Using Origin-Destination Flow Graph and Public Transit Information to Enhance Short-Term Ridership Prediction in Bike-Sharing Systems

Kaifa Lu

International Center for Adaptation Planning & Design (iAdapt)

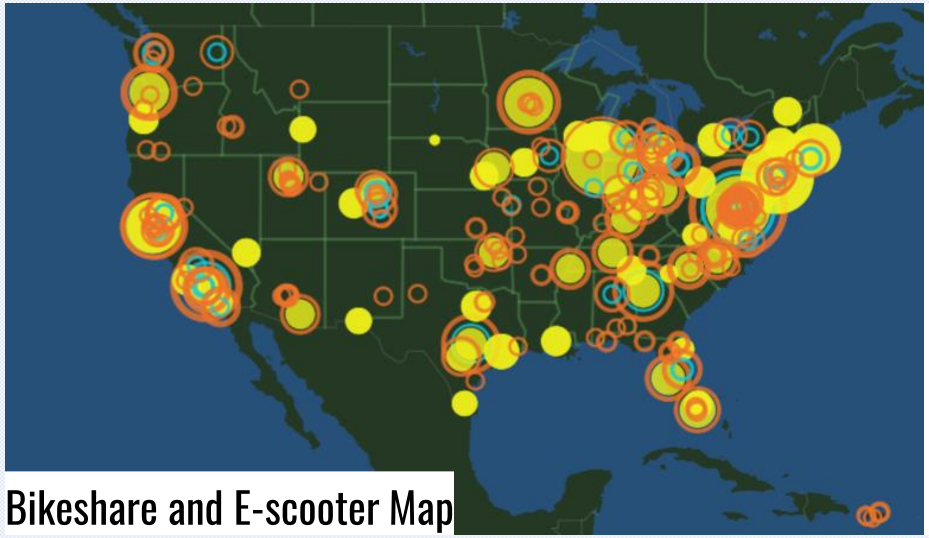
College of Design, Construction and Planning

University of Florida

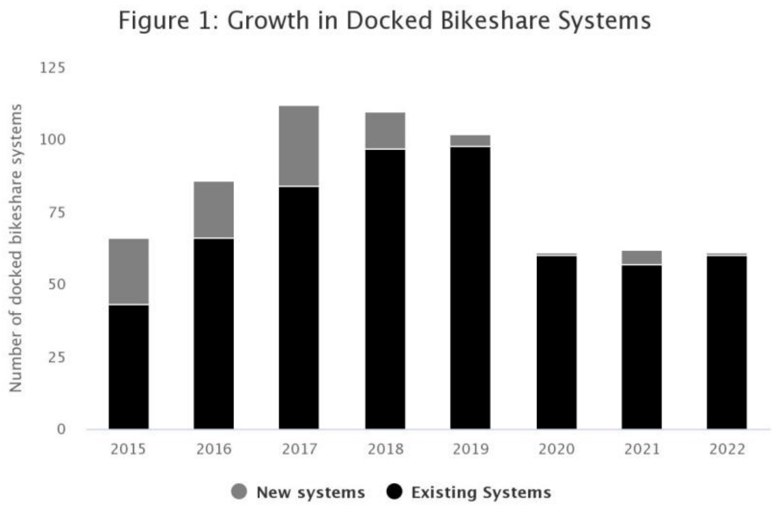
April 25th, 2023

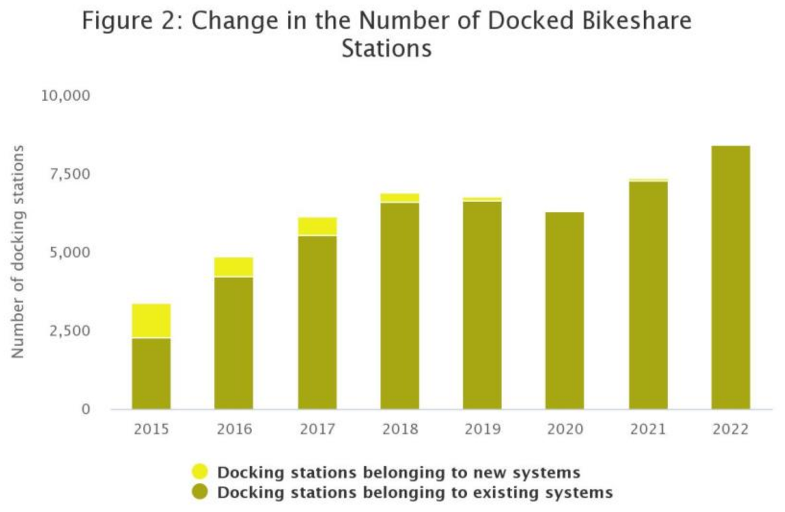
INTERNATIONAL CENTER FOR ADAPTATION PLANNING AND DESIGN | UF UNIVERSITY of FLORIDA

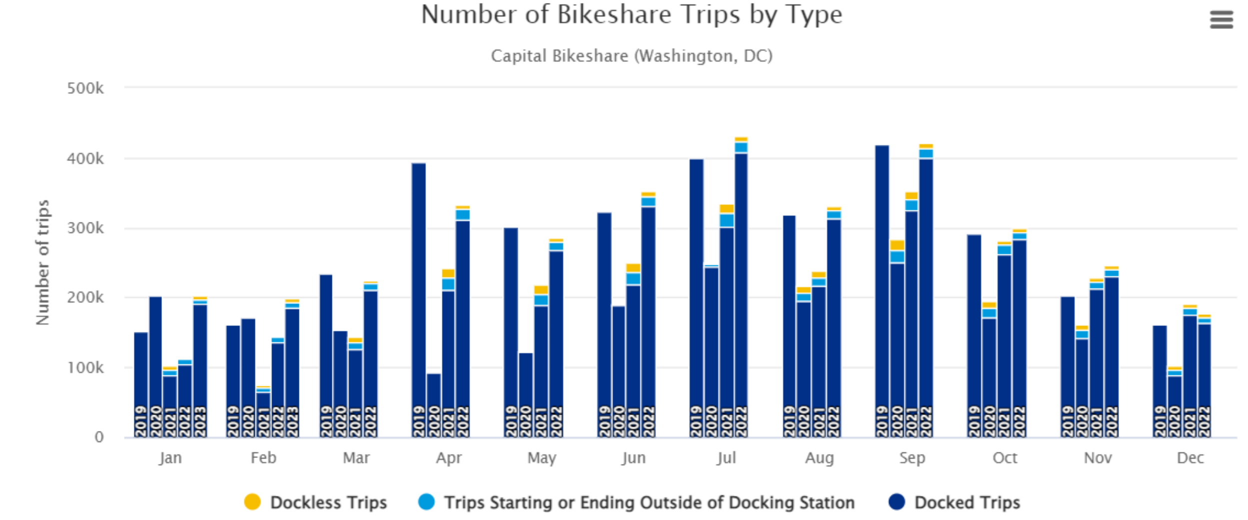
Bike-Sharing Market Is Blooming in Transportation Landscape



Bikeshare and E-scooter Map







- As of July 2022, 61 docked bikeshare systems open to the general public operated 8,473 docking stations in the U.S.
- In Washington D.C., there are about more than 200,000 bikeshare trips each month, most of which are docked trips.
- Bike-sharing system is blooming in transportation landscape.

Source: https://data.bts.gov/stories/s/Bikeshare-and-e-scooters-in-the-U-S-/fwcs-jprj/

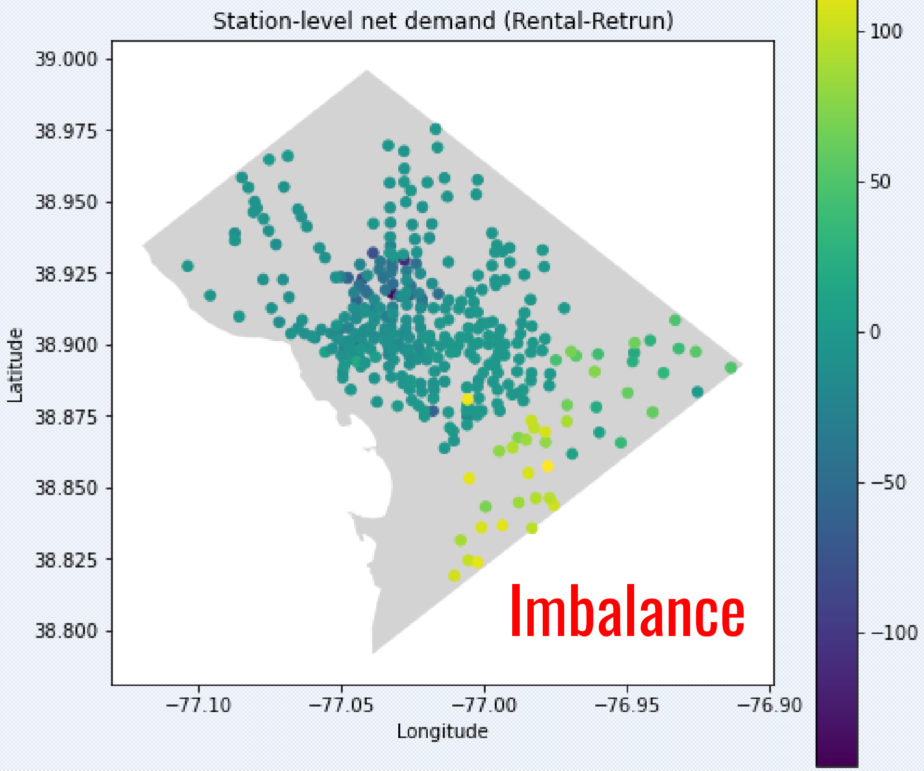
Bike-Sharing System Is Shaping New Transportation Landscape

Potentials

- Reducing people's dependencies on private vehicles
- Encouraging the use of public transit
- Improving accessibility, mobility, affordability

However,

Issue still exists



- Potential solution: Accurate spatiotemporal prediction of bike-sharing ridership is crucial to assist decision-makers in formulating effective bike dispatching and rebalancing plans and mitigating spatiotemporal imbalance.

Background

INTERNATIONAL CENTER FOR ADAPTATION PLANNING AND DESIGN

UF UNIVERSITY of FLORIDA

Research Question

How to improve the accuracy of short-term ridership prediction in bike-sharing system?

Literature Review

Types of dealing with docked bike ridership forecasting problems :

- *Forecast unit*: (1) at an individual station level (Lin et al. 2018); (2) at an aggregated group level (Li et al. 2015). The latter is preferred due to high flexibility of sharing bikes, and ridership biases and uncertainties among adjacent stations, which makes station-level prediction highly challenging.
- *Forecast model*:
  - Parametric statistical models (Kaltenbrunner et al. 2010): time-series prediction model, i.e., ARIMA (Jaber et al. 2022)
  - Machine learning approaches (Cho et al., 2021): random forest and support vector machine (Gao and Chen, 2022), convolutional neural networks (Li et al. 2023), and graph neural networks (Lin et al. 2018)

However, there is a lack of a feature engineering approach to predicting bike-sharing ridership that can compromise among data need, computing power, model interpretation, and prediction accuracy.

Background

INTERNATIONAL CENTER FOR ADAPTATION PLANNING AND DESIGN

UF UNIVERSITY of FLORIDA

Research Question

How to improve the accuracy of short-term ridership prediction in bike-sharing system?

Literature Review

The key to feature engineering is how to extract features significantly contributing to the accuracy of bike ridership prediction:

- *Temporal feature*: hour of a day, day of a week, holidays (Jaber et al. 2022)
- *Meteorology*: temp, humidity, wind, precipitation (Cantelmo et al. 2020)
- *Built environment*: design, density, diversity, accessibility (Liu & Lin, 2019)
- *Public transit*: metro & bus schedules and ridership (Fan et al. 2019)
- *Origin-destination flow graph*: there are unexplored features present in time-series data of bike trips than may improve the prediction precision.

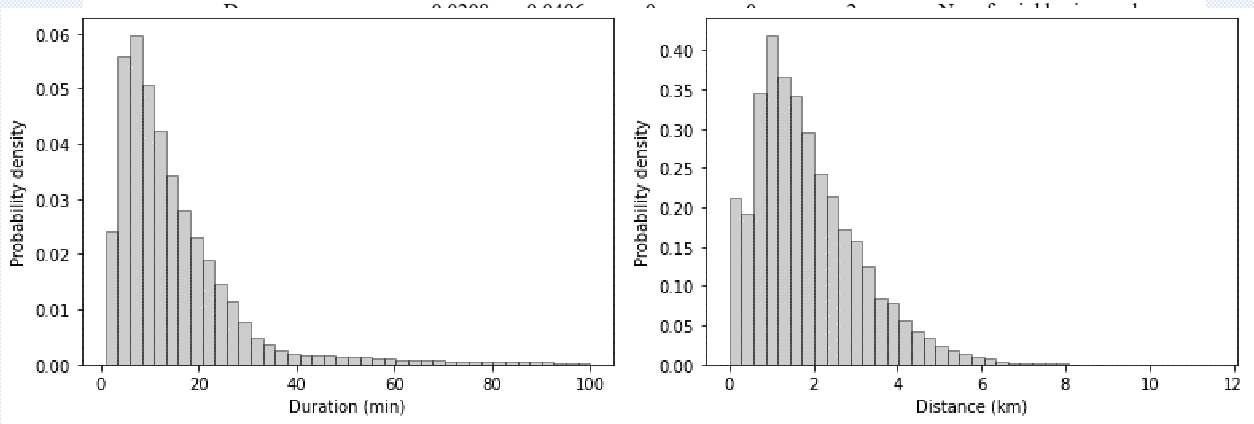
Most of feature engineering approaches was based on the association between bike-sharing ridership and the first three feature groups.

A big research gap: whether bike trip flow and public transit information can further improve the accuracy of bike-sharing ridership prediction?

Approach: Data and Methods

Table 2 Descriptive statistics affiliated to the 100 clustered centroids in Washington D.C.

Metric	Statistical indicator					Variable description
	Mean	Std	Min	Median	Max	
Hourly ridership	1.7971	3.9791	0	0	72	Hourly volume of bike trips
Graph attributes						
Relative humidity	62.30	20.80	15	62	100	Continuous (%)
Wind speed	9.2003	4.7373	0	9	29	Continuous (mph)
Wind direction	174.97	109.70	0	180	360	Continuous
Precipitation	0.0021	0.0090	0	0	0.1025	Continuous (inch)



- **Data:**
  - Bikeshare trip data in Washington D.C. during March 2020.
  - GTFS data in Washington D.C. during March 2020.
  - OpenStreetMap and socio-demographics & built environment
  - Hourly meteorological data during March 2020

- **Methods:**
  - Modelling paradigm: Network Science & Machine Learning.
  - Research method: A data-driven variable selection and feature engineering approach for short-term bike-sharing ridership prediction.
  - Research design:
    - Data preprocessing and preparation
      - Bike station clustering: K-means clustering to group 352 bike stations into 100 clusters based on spatial proximity
      - OD flow graph construction and feature extraction: Nodes, Edges, and Weights & Degree, Betweenness, and PageRank
      - Public transit information: static and dynamic number of transit (bus or metro) stops around clustered centroids
      - Data aggregation: bike trip data, meteorology, built environment, OD graph and public transit information at the same scale

Approach: Data and Methods

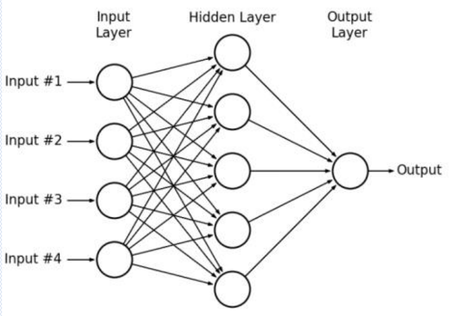
- **Methods (continued):**
  - Research design (continued):
    - Variable selection: A Spatial Vector Autoregressive LASSO (SpVAR-LASSO) Model
      - Dependent variables: $y_{i,t}$ - Bike-sharing ridership data at 100 clustered centroids
      - Independent variables: $\boldsymbol{x}_i$ - Population density, housing unit density, land use diversity, bike lane density, static and dynamic number of metro and bus stops around bike centroids, $\boldsymbol{y}_{i,t-1\sim t-n}$ - bike-sharing ridership data in the last 168 hours (1 week)
      - Model form:

$$y_{i,t} = \beta_0 + \boldsymbol{\beta_1 x}_i + \beta_2 \sum_j w_{ij} y_{j,t} + \boldsymbol{\beta_3 y}_{i,t-1\sim t-n} + \gamma \sum_j |\boldsymbol{\beta}_j| + \varepsilon_{i,t};\ \boldsymbol{\beta} = \underset{\boldsymbol{\beta}}{\text{argmin}}\left[(\boldsymbol{Y}_t - (\boldsymbol{X}, \boldsymbol{Y}_{t-1\sim t-n})\boldsymbol{\beta})'(\boldsymbol{Y_t} - (\boldsymbol{X}, \boldsymbol{Y}_{t-1\sim t-n})\boldsymbol{\beta}) + \gamma \sum_j |\boldsymbol{\beta}_j|\right]$$

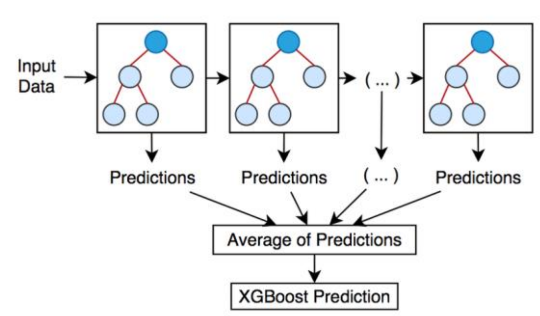
    - Short-term bike-sharing ridership prediction: Baseline model (HA, ARIMA), OLS, MLP, and **XGBoost** (a feature engineering approach)
      - Dependent variables: Y – Bike-sharing ridership data at 100 clustered centroids
      - Independent variables: X – Temporal variables, spatial variables, graph attributes, public transit information, and meteorology selected from the SpVAR-LASSO Model
      - Model form:

OLS: $Y = AX + b + \varepsilon$

MLP:

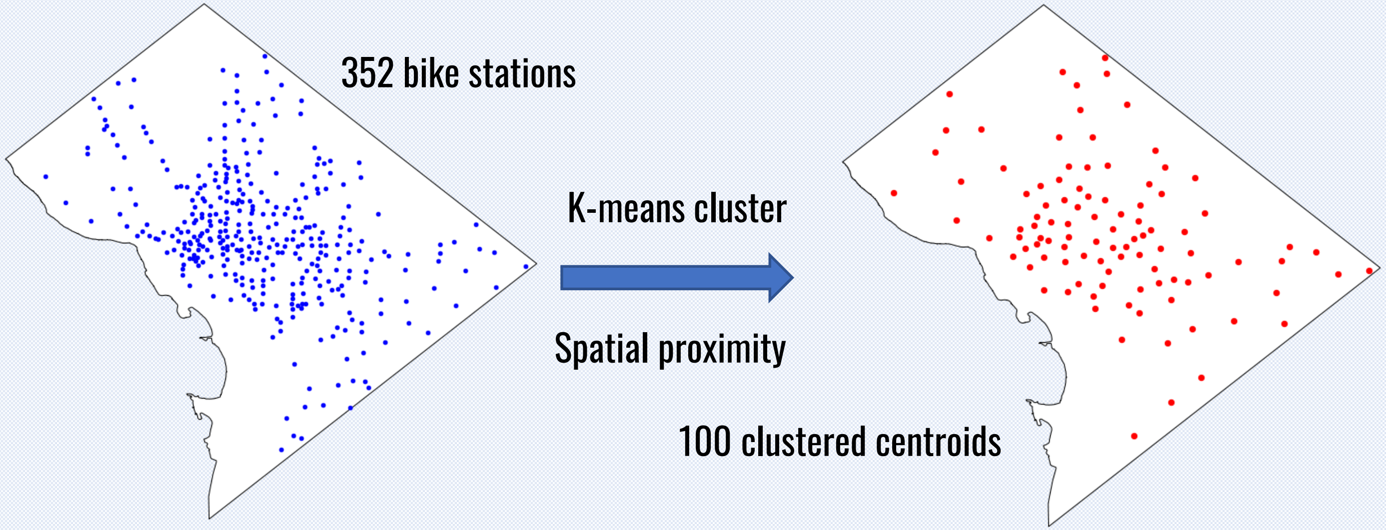


XGBoost:

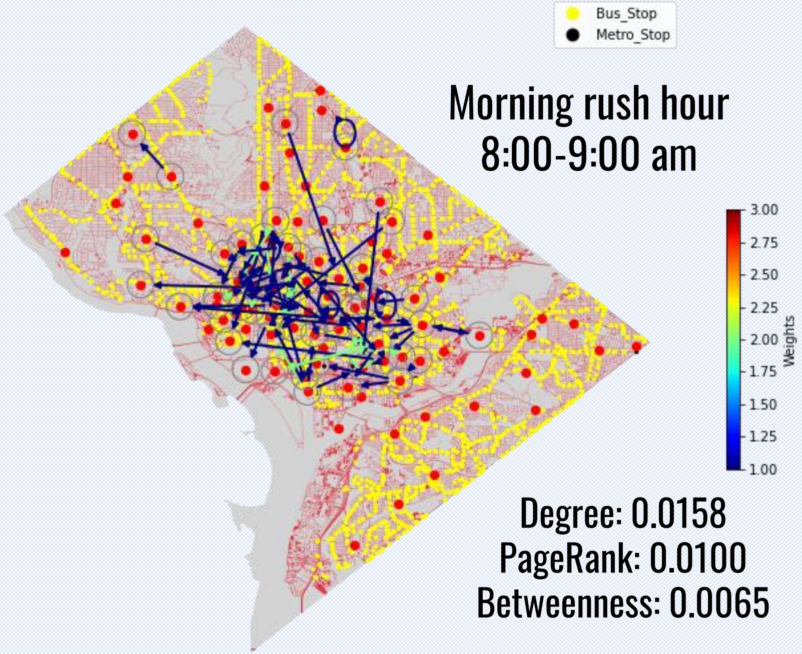


Results – Data Preprocessing and Aggregation

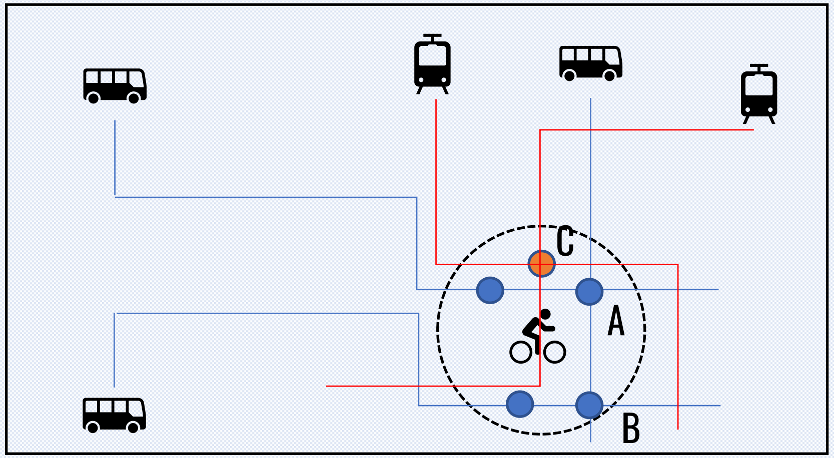
- **Bike station clustering:** 352 bike stations into 100 bike centroids



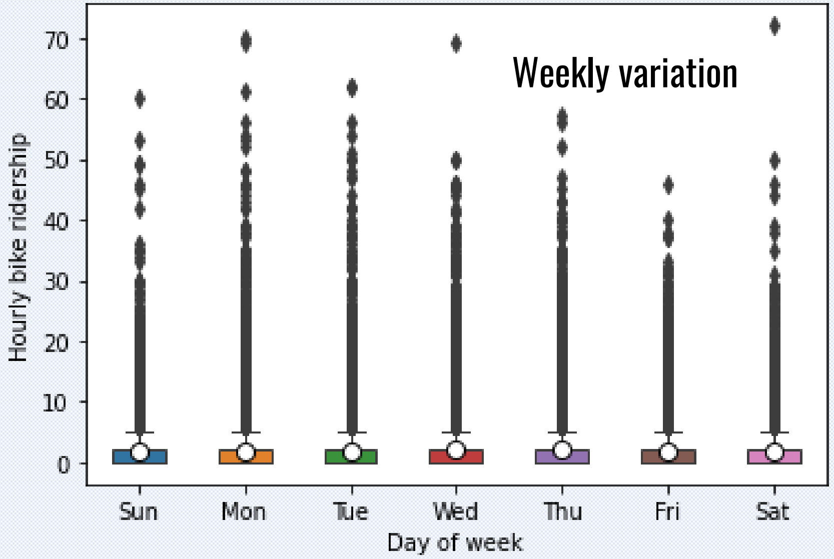
- **Graph construction and attributes:**
  - Nodes: 100 centroids of 352 bike stations.
  - Edges: A travel flow from A to B.
  - Weights: The hourly number of trips from A to B.

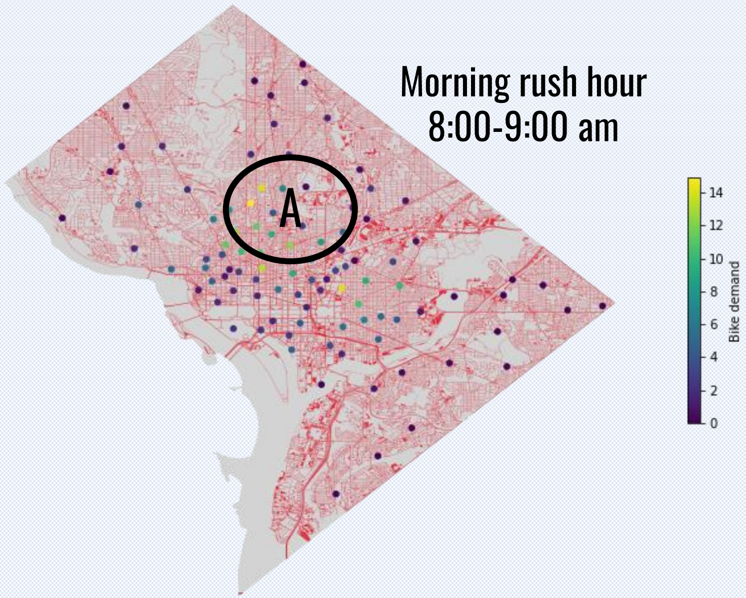


- **Public transit information:**
  - Dynamic number of metro stops nearby: 2
  - Dynamic number of bus stops nearby: 6
  - Static number of metro stops nearby: 1
  - Static number of bus stops nearby: 4

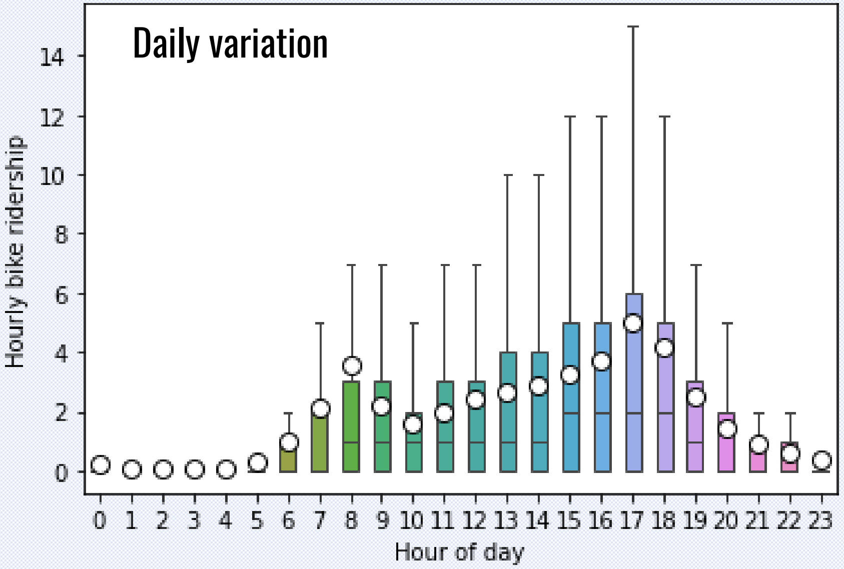


Results – Spatiotemporal Distribution of Bike Ridership

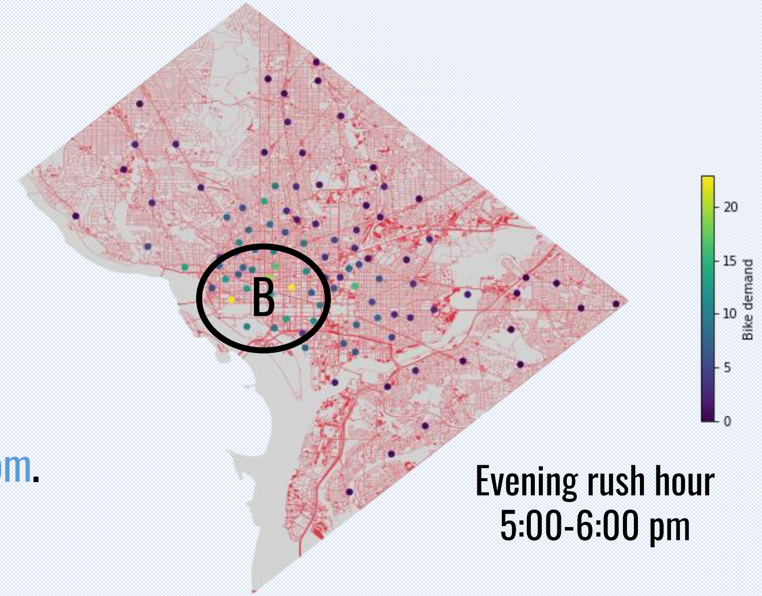




- **Spatial distribution:**
  - Downtown > Surroundings.
  - Implying land use patterns:
  - A: more likely to be residential areas.
  - B: more likely to be commercial areas or workplace

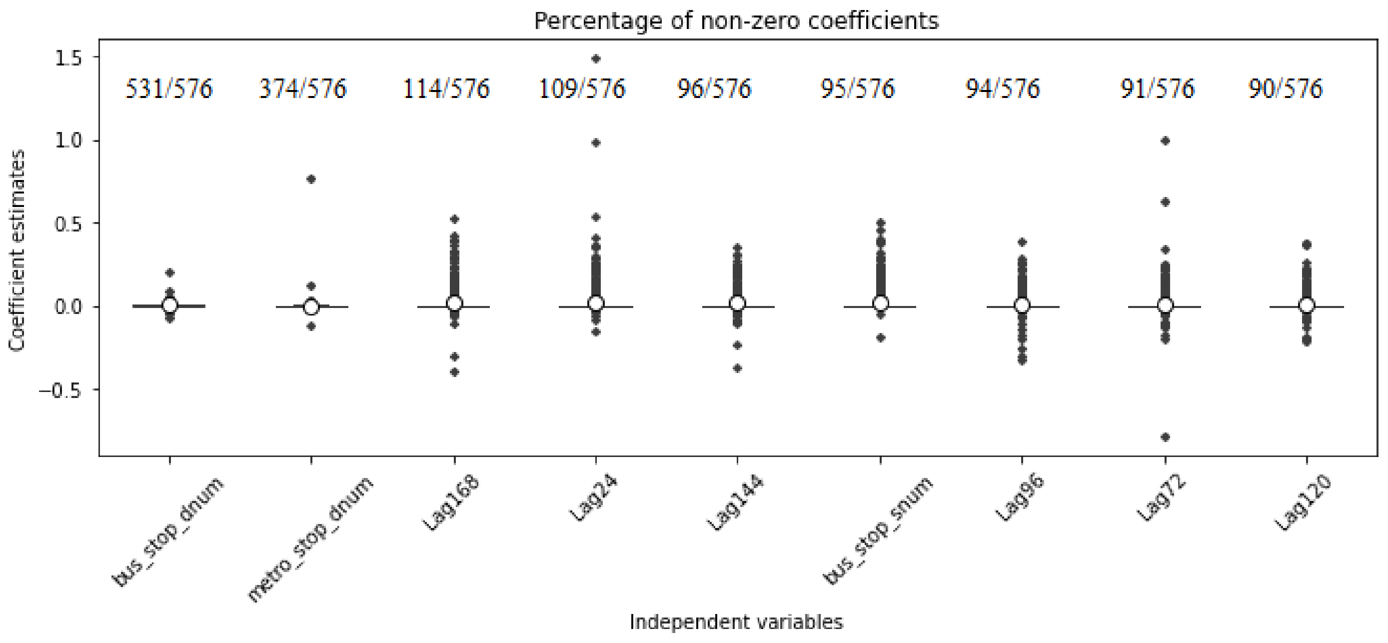


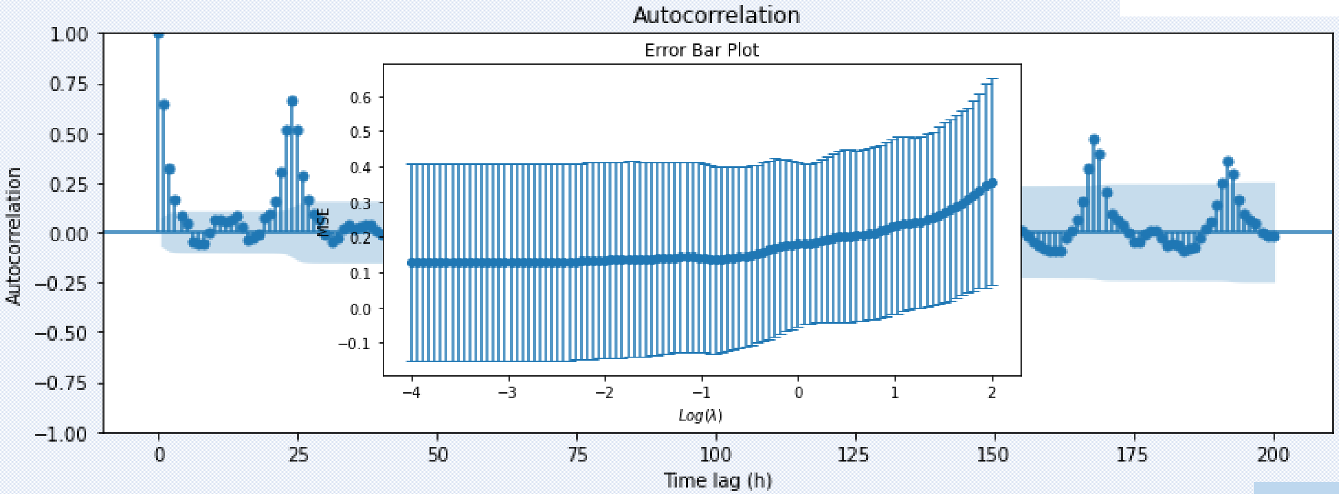
- **Temporal distribution:**
  - No significant weekly variation.
  - Significant daily variation.
  - Two peaks occurred at 8-9 am and 5-6 pm.



Results – Variable Selection for Ridership Prediction

- **SpVAR-LASSO**: both temporal and spatial information.
  - Hyperparameter:
    - Time lag: 168 (based on ACF)
    - Penalty: γ=0.0001 (based on CV)
  - Experiment design:
    - Perform 576 regressions (CV) for each hour
    - Count the number of each variable with non-zero coefficient estimates





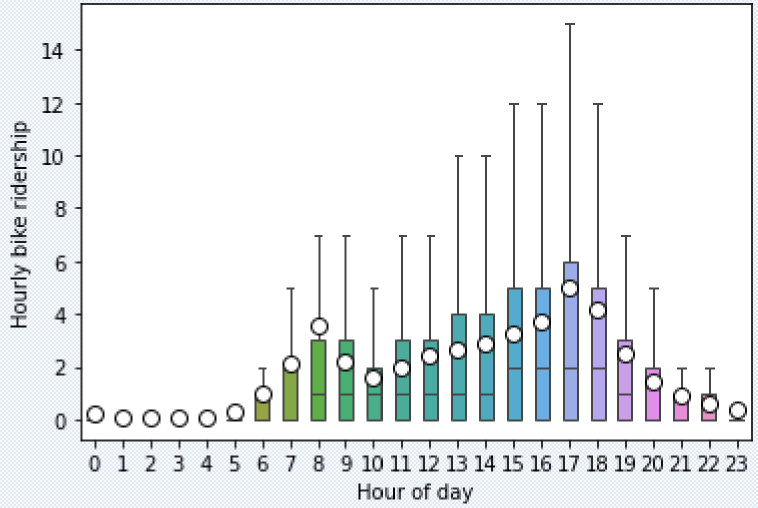
- **Variable Selection and Importance:**
  - Spatial variables: dynamic number of bus and metro stops, as well as static number of bus stops.
  - Temporal variables: bike demand data in the last 168th, 24th, 144th, 96th, 72nd, and 120th hours.

Results – Short-Term Bike Ridership Prediction

- **Experimental Design:**
  - Model inputs (Table 1).
  - Experimental design:
    - Basic scenario (S0): T+M as model inputs
    - Effect of spatial dependency (S1): T+M+S as model inputs
    - Effect of graph attributes (S2): T+M+S+G as model inputs
    - Effect of public transit information (S3): T+M+S+PT as model inputs
    - Effect of all features (S4): T+M+S+G+PT as model inputs
- Model: HA, ARIMA, OLS, MLP, XGBoost.
- Data: training (80%) and testing (20%) data
  - All data set (D0): All data sets as model inputs
  - Partial data set 1 (D1): Non-zero data sets
  - Partial data set 2 (D2): Rush-hour data sets
- Evaluation indices:
  - MSE and r^2

Table 3 Variables in the XGBoost model

Y: Dependent variable	Hourly ridership of sharing bikes in the next hour (+ 1 h)
X: Independent variable	
Temporal features (T)	Time-lagged bike-sharing ridership (-24, -72, -96, -120, -144, -168 h)
Meteorological factors (M)	Visibility, temperature, dew point temperature, relative humidity, wind speed & direction, precipitation
Spatial features (S)	Geographically weighted bike-sharing ridership
Time-lagged graphs (G)	Time-lagged degree, PageRank, and betweenness (-24, -72, -96, -120, -144, -168 h)
Public transit (PT)	Dynamic number of metro & bus stops, static number of bus stops around centroids



Results – Short-Term Bike Ridership Prediction

- **Prediction Accuracy:**
  - XGBoost > MLP > OLS > HA & ARIMA
  - Effect of feature inputs:
    - Key: temporal & spatial features
    - Graph + Public Transit can further improve accuracy
  - Effect of dataset:
    - All > rush-hour > nonzero dataset
    - In any dataset, public transit information is important
    - In all or nonzero datasets, graph is less important; while in rush-hour dataset, graph is more important.

Table 4 Comparison of prediction accuracy

Model	Dataset / Scenario	All dataset (N=57,600) MSE	All dataset (N=57,600) r^2	Non-zero dataset (N=23,162) MSE	Non-zero dataset (N=23,162) r^2	Rush-hour dataset (N=19,200) MSE	Rush-hour dataset (N=19,200) r^2
HA	T	15.57	-2.93	24.79	-5.22	29.27	-1.64
ARIMA		19.97	-0.58	N/A			
OLS	T+M	4.54/4.33	0.65/0.64	9.45/9.36	0.58/0.59	7.43/7.61	0.68/0.66
	T+M+S	4.14/3.90	0.68/0.68	8.71/8.41	0.61/0.63	6.80/6.25	0.70/0.73
	T+M+S+G	4.10/3.89	0.68/0.68	8.62/8.31	0.62/0.64	6.57/6.79	0.72/0.69
	T+M+S+PT	4.12/3.88	0.68/0.68	8.65/8.40	0.61/0.63	6.63/6.84	0.71/0.69
	T+M+S+G+PT	4.08/3.87	0.68/0.68	8.57/8.30	0.62/0.64	6.55/6.77	0.72/0.69
MLP	T+M	3.31/3.88	0.74/0.68	7.53/7.38	0.66/0.68	5.48/6.14	0.76/0.72
	T+M+S	3.10/3.57	0.76/0.71	6.74/6.62	0.70/0.71	4.97/4.86	0.78/0.79
	T+M+S+G	2.90/3.47	0.77/0.71	6.34/6.81	0.72/0.70	4.34/5.49	0.81/0.75
	T+M+S+PT	2.97/3.46	0.77/0.71	6.53/6.68	0.71/0.71	4.55/5.57	0.80/0.75
	T+M+S+G+PT	2.88/3.49	0.78/0.71	6.22/6.86	0.72/0.70	4.68/5.84	0.80/0.74
XGBoost	T+M	3.07/3.70	0.76/0.70	6.46/7.14	0.71/0.69	4.26/5.45	0.82/0.75
	T+M+S	2.83/3.41	0.78/0.72	5.98/6.56	0.73/0.71	4.16/4.96	0.82/0.79
	T+M+S+G	2.82/3.42	0.78/0.72	5.95/6.63	0.73/0.71	4.02/5.46	0.83/0.75
	T+M+S+PT	2.68/3.27	0.79/0.73	5.80/6.37	0.74/0.72	3.96/5.19	0.83/0.77
	T+M+S+G+PT	**2.67/3.25**	**0.79/0.73**	**5.71/6.44**	**0.75/0.72**	**3.93/5.30**	**0.83/0.76**

Note: a/b in MSE and R^2 columns represent training and testing MSE and R^2, respectively.

Results – Short-Term Bike Ridership Prediction

INTERNATIONAL CENTER FOR ADAPTATION PLANNING AND DESIGN

UF UNIVERSITY of FLORIDA

All dataset

OLS Regression Results

Dep. Variable:	lag_0	R-squared:	0.681
Model:	OLS	Adj. R-squared:	0.681
Method:	Least Squares	F-statistic:	3962.
Date:	Mon, 24 Apr 2023	Prob (F-statistic):	0.00
Time:	14:10:01	Log-Likelihood:	-1.2196e+05
No. Observations:	57600	AIC:	2.440e+05
Df Residuals:	57568	BIC:	2.443e+05
Df Model:	31		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	-0.9855	0.232	-4.247	0.000	-1.440	-0.531
lag_24	0.3410	0.003	100.195	0.000	0.334	0.348
lag_96	0.0434	0.003	12.732	0.000	0.037	0.050
lag_120	0.0370	0.004	9.884	0.000	0.030	0.044
lag_144	0.0804	0.004	21.966	0.000	0.073	0.088
lag_168	0.1610	0.003	48.888	0.000	0.155	0.167
gw_lag_0	0.4896	0.006	77.199	0.000	0.477	0.502
degree_24	-2.0401	0.285	-7.165	0.000	-2.598	-1.482
pagerank_24	0.4538	0.368	1.232	0.218	-0.268	1.176
betweenness_24	-1.5621	0.470	-3.322	0.001	-2.484	-0.640
degree_96	-2.0880	0.282	-7.397	0.000	-2.641	-1.535
pagerank_96	1.8029	0.381	4.732	0.000	1.056	2.550
betweenness_96	0.4549	0.474	0.960	0.337	-0.474	1.384
degree_120	-1.2003	0.279	-4.306	0.000	-1.747	-0.654
pagerank_120	1.2817	0.384	3.339	0.001	0.529	2.034
betweenness_120	-0.0568	0.475	-0.120	0.905	-0.988	0.874
degree_144	-1.7707	0.276	-6.426	0.000	-2.311	-1.231
pagerank_144	1.3110	0.387	3.388	0.001	0.553	2.069
betweenness_144	-0.5192	0.479	-1.084	0.279	-1.458	0.420
degree_168	-1.5701	0.278	-5.651	0.000	-2.115	-1.025
pagerank_168	1.2008	0.394	3.045	0.002	0.428	1.974
betweenness_168	-1.1023	0.476	-2.315	0.021	-2.035	-0.169
metro_dnum_0	-0.0004	0.000	-1.472	0.141	-0.001	0.000
bus_dnum_0	-0.0023	0.000	-14.119	0.000	-0.003	-0.002
bus_snum_0	0.0986	0.006	16.894	0.000	0.087	0.110
visibility	0.0087	0.007	1.259	0.208	-0.005	0.022
temperature	0.0353	0.008	4.417	0.000	0.020	0.051
dew_point_temp	-0.0329	0.008	-3.969	0.000	-0.049	-0.017
relative_humidity	0.0069	0.002	3.120	0.002	0.003	0.011
wind_speed	-0.0135	0.002	-6.352	0.000	-0.018	-0.009
wind_direction	-0.0001	8.46e-05	-1.346	0.178	-0.000	5.19e-05
precipitation	1.6520	0.944	1.750	0.080	-0.199	3.503

Omnibus:	44407.355	Durbin-Watson:	1.800
Prob(Omnibus):	0.000	Jarque-Bera (JB):	4815275.152
Skew:	3.006	Prob(JB):	0.00
Kurtosis:	47.387	Cond. No.	2.32e+04

Non-zero dataset

OLS Regression Results

Dep. Variable:	lag_0	R-squared:	0.621
Model:	OLS	Adj. R-squared:	0.620
Method:	Least Squares	F-statistic:	1223.
Date:	Mon, 24 Apr 2023	Prob (F-statistic):	0.00
Time:	14:24:25	Log-Likelihood:	-57684.
No. Observations:	23162	AIC:	1.154e+05
Df Residuals:	23130	BIC:	1.157e+05
Df Model:	31		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	-0.7840	0.557	-1.409	0.159	-1.875	0.307
lag_24	0.3285	0.005	61.707	0.000	0.318	0.339
lag_96	0.0435	0.005	8.045	0.000	0.033	0.054
lag_120	0.0396	0.006	6.665	0.000	0.028	0.051
lag_144	0.0856	0.006	14.741	0.000	0.074	0.097
lag_168	0.1547	0.005	29.652	0.000	0.144	0.165
gw_lag_0	0.5192	0.011	48.971	0.000	0.498	0.540
degree_24	-3.6869	0.638	-5.776	0.000	-4.938	-2.436
pagerank_24	2.3008	1.050	2.192	0.028	0.243	4.358
betweenness_24	-0.7553	0.870	-0.868	0.385	-2.460	0.950
degree_96	-3.1206	0.602	-5.184	0.000	-4.301	-1.941
pagerank_96	3.3281	[illegible]	[illegible]	[illegible]	1.303	5.353
betweenness_96	1.7308	[illegible]	[illegible]	[illegible]	-0.013	3.474
degree_120	-1.5707	[illegible]	[illegible]	[illegible]	-2.728	-0.413
pagerank_120	2.0223	[illegible]	[illegible]	[illegible]	-0.075	4.119
betweenness_120	1.4860	[illegible]	[illegible]	[illegible]	-0.282	3.254
degree_144	-2.5494	[illegible]	[illegible]	[illegible]	-3.673	-1.426
pagerank_144	2.5721	[illegible]	[illegible]	[illegible]	0.495	4.649
betweenness_144	1.0238	0.903	[illegible]	[illegible]	-0.747	2.794
degree_168	-2.1754	0.575	[illegible]	[illegible]	-3.303	-1.048
pagerank_168	3.7155	1.130	3.289	0.001	1.502	5.929
betweenness_168	0.0161	0.917	0.018	0.986	-1.781	1.813
metro_dnum_0	-0.0002	0.000	-0.388	0.698	-0.001	0.001
bus_dnum_0	-0.0023	0.000	-8.571	0.000	-0.003	-0.002
bus_snum_0	0.0372	0.011	3.350	0.001	0.015	0.059
visibility	0.0196	0.023	0.840	0.401	-0.026	0.065
temperature	0.0599	0.017	3.592	0.000	0.027	0.093
dew_point_temp	-0.0442	0.017	-2.567	0.010	-0.078	-0.010
relative_humidity	0.0072	0.005	1.453	0.146	-0.003	0.017
wind_speed	-0.0188	0.005	-3.810	0.000	-0.028	-0.009
wind_direction	-0.0003	0.000	-1.562	0.118	-0.001	8.3e-05
precipitation	-0.4739	4.026	-0.118	0.906	-8.364	7.416

Omnibus:	13473.526	Durbin-Watson:	1.806
Prob(Omnibus):	0.000	Jarque-Bera (JB):	459443.214
Skew:	2.224	Prob(JB):	0.00
Kurtosis:	24.361	Cond. No.	4.92e+04

Temporal + Spatial + Degree + PageRank + Public Transit + Wind + Temperature

Rush-hour dataset

OLS Regression Results

Dep. Variable:	lag_0	R-squared:	0.713
Model:	OLS	Adj. R-squared:	0.713
Method:	Least Squares	F-statistic:	1538.
Date:	Mon, 24 Apr 2023	Prob (F-statistic):	0.00
Time:	14:14:25	Log-Likelihood:	-45352.
No. Observations:	19200	AIC:	9.077e+04
Df Residuals:	19168	BIC:	9.102e+04
Df Model:	31		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	-0.2863	0.519	-0.551	0.581	-1.304	0.732
lag_24	0.3951	0.006	69.658	0.000	0.384	0.406
lag_96	0.0344	0.006	5.927	0.000	0.023	0.046
lag_120	0.0297	0.007	4.516	0.000	0.017	0.043
lag_144	0.0762	0.006	11.960	0.000	0.064	0.089
lag_168	0.1537	0.005	28.033	0.000	0.143	0.164
gw_lag_0	0.5155	0.011	48.634	0.000	0.495	0.536
degree_24	-5.2390	0.651	-8.047	0.000	-6.515	-3.963
pagerank_24	1.0737	1.235	0.870	0.385	-1.346	3.494
betweenness_24	0.2846	0.998	0.285	0.776	-1.671	2.240
degree_96	-1.7893	0.592	-3.025	0.002	-2.949	-0.630
pagerank_96	4.1715	1.384	3.015	0.003	1.459	6.884
betweenness_96	1.2429	1.040	1.195	0.232	-0.796	3.281
degree_120	-0.3456	0.575	-0.601	0.548	-1.473	0.782
pagerank_120	1.4789	1.405	1.052	0.293	-1.275	4.233
betweenness_120	-0.6703	1.030	-0.651	0.515	-2.689	1.349
degree_144	-2.5175	0.557	-4.518	0.000	-3.610	-1.425
pagerank_144	1.9807	1.417	1.398	0.162	-0.797	4.759
betweenness_144	2.0144	1.038	1.940	0.052	-0.020	4.049
degree_168	-2.2351	0.552	-4.050	0.000	-3.317	-1.153
pagerank_168	0.2867	1.428	0.201	0.841	-2.512	3.086
betweenness_168	-0.0180	1.028	-0.017	0.986	-2.033	1.997
metro_dnum_0	-0.0010	0.000	-2.098	0.036	-0.002	-6.64e-05
bus_dnum_0	-0.0016	0.000	-5.774	0.000	-0.002	-0.001
bus_snum_0	0.0772	0.013	6.112	0.000	0.052	0.102
visibility	-0.0242	0.016	-1.558	0.119	-0.055	0.006
temperature	0.0286	0.018	1.626	0.104	-0.006	0.063
dew_point_temp	-0.0034	0.019	-0.183	0.855	-0.040	0.033
relative_humidity	-0.0017	0.005	-0.347	0.729	-0.012	0.008
wind_speed	-0.0198	0.005	-4.152	0.000	-0.029	-0.010
wind_direction	-0.0009	0.000	-4.635	0.000	-0.001	-0.001
precipitation	1.4465	2.185	0.662	0.508	-2.836	5.729

Omnibus:	11288.765	Durbin-Watson:	1.799
Prob(Omnibus):	0.000	Jarque-Bera (JB):	465101.489
Skew:	2.197	Prob(JB):	0.00
Kurtosis:	26.708	Cond. No.	2.55e+04

Summary

Conclusions

Bike Ridership Distribution

- Significant daily variations and spatial distribution patterns

SpVAR-LASSO: Variable Selection

- Bike ridership prediction: public transit, bike ridership in the last 168^{th},24^{th},144^{th},96^{th},72^{nd},120^{th} hours.

XGBoost: Bike Ridership Prediction

- Graph (rush-hour) and public transit can improve accuracy.

Discussion

Innovation:

(1) To propose a feature engineering approach for bike ridership prediction

(2) To reveal the impact of OD flow graph and transit information on the prediction accuracy

Implication:

(1) Bike fleet rebalancing to reduce the spatiotemporal imbalance

(2) Insights into the interaction between bike-sharing system and public transit

References

Baumanis, C., Hall, J., Machemehl, R., 2023. A machine learning approach to predicting bicycle demand during the COVID-19 pandemic. Research in Transportation Economics.

Bigazzi, A., Wong, K., 2020. Electric bicycle mode substitution for driving, public transit, conventional cycling, and walking. Transportation Research Part D: Transport and Environment 85.

Cantelmo, G., Kucharski, R., Antoniou, C., 2020. Low-dimensional model for bike-sharing demand forecasting that explicitly accounts for weather data. Transportation Research Record: Journal of the Transportation Research Board 2674, 132-144.

Chen, L., Ma, X., Nguyen, T.-M.-T., Pan, G., Jakubowicz, J., 2017. Understanding bike trip patterns leveraging bike sharing system open data. Frontiers of Computer Science 11, 38-48.

Cho, J.-H., Ham, S.W., Kim, D.-K., 2021. Enhancing the accuracy of peak hourly demand in bike-sharing systems using a graph convolutional network with public transit usage data. Transportation Research Record: Journal of the Transportation Research Board 2675, 554-565.

Fan, A., Chen, X., Wan, T., 2019. How have travelers changed mode choices for first/last mile trips after the introduction of bicycle-sharing systems: An empirical study in Beijing, China. Journal of Advanced Transportation 2019, 1-16.

Ferguson, B., Sanguinetti, A., 2021. Facilitating micromobility for first and last mile connection with public transit through environmental design: A case study of California Bay Area rapid transit stations. Proceedings of the Design Society 1, 1577-1586.

Gao, C., Chen, Y., 2022. Using machine learning methods to predict demand for bike sharing, in: Stienmetz, J.L., Ferrer-Rosell, B., Massimo, D. (Eds.), Information and Communication Technologies in Tourism 2022. Springer.

Hartigan, J.A., Wong, M.A., 1979. Algorithm AS 136: A K-Means Clustering Algorithm. Applied Statistics 28.

Hasselwander, M., Nieland, S., Dematera-Contreras, K., Goletz, M., 2022. MaaS for the masses: Potential transit accessibility gains and required policies under mobility-as-a-service.

Jaber, A., Csonka, B., Juhasz, J., 2022. Long term time series prediction of bike sharing trips: A cast study of Budapest City, 2022 Smart City Symposium Prague (SCSP), pp. 1-5.

Jin, C., Lee, G., 2019. Exploring spatiotemporal dynamics in a housing market using the spatial vector autoregressive Lasso: A case study of Seoul, Korea. Transactions in GIS 24, 27-43.

Kaltenbrunner, A., Meza, R., Grivolla, J., Codina, J., Banchs, R., 2010. Urban cycles and mobility patterns: Exploring and predicting trends in a bicycle-based public transport system. Pervasive and Mobile Computing 6, 455-466.

Li, X., Xu, Y., Zhang, X., Shi, W., Yue, Y., Li, Q., 2023. Improving short-term bike sharing demand forecast through an irregular convolutional neural network. Transportation Research Part C: Emerging Technologies 147.

Li, Y., Zheng, Y., Zhang, H., Chen, L., 2015. Traffic prediction in a bike-sharing system, Proceedings of the 23rd SIGSPATIAL International Conference on Advances in Geographic Information Systems, pp. 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.

References

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, H.-C., Lin, J.-J., 2019. Associations of built environments with spatiotemporal patterns of public bicycle use. Journal of Transport Geography 74, 299-312.

Madapur, B., Madangopal, S., Chandrashekar, M.N., 1970. Micro-mobility infrastructure for redefining urban mobility. European Journal of Engineering Science and Technology 3, 71-85.

Oeschger, G., Carroll, P., Caulfield, B., 2020. Micromobility and public transport integration: The current state of knowledge. Transportation Research Part D: Transport and Environment 89.

Peng, B., Yang, K., Dong, X., 2023. Variable selection for quantile autoregressive model: Bayesian methods versus classical methods. Journal of Applied Statistics, 1-33.

Reck, D.J., Haitao, H., Guidon, S., Axhausen, K.W., 2021. Explaining shared micromobility usage, competition and mode choice by modelling empirical data from Zurich, Switzerland. Transportation Research Part C: Emerging Technologies 124.

Saberi, M., Ghamami, M., Gu, Y., Shojaei, M.H., Fishman, E., 2018. Understanding the impacts of a public transit disruption on bicycle sharing mobility patterns: A case of Tube strike in London. Journal of Transport Geography 66, 154-166.

Saberi, M., Mahmassani, H.S., Brockmann, D., Hosseini, A., 2016. A complex network perspective for characterizing urban travel demand patterns: graph theoretical analysis of large-scale origin–destination demand networks. Transportation 44, 1383-1402.

Tavassoli, K., Tamannaei, M., 2020. Hub network design for integrated Bike-and-Ride services: A competitive approach to reducing automobile dependence. Journal of Cleaner Production 248.

Xu, Y., Paliwal, M., Zhao, X., 2021. Real-time forecasting of dockless scooter-sharing demand: A context-aware spatio-temporal multi-graph convolutional network approach. arXiv preprint arXiv:2111.01355.

Yang, Y., Heppenstall, A., Turner, A., Comber, A., 2019. A spatiotemporal and graph-based analysis of dockless bike sharing patterns to understand urban flows over the last mile. Computers, Environment and Urban Systems 77.

Yang, Y., Heppenstall, A., Turner, A., Comber, A., 2020. Using graph structural information about flows to enhance short-term demand prediction in bike-sharing systems. Computers, Environment and Urban Systems 83.

Yang, Y., Shao, X., Zhu, Y., Yao, E., Liu, D., Zhao, F., Zou, Y., 2023. Short-term forecasting of dockless bike-sharing demand with the built environment and weather. Journal of Advanced Transportation 2023, 1-13.

Zellner, A., 1962. An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. Journal of the American statistical Association 57, 348-368.

Zhou, Y., Li, Y., Zhu, Q., Chen, F., Shao, J., Luo, Y., Zhang, Y., Zhang, P., Yang, W., 2019. A reliable traffic prediction approach for bike- sharing system by exploiting rich information with temporal link prediction strategy. Transactions in GIS 23, 1125-1151.

Zuo, T., Wei, H., Chen, N., Zhang, C., 2020. First-and-last mile solution via bicycling to improving transit accessibility and advancing transportation equity. Cities 99.