(11): 793–794

Fishman E (2016). Bikeshare: A review of recent literature. Transport Reviews, 36(1): 92–113

Fishman E, Allan V (2019). Chapter six: Bike share. In: Fishman E, ed. Advances in Transport Policy and Planning. Volume 4. The Sharing Economy and the Relevance for Transport. Cambridge: Elsevier Inc., 121–152

He S, Shin K (2020). Towards fine-grained flow forecasting: A graph attention approach for bike sharing systems. In: Proceedings of the Web Conference. Taipei: ACM, 88–98

Hulot P, Aloise D, Jena S (2018). Towards station-level demand prediction for effective rebalancing in bike-sharing systems. In: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. London: ACM, 378–386

Hyland M, Hong Z, Pinto H, Chen Y (2018). Hybrid cluster-regression approach to model bikeshare station usage. Transportation Research Part A: Policy and Practice, 115: 71–89

Hyndman R, Khandakar Y (2008). Automatic time series forecasting: The forecast package for R. Journal of Statistical Software, 27(3): 1–22

Jia W, Tan Y, Liu L, Li J, Zhang H, Zhao K (2019). Hierarchical prediction based on two-level Gaussian mixture model clustering for bike-sharing system. Knowledge-Based Systems, 178: 84–97

Kim D, Cho Y, Kim D, Park C, Choo J (2022). Residual correction in real-time traffic forecasting. In: Proceedings of the 31st ACM International Conference on Information & Knowledge Management. Atlanta, GA: ACM, 962–971

Kocer U U (2013). Forecasting intermittent demand by Markov chain model. International Journal of Innovative Computing, Information & Control, 9(8): 3307–3318

Kontoyiannis I, Algoet P, Suhov Y, Wyner A (1998). Nonparametric entropy estimation for stationary processes and random fields, with applications to English text. IEEE Transactions on Information Theory, 44(3): 1319–1327

Kou Z, Cai H (2019). Understanding bike sharing travel patterns: An analysis of trip data from eight cities. Physica A, 515: 785–797

Kou Z, Cai H (2021a). Comparing the performance of different types of bike share systems. Transportation Research Part D: Transport and Environment, 94: 102823

Kou Z, Cai H (2021b). Incorporating spatial network information to improve demand prediction for bike share system expansion. In: Proceedings of the 10th International Workshop on Urban Computing. Beijing

Kou Z, Wang X, Chiu S, Cai H (2020). Quantifying greenhouse gas emissions reduction from bike share systems: A model considering real-world trips and transportation mode choice patterns. Resources, Conservation and Recycling, 153: 104534

Li Y, Zheng Y (2020). Citywide bike usage prediction in a bike-sharing system. IEEE Transactions on Knowledge and Data Engineering, 32(6): 1079–1091

Li Y, Zheng Y, Zhang H, Chen L (2015). Traffic prediction in a bike-sharing system. In: Proceedings of the 23rd SIGSPATIAL International Conference on Advances in Geographic Information Systems. Seattle, WA: ACM, 1–10

Lin L, He Z, Peeta S (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: 258–276

Lin P, Weng J, Hu S, Alivanistos D, Li X, Yin B (2020). Revealing spatio-temporal patterns and influencing factors of dockless bike sharing demand. IEEE Access, 8: 66139–66149

Liu C, Gao X, Wang X (2022). Data adaptive functional outlier detection: Analysis of the Paris bike sharing system data. Information Sciences, 602: 13–42

Lu X, Wetter E, Bharti N, Tatem A, Bengtsson L (2013). Approaching the limit of predictability in human mobility. Scientific Reports, 3(1): 2923

Luo H, Zhao F, Chen W, Cai H (2020). Optimizing bike sharing systems from the life cycle greenhouse gas emissions perspective. Transportation Research Part C: Emerging Technologies, 117: 102705

Médard de Chardon C, Caruso G (2015). Estimating bike-share trips using station level data. Transportation Research Part B: Methodological, 78: 260–279

Niu Z, Zhou M, Wang L, Gao X, Hua G (2016). Ordinal regression with multiple output CNN for age estimation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV: IEEE, 4920–4928

Nochai R, Nochai T (2006). ARIMA model for forecasting oil palm price. In: Proceedings of the 2nd IMT-GT Regional Conference on Mathematics, Statistics and Applications. Penang: Academia, 13–15

Regue R, Recker W (2014). Proactive vehicle routing with inferred demand to solve the bikesharing rebalancing problem. Transportation Research Part E: Logistics and Transportation Review, 72: 192–209

Saboia J L M (1977). Autoregressive integrated moving average (ARIMA) models for birth forecasting. Journal of the American Statistical Association, 72(358): 264–270

Senter H (2008). In: Kutner M H, Nachtsheim C J, Neter J, Li W, eds. Applied Linear Statistical Models. 5th ed. Boston: McGraw-Hill

Shaheen S, Guzman S, Zhang H (2010). Bikesharing in Europe, the Americas, and Asia. Transportation Research Record: Journal of the Transportation Research Board, 2143(1): 159–167

Shannon C E (1948). A mathematical theory of communication. Bell System Technical Journal, 27(3): 379–423

Singhvi D, Singhvi S, Frazier P I, Henderson S G, Mahony E O, Shmoys D B, Woodard D B (2015). Predicting bike usage for New York City's bike sharing system. In: AAAI Workshop: Computational Sustainability. Austin, TX: AAAI

Smith D A, Shen Y, Barros J, Zhong C, Batty M, Giannotti M (2020). A compact city for the wealthy? Employment accessibility inequalities between occupational classes in the London metropolitan region 2011. Journal of Transport Geography, 86: 102767

Song C, Qu Z, Blumm N, Barabási A L (2010). Limits of predictability in human mobility. Science, 327(5968): 1018–1021

Straits Research (2021). Bike sharing market new research analysis and forecast 2030. Online Report

Stromberg J (2015). Bike share users are mostly rich and white. Here's why that's hard to change

US Census Bureau (2012). 2009–2011 American Community Survey 3-year Public Use Microdata Samples

Yang Z, Chen J, Hu J, Shu Y, Cheng P (2019). Mobility modeling and data-driven closed-loop prediction in bike-sharing systems. IEEE Transactions on Intelligent Transportation Systems, 20(12): 4488–4499

Yang Z, Hu J, Shu Y, Cheng P, Chen J, Moscibroda T (2016). Mobility modeling and prediction in bike-sharing systems. In: Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services. Singapore: ACM, 165–178

Yu B, Yin H, Zhu Z (2018). Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting. In: Proceedings of the 27th International Joint Conference on Artificial Intelligence. Stockholm: ACM, 3634–3640

Zhang Y, Thomas T, Brussel M J G, van Maarseveen M F A M (2016). Expanding bicycle-sharing systems: Lessons learnt from an analysis of usage. PLoS One, 11(12): e0168604

Zheng V Z, Choi S, Sun L (2023). Enhancing deep traffic forecasting models with dynamic regression. arXiv preprint, arXiv:2301.06650

Zhou X (2015). Understanding spatiotemporal patterns of biking behavior by analyzing massive bike sharing data in Chicago. PLoS One, 10(10): e0137922

Zhou Y, Wang L, Zhong R, Tan Y (2018). A Markov chain based demand prediction model for stations in bike sharing systems. Mathematical Problems in Engineering, 8028714

Zhou Y, Yu Y, Wang Y, He B, Yang L (2023). Mode substitution and carbon emission impacts of electric bike sharing systems. Sustainable Cities and Society, 89: 104312