



HAL
open science

Contextual Data Integration for Bike-sharing Demand Prediction with Graph Neural Networks in Degraded Weather Conditions

Romain Rochas, Angelo Furno, Nour-Eddin El Faouzi

► **To cite this version:**

Romain Rochas, Angelo Furno, Nour-Eddin El Faouzi. Contextual Data Integration for Bike-sharing Demand Prediction with Graph Neural Networks in Degraded Weather Conditions. 2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC), Sep 2023, Bilbao, Spain. pp.5436 - 5441, 10.1109/itsc57777.2023.10422448 . hal-04796111

HAL Id: hal-04796111

<https://hal.science/hal-04796111v1>

Submitted on 3 Dec 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Contextual Data Integration for Bike-sharing Demand Prediction with Graph Neural Networks in Degraded Weather Conditions

Romain Rochas¹, Angelo Furno¹ and Nour-Eddin El Faouzi¹

Abstract—Demand for bike sharing is impacted by various factors, such as weather conditions, events, and the availability of other transportation modes. This impact remains elusive due to the complex interdependence of these factors or location-related user behavior variations. It is also not clear which factor is additional information which are not already contained in the historical demand. Intermodal dependencies between bike-sharing and other modes are also underexplored, and the value of this information has not been studied in degraded situations. The proposed study analyzes the impact of adding contextual data, such as weather, time embedding, and road traffic flow, to predict bike-sharing Origin-Destination (OD) flows in atypical weather situations. Our study highlights a mild relationship between prediction quality of bike-sharing demand and road traffic flow, while the introduced time embedding allows outperforming state-of-the-art results, particularly in the case of degraded weather conditions. Including weather data as an additional input further improves our model with respect to the basic ST-ED-RMGC prediction model by reducing of more than 20% the prediction error in degraded weather condition.

I. INTRODUCTION

In recent years, there has been a significant surge in the adoption of soft mobility options, driven by factors such as health considerations, escalating oil prices, and personal ecological concerns. This growing trend towards soft mobility, which includes biking and other forms of non-motorized transportation, has prompted the need for accurate demand prediction on the network. Such prediction improvements hold substantial benefits for both users and operators of Bike-Sharing Systems (BSS), particularly in addressing challenges like bike rebalancing [4]. In practice, the bike-sharing systems serve as connection on the multi-modal transport network [2], allowing for reduced travel times, cost-effectiveness, and a smaller spatial and environmental footprint compared to traditional modes of transportation [3]. While demand forecasting has historically relied on statistical and machine learning techniques, the advent of deep learning has ushered in a new era of forecasting algorithms. Among these, GNNs [6] appeared in 2014, but it has only been applied to the transportation domain from 2017 [7] and has emerged as the state-of-the-art models for capturing spatial dependencies in transportation prediction. Various models exist for bike-sharing demand prediction, ranging from station level forecast [9], to cluster-based prediction [8],

and many other variations deeply detailed in surveys such as [14]. However, only a limited number of studies have focused on Origin-Destination (OD) prediction [11]. OD prediction poses unique challenges, as the final destinations are usually unknown, and OD matrices generated are sparse [13], [12], [18]. Moreover, forecasting demand in the case of atypical phenomena, such as non-recurrent events, remains a challenge [5]). Bike-sharing demands exhibit temporal and calendar dependencies tied to factors such as weekdays, public holidays, workdays, and school holidays [1]. Furthermore, they are influenced by weather conditions such as wind, humidity, temperature, and particularly rainfall [1],[15],[17]. While this introduction does not aim to comprehensively address all the factors impacting bike-sharing demand, in-depth studies provide further insights [10]. We have chosen to work on an OD prediction algorithm, the ST-ED-RMGC [18] as it is one of the few addressing the issue of forecasting by OD, it is a recent state-of-the-art model, and it integrates proven prediction modules (Multi-Graph-Convolution, Encoder-Decoder and Residual Module) for traffic prediction. In contrast to the original paper, we have chosen to work with bike-sharing data, incorporating contextual information such as time, weather conditions, and multi-modal data, including road car flow. Specifically, we evaluate the inclusion of contextual data for predicting Weather-related scenarios. We have already worked on this algorithm to make an empirical analysis of the forecasting accuracy of the algorithm when predicting bike-sharing demand under the constraint of weather scenarios. This paper goes further by integrating contextual data within the training phase. For further information, the reader can refer to our last paper [19]. The main contributions of this study are the following:

- We propose the integration of an embedding module in the ST-ED-RMGC to capture the calendar-based dependencies of bike-sharing demand.
- We propose the integration of contextual weather data to account for the meteorological sensitivity of bike-sharing demand.
- We propose the integration of road car flow data by zone to capture multi-modal dependencies under rainy scenarios.
- We evaluate the addition of these contextual data not only on the global prediction performances of the neural network, but also on the prediction quality during atypical weather scenarios

The rest of the paper is organised as follows. In Sec. II, we present the problem being addressed, introduce the ST-

This work was not supported by any organization

¹Romain Rochas, Angelo Furno and Nour-Eddin El Faouzi are with LICIT-ECO7 UMR T9401, University Gustave Eiffel, University of Lyon, ENTPE, Lyon, France. romain.rochas@entpe.fr, angelo.furno@univ-eiffel.fr, nour-eddin.elfaouzi@univ-eiffel.fr

ED-RMGC model and describe the proposed improvements by adding embedding and contextual data. In Sec. III, we present the available dataset and the spatial aggregation applied to road loop detectors and bike sharing stations. Sec. IV, details the results obtained from the various scenarios tested. Sec. V concludes the study while outlining future research directions.

II. METHODOLOGY

This section details the problem and the methodology proposed initially in [18], as well as our proposed enhancements to the original model that leverage contextual data in several forms to enhance the prediction accuracy in degraded weather conditions.

Original Problem Statement

Let \mathcal{Z} represent the set of areas corresponding to a partitioning of the analyzed territory, and let $OD_{\mathcal{Z}} = \{(z_i, z_j) | z_i, z_j \in \mathcal{Z}, i \neq j\}$ denote the set of all pairs of zones. We define $Y_{(z_i, z_j)}^t \in \mathbb{R}$ as the bike-sharing demand between (z_i, z_j) at t time-step. Specifically, $Y_{(z_i, z_j)}^t$ represents the outflow demand for destination z_j that left from z_i during time slot t . Let ${}^B X_{(z_i, z_j)}^t = [Y_{(z_i, z_j)}^{t-7d}, Y_{(z_i, z_j)}^{t-d}, Y_{(z_i, z_j)}^{t-2}, Y_{(z_i, z_j)}^{t-1}]$ be the associated sequence of previously observed bike-sharing demands which allows to predict $Y_{(z_i, z_j)}^t$, where $t-2$ and $t-1$ represent respectively the period t minus 1 and 2 time-steps, and d represent a time-step of 1 day. Let G_u be a complete graph associated to a relationship u on $OD_{\mathcal{Z}}$, and let A_u be its adjacency matrix which takes into account a spatial dependency. Multiple u relationships can be jointly considered, and \ddot{A} denotes the concatenation of all these adjacency matrices. All the considered adjacency matrices are detailed in II-B. The temporal dependencies are instead taken into account via the historical demand sequence data and the inherent periodicity of such data. Let $Y^t = [Y_{(z_{i_1}, z_{j_1})}^t, \dots, Y_{(z_{i_N}, z_{j_N})}^t] \in \mathbb{R}^{N \times 1}$, where N is the number of ODs, and ${}^B X^t = [{}^B X_{(z_{i_1}, z_{j_1})}^t, \dots, {}^B X_{(z_{i_N}, z_{j_N})}^t] \in \mathbb{R}^{N \times 4}$. The problem can be described as follows:

$$Y^t = F({}^B X^t, \ddot{A}) \quad (1)$$

where F is the prediction function.

A. ST-ED-RMGC

Figure 1 displays the overall architecture of the model. The model is an encoder-decoder based model, where the encoder is composed by a temporal encoder which takes into account the temporal dependencies of all OD pairs (and not just one at a time), and a spatial encoder with several Residual-Multi-Graph-Convolutional network (RMGC) which takes several adjacency matrix and the graph OD demand as inputs. The RMGC combines a residual module with a multi-graph convolution, to capture the spatial correlation between OD pairs. The residual module is introduced to tackle the issue of gradient explosion/gradient vanishing in complex deep networks. The multi-graph convolution applies graph convolution on stacked weighted adjacency matrices.

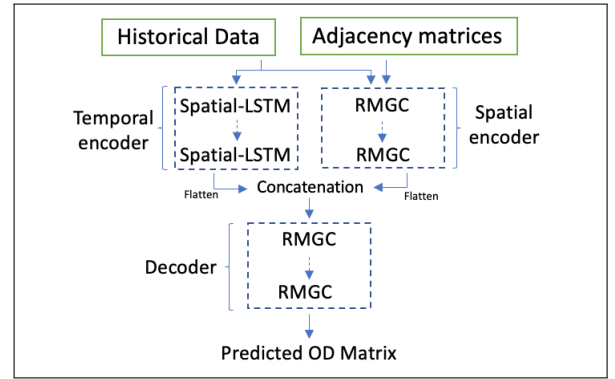


Fig. 1. Framework of the ST-ED-RMGC model, adapted from [18]

B. Adjacency Matrices

The relationships introduced in Sec. II can refer to a spatial relationship or a semantic relationship. As proposed in [18], a spatial relationship between two OD pairs can be expressed via two relationships: an origin-based, and a destination-based, denoting a relationship on the origins (respectively destinations) of the two OD pairs. The neighbourhood (A_n^O, A_n^D) and centroid distances (A_d^O, A_d^D) are used as spatial relationships. We defined semantic relationships in the same way, where a semantic relationship between two areas is not dependent on their geographical location. Thus, the functionality between two areas is translated into an origin-based and destination-based functionality adjacency matrix (A_f^O, A_f^D) . The functionality of an area is represented by a vector containing the socio-economic information of an area, such as the presence of a railway station, or the density of housing. A relationship on the correlation between the historical demands of two OD pairs is also defined (A_{corr}) . In this paper, we then define seven different adjacency matrices, which are all concatenated in \ddot{A} to capture the dependencies between zones.

$$\ddot{A} = [A_n^O, A_n^D, A_d^O, A_d^D, A_f^O, A_f^D, A_{corr}] \quad (2)$$

C. Road sensor data integration

Differently from the original ST-ED-RMGC approach, which only uses bike-sharing historical data as the main input to the forecasting model, we also decided to integrate, as an additional input, information related to the (historical) car traffic flow, as observed in the proximity of the origin of the OD bike-sharing flow we are trying to predict. The main idea is that this information could allow capturing the inter-dependence between car and bike usages, which could be helpful to anticipate, *e.g.*, modal changes such as people switching from bikes to cars in the presence of degraded weather conditions or other kinds of perturbations. From an architectural perspective, the data from road sensors are concatenated to ${}^B X_{(z_i, z_j)}^t \forall i, j, k$. Let I_i^t be the total car traffic flow recorded in zone i during time slot t . We define a new feature vector:

$${}^I X_{(z_i, z_j)}^t = [{}^B X_{(z_i, z_j)}^t, I_i^{t-T_{I1}}, \dots, I_i^{t-T_{I2}}] \forall i, j \quad (3)$$

where $(T_{I1} - T_{I2} + 1)$ is the depth of historical car flow observations considered, with $T_{I1} > T_{I2}$

D. Weather data integration

In the same way, we decided to integrate information related to the historical weather, as well as weather forecast. Knowing that bike-share use is impacted before, during or after rain, the interest here is in capturing patterns in changes in bike-share use related to degraded weather. Let W^t a sequence of features from the weather dataset, which will be concatenated to ${}^B X_{(z_i, z_j)}^t \forall i, j$. W^t can be composed of historical data such as $[v^{t-T_{W1}}, \dots, v^{t-T_{W2}}]$, where v can be set as hr , corresponding to the hourly rainfall ($mm.h^{-1}$), or hd , corresponding to the hourly rainfall duration (min per hour), depending on the considered model variation. $(T_{W1} - T_{W2} + 1)$ is the number of weather historical time-steps provided as an input. Note that $T_{W1} > T_{W2}$ and T_{W2} can be negative or zero, which corresponds to the use of a weather prediction for the next $-T_{W2}$ time-steps. W^t can also contain weather forecast such as dcr^t ($mm.h^{-1}$ per day), *i.e.*, the daily cumulative rainfall recorded the day associated to the date t . We thus define the new feature vector:

$${}^W X_{(z_i, z_j)}^t = [{}^B X_{(z_i, z_j)}^t, W^t] \forall i, j. \quad (4)$$

We can finally define a new feature vector including both contextual weather and sensor information as follows:

$${}^{W,I} X_{(z_i, z_j)}^t = [{}^B X_{(z_i, z_j)}^t, W^t, I_i^{t-T_{I1}}, \dots, I_i^{t-T_{I2}}] \forall i, j \quad (5)$$

E. Time Encoding

We have integrated a time-encoding component to take into account temporal contextual information. Demand for bike sharing, and more generally, any transport mode, is calendar-dependent. It is normally linked to working hours, and therefore directly related to the day of the week or the time of the day. Demand in university zones is for instance impacted by school vacations. Similarly, in business districts, employees with children can take vacations accordingly. For these reasons, we decided to integrate different calendar-related features as contextual data: weekday, whether we are trying to predict a business day, a school holiday, a day of departure for school holidays or a return day from school holidays. Each of these r features have been encoded as a binary vector T_i whose dimension is equal to the number of classes in the feature i , except for the feature "hour". As for the feature "hour", we have labelled the 24 possible hours in 18 classes. From 1 to 17 for the first 17 hours from 6 a.m., and 0 for the others. It avoids adding unnecessary information, as the bike ODs during the night hours are almost zero. As for the others features, they have been labelled from 0 to the number of possible classes (*e.g.*, 0,1,..6 to label the 7 days of the week). Then, each labelled features has been encoded in a binary vector of the dimension of the number of classes in the labelled feature. The final output of this module is a list of vectors: T_1, \dots, T_r .

F. Time Embedding

We have chosen to implement an Embedding module (see Fig 2) to take into account temporal contextual information and the dependencies between them. The embedding

module contains r embedding layers in parallel, each of them corresponding to a unique $T_i, i \in [1, r]$. Each of the outputs of the embedding layers passes through a dense layer, before being concatenated and finally passing into a dense module (composed of 3 dense layers). Dense layers capture the dependencies between each of the time-encoded vectors. The output of the dense module is E_T , a vector of dimension p . This vector represents the temporal information of the date t that we are trying to predict. As the time embedding is the same for each ODs (*i.e.*, the time embedding does not capture spatial dependencies), E_T is then stacked as many times as the number of selected ODs, which gives a 2D matrix $[E_T, \dots, E_T]$ of size $[N, p]$. Finally, this matrix is concatenated (Fig 3) to the feature vector X to obtain a new feature vector of dimension $[N, L+p]$.

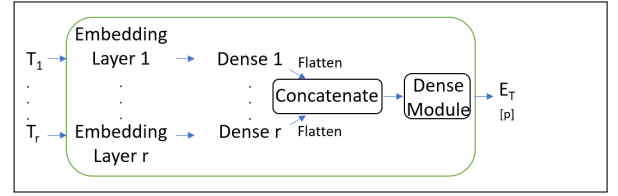


Fig. 2. Embedding module. The Dense Module is composed of 3 dense layers in series.

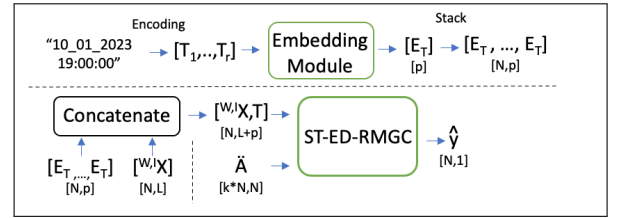


Fig. 3. Integration of the embedding module into the existing architecture.

III. DATASETS

A. Bike-sharing Data

We work with bike-sharing data provided by JC Decaux. We have actual drop-in and drop-out data per station. Working by OD between stations does not make sense because the dynamic OD matrix will be sparse and would not represent user mobility. Stations have physical limitations: a finite set of spatially localised stations with limited vehicle capacity. Many users, therefore, use the bikes available at stations within a certain radius of their departure area to reach a certain radius of their arrival destination. It was thus necessary to aggregate the bike share data not only temporally, but also spatially. The stations were first aggregated by IRIS zones (Fig 4), which are zones developed by the French Institute of Statistics. This segmentation divides the conurbation of Lyon into small geographical areas, each grouping approximately 2000 inhabitants¹. The IRIS areas are thus used to group bike-sharing stations according to socio-economic criteria. To reduce the sparsity of the data, we then decided to further aggregate pairs of IRIS zones according to an iterative

¹<https://www.insee.fr/fr/metadonnees/definition/c1523>

procedure described below, based on three criteria while maintaining the spatial homogeneity of the division. After this aggregation, we end up with a very sparse OD matrix (because all zones are connected). We filtered the ODs by setting a proportion p_{bike} of total demand between 7 a.m. and 9 p.m. and taking the minimum number of ODs needed to reach this proportion.

B. Spatial Aggregation

The adopted aggregation (Fig 4) procedure builds upon the following three criteria: the proximity between two IRIS zones, the sum of their area and the common perimeter between two adjacent areas. Specifically, at each iteration, we look for the two IRIS zones i^* and j^* minimizing the following objective:

$$(i^*, j^*) = \underset{(i,j) \in \mathcal{Z}_{\text{adj}}}{\text{argmin}} \left(\left\{ \frac{1}{P_{i,j}} (s_i + s_j) \right\} \right) \quad (6)$$

where \mathcal{Z}_{adj} is the set of adjacent pairs of zones in the current spatial aggregation, s_i the surface of the zone i , and $P_{i,j}$ is the common perimeter between i and j . Zones are aggregated up to a fixed number of zones.

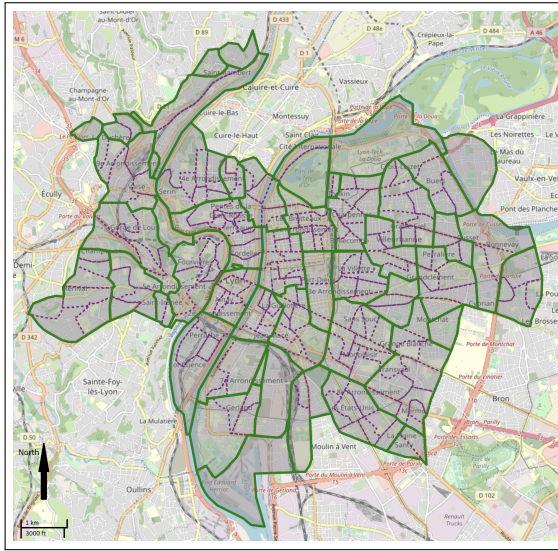


Fig. 4. IRIS zones without any aggregation (dashed line) and the aggregation into 50 zones (solid line).

C. Car Flow Data

Data from loop detectors provide information on the flow and loop occupancy rate of cars, recorded on the loop. Our initial dataset provided counts aggregated every 6 minutes. The data were first aggregated by hour and by loop, keeping a sample if the loop recorded more than 4 periods of 6 minutes in an hour. A period is kept if it does not represent erroneous data, *i.e.*, an abnormal loop occupancy rate ($> 50\%$), or a NaN value flow. Otherwise, we remove the data, as they are not statistically correct. Each of the aggregated areas contains several loops. Data were then aggregated by zone. At the end of the preprocessing, only about 10 periods of 1 hour could not be determined. For such hours we simply adopted a linear interpolation.

D. Weather Data

For the integration of the weather context as described in Sec. II-D, we leveraged data provided by Meteo France² and collected hourly from 2 weather stations of the city of Lyon, reporting weather conditions (*e.g.*, hourly rainfall intensity and duration, temperature, *etc.*) from January 1st, 2019, to December 31st, 2020. The two stations have been aggregated by taking the average of the values on both. Weather data are therefore independent of the zone, but dependent on the hourly time slot t .

IV. EXPERIMENTS

The train dataset used in our experiments includes hours from January 8th, 2019 to November 9th, 2019, while the test dataset covers the 7 a.m. and 9 p.m. period for all days between December 1th, 2019, and March 10th, 2020. The 220 IRIS zones from Lyon and Villeurbanne have been aggregated into 50 zones (Fig 4), which correspond to 2500 OD pairs. We decided to perform the prediction exclusively on the minimum number of ODs necessary to represent $p_{\text{bike}} = 60\%$ of the bike-sharing demand from 7 a.m. to 9 p.m. This meant that we only worked with 130 of the 2500 ODs of our territory. The algorithm was trained on all the hours of the day (contrary to the test scenarios limited between 7 a.m. and 9 p.m.).

A. Weather-related Test Scenarios

In addition to the global test dataset defined above and used to evaluate the overall performance of each model, we identified *weather-related scenarios* to assess the accuracy of bike-sharing demand prediction in degraded weather conditions. Specifically, we used such test scenarios to comparatively evaluate the proposed models, which include contextual data, with respect to the baseline ST-ED-RMGC model that does not include them. The weather-related scenarios are built as subsets of the global test dataset, based on the daily cumulative rainfall levels (dcr) as well as the hourly amount of rain (hr). Indeed, the user's choice of bike sharing as travel mode can be strongly influenced by an ongoing rain episode, an expected rainfall, as well as an already concluded rainfall episode [16]. We thus include in the weather-related scenarios hours (between 7 a.m. and 9 p.m.) with rain ($hr > 0$), without rain ($hr = 0$), and with daily cumulative rainfall falling within specific positive ranges, as detailed in Tab I.

TABLE I
WEATHER-RELATED SCENARIOS

n°	scenario	nb of hours	% of 0 values	n°	scenario	nb of hours	% of 0 values
0	test dataset between 7 a.m. and 9 p.m.	1500	12	3	$dcr = 0$	960	11
1	$hr = 0$	1344	12	4	$dcr \in [0, 1]$	120	11
2	$hr > 0$	156	20	5	$dcr \in [1, 3]$	240	13

²<https://meteofrance.com>

B. Model Settings

The architecture parameters, such as the number of neurons and the number of layers, are those proposed in the original ST-ED-RMGC paper [18]. The default values of other hyperparameters in our experiments are set up as follows. We set the length of a time-step to 1 hour. The optimizer used in the model is Adam with a learning rate of $5e^{-5}$ and a decay of $1e^{-6}$. We set the dropout to 0.7, the batch size to 16, and the number of epochs to 80 for all models. The loss function used to train the model is based on the Mean-Squared Error (MSE). These hyperparameters were chosen by validation with respect to the reference model, corresponding to the basic ST-ED-RMGC corresponding to the first line in Tab II. All other models were trained with these same hyperparameters, ensuring that each of them was fully trained with no overfitting. In Tab II, the column *Model Type* denotes the type of embedding or contextual data which have been included (or not) within a specific model. *X* denotes the usual feature vector without any additional information, *T* denotes the one with time embedding, *W* the one with weather data, *I* the one with road car flow data. *WIT* indicates a combination of *W*, *I*, and *T*. The column *Model Number* represents the number associated to the model variation being considered. *e.g.*, the 5 models with type *I* will be called *I1*, *I2*, *I3*, *I4*, and *I5*. The column *Included Context Elements* corresponds to the detailed description of the contextual elements considered within the model variation.

C. Metrics

The metrics used to evaluate the models are the MSE and the Mean Absolute Percentage Error (MAPE). The OD matrix contains a large number of demands close to 0, that is why the MAPE is high (> 0.5), since many predictions are close in absolute terms but the error is relatively high (e.g. real demand equal to 1 and prediction equal to 3). We highlight here that the metrics depend on the demand. It is therefore not possible to directly compare the performance of two different scenarios with each other, because the related datasets are different in terms of bike-sharing demand. It is nonetheless possible to compare models between them on the same scenario, as detailed below.

D. Results

Tab. II reports the prediction results (*MSE* and *MAPE*) for scenario $n^{\circ}0$ from Tab. I Tables III:V report instead the performance on the weather-related scenarios, with specific focus on the model variations including weather-related (Tab. III), flow-related (Tab. IV) and time embedding (Tab. V) features, respectively.

1) *Model with weather features*: Both in the case of scenarios without rain ($hr = 0, dcr = 0$, Tab. III), and in the scenario $n^{\circ}0$ (Tab. II), models with weather features show no significant change in the prediction quality compared to the baseline simpler model (X). However, in scenarios with rainfall ($hr > 0$), all models with weather information outperform the simple model. The model (W7) achieves the lowest MSE and MAPE in the scenario ($hr > 0$,

TABLE II
OVERALL PERFORMANCE COMPARISON

Model Type	Model Number	Included Context Elements	MSE	MAPE
X			8.70	0.570
T		E_T	8.22	0.532
W	1	hr^{t-1}	8.63	0.560
	2	hr^{t-2}, hr^{t-1}	8.90	0.582
	3	$hr^{t-3}, hr^{t-2}, hr^{t-1}$	8.62	0.549
	4	hr^t	8.67	0.583
	5	hr^{t-1}, hr^t	8.66	0.553
	6	hr^{t-1}, hd^{t-1}	8.86	0.575
	7	hr^t, hd^t	8.56	0.55
I	1	I^{t-1}	8.78	0.588
	2	I^{t-2}, I^{t-1}	8.70	0.572
	3	I^t	8.68	0.562
	4	I^{t-1}, I^t	8.47	0.555
	5	I^{t-2}, I^{t-1}, I^t	8.44	0.562
WIT		$hr^{t-1}, hr^t, hr^{t-1}, I^{t-1}, I^t, E_T$	8.21	0.528

Tab. III), with approximately 20% and 27% improvement, respectively, compared to the (X) model. Notably, for this scenario, the (W7) model outperforms all other models, surpassing the second-best model (WIT) by 4% in MSE and 9% in MAPE. This means that the duration of the rain is directly linked to the bike-sharing demand. Compared with the other models, the (W7) model knows how long the rain will last over the next hour, and therefore the amount of time it will not rain in the next hour. There is strong reason to believe that as soon as the rain stops, a certain proportion of users will quickly resume their regular BSS use. For example, if employees who use bike-sharing are in the office, they will wait a few minutes for the rain to stop before moving.

2) *Model with flow features*: Tab IV and Tab II occasionally exhibit a slight improvement in the prediction quality with respect to the baseline (X) model. The improvement ranges from 0% to 6% in MSE and 0% to 3% in MAPE, depending on the models and scenarios (I1, I2, I3, I4, I5).

TABLE III
MODEL WITH WEATHER-RELATED DATA

scenario	model	mse	mape	scenario	model	mse	mape
$hr = 0$	X	8.71	0.551	$dcr \in [0, 1]$	X	8.57	0.578
	W1	8.69	0.547		W1	8.44	0.565
	W2	8.94	0.568		W2	8.86	0.59
	W3	8.7	0.537		W3	8.66	0.553
	W4	8.87	0.58		W4	8.47	0.597
	W5	8.8	0.548		W5	8.68	0.562
	W6	8.92	0.563		W6	8.5	0.577
	W7	8.8	0.55		W7	8.48	0.562
WIT	8.33	0.521	WIT	7.9	0.538		
$hr > 0$	X	8.68	0.747	$dcr \in [1, 3]$	X	8.7	0.607
	W1	8.1	0.688		W1	8.79	0.597
	W2	8.59	0.716		W2	9.21	0.619
	W3	7.99	0.663		W3	8.85	0.579
	W4	6.97	0.611		W4	8.43	0.605
	W5	7.48	0.606		W5	8.47	0.571
	W6	8.35	0.689		W6	9	0.605
	W7	6.88	0.546		W7	8.29	0.562
WIT	7.15	0.597	WIT	8.14	0.541		
$dcr = 0$	X	8.86	0.541				
	W1	8.84	0.539				
	W2	9.06	0.56				
	W3	8.8	0.531				
	W4	9.03	0.569				
	W5	8.95	0.538				
	W6	9.07	0.556				
	W7	8.96	0.541				
WIT	8.52	0.518					

However, considering the marginal improvement achieved by integrating these features in the specific analysed scenarios, it is not possible to draw conclusive evidence regarding the usefulness of their utilization.

TABLE IV
MODEL WITH FLOW DATA

scenario	model	mse	mape	scenario	model	mse	mape
$hr = 0$	X	8.71	0.551	$hr > 1$	X	7.22	0.978
	I1	8.78	0.568		I1	7.91	1.05
	I2	8.7	0.553		I2	7.19	0.991
	I3	8.71	0.543		I3	6.78	0.974
	I4	8.51	0.535		I4	7.03	0.999
	I5	8.45	0.542	I5	7.31	0.985	
$hr > 0$	X	8.68	0.747				
	I1	8.87	0.78				
	I2	8.68	0.745				
	I3	8.4	0.741				
	I4	8.12	0.737				
	I5	8.44	0.745				

3) *Model with time embedding*: The application of time embedding results in an overall improvement in the results compared to the simple model (X), with lower MSE and MAPE in the order of 5.5% and 6.6%, respectively. Upon examining Tab V, it can be noted that predictions are superior regardless of the scenario. A scenario without rain demonstrates improved performance, with a reduced MSE and MAPE of 4.8% and 6.2%, respectively. The improvement is even more pronounced in scenarios with rainfall. For example, dates with rainfall presence ($hr > 0$) exhibit a significantly smaller MSE and MAPE, reduced by 12.4% and 10.8% respectively.

TABLE V
TIME EMBEDDING MODEL COMPARISON

scenario	model	mse	mape	scenario	model	mse	mape
$hr = 0$	X	8.71	0.551	$dcr = 0$	X	8.86	0.541
	T	8.29	0.517		T	8.49	0.515
$hr > 0$	X	8.68	0.747				
	T	7.6	0.666				

V. CONCLUSION AND FUTURE WORK

The results of this study emphasize the relevance of contextual data in OD bike sharing demand forecast, especially during degraded weather-related situations. In future work, we plan to study scenarios without rain ($hr = 0$) among the days where a certain amount of rain has occurred (dcr scenarios within a positive interval). This would enable studying the extent to which users are impacted by the risk or occurrence of rain and assess whether weather data can better anticipate user choices, even when it ultimately does not rain. The inter-modal relationships between (car) traffic flow in a specific area and BSS demand have not been extensively explored. A finer zoning approach might lead to more interesting conclusions. Finally, further research in this direction includes focusing on other inter-modal travel demand relationships, such as public transportation and BSS demand.

ACKNOWLEDGMENT

This research is supported by the French ANR research projects PROMENADE (grant number ANR-18-CE22-0008) and MOBITIC (grant number ANR-19-CE22-0010).

REFERENCES

- [1] K. Kim. Investigation on the effects of weather and calendar events on bike-sharing according to the trip patterns of bike rentals of stations. *Journal of Transport Geography*, 66:309–320, 2018.
- [2] C. Zhang, L. Zhang, Y. Liu, and X. Yang, “Short-term prediction of bike-sharing usage considering public transport: A lstm approach,” in 2018 21st International Conference on Intelligent Transportation Systems (ITSC). IEEE, 2018, pp. 1564–1571.
- [3] R. Buehler, A. Hamre, et al. Economic benefits of capital bikeshare: A focus on users and businesses. Technical report, Mid-Atlantic Universities Transportation Center, 2014.
- [4] Hua, M., Chen, J., Chen, X., Gan, Z., Wang, P. and Zhao, D. (2020), Forecasting usage and bike distribution of dockless bike-sharing using journey data. *IET Intell. Transp. Syst.*, 14: 1647-1656. <https://doi.org/10.1049/iet-its.2020.0305>
- [5] An, J., Fu, L., Hu, M., Chen, W., and Zhan, J. (2019). “A novel fuzzy-based convolutional neural network method to traffic flow prediction with uncertain traffic accident information.” *IEEE Access*, 7, 20708–20722
- [6] Bruna, J., Zaremba, W., Szlam, A., and Lecun, Y. (2014). “Spectral networks and locally connected networks on graphs.” *International Conference on Learning Representations (ICLR2014)*, CBL5, April 2014
- [7] Yu, B., Yin, H., and Zhu, Z. (2017). “Spatio-temporal graph convolutional neural network: A deep learning framework for traffic forecasting.” *CoRR*, abs/1709.04875 casting.” *CoRR*, abs/1707.01926
- [8] Chen, L., Jakubowicz, J., Zhang, D., Wang, L., Yang, D., Ma, X., Li, S., Wu, Z., Pan, G., and Thi Mai Trang, N. (2016). “Dynamic cluster-based over-demand prediction in bike sharing systems.” 841–852 (09)
- [9] Chai, D., Wang, L., and Yang, Q. (2018). “Bike flow prediction with multi-graph convolutional networks.” *CoRR*, abs/1807.10934
- [10] Ezgi Eren, Volkan Emre Uz, A review on bike-sharing: The factors affecting bike-sharing demand, *Sustainable Cities and Society*, 2020, <https://doi.org/10.1016/j.scs.2019.101882>.
- [11] Li, Y., Shuai, B. Origin and destination forecasting on dockless shared bicycle in a hybrid deep-learning algorithms. *Multimed Tools Appl* 79, 5269–5280 (2020). <https://doi.org/10.1007/s11042-018-6374-x>
- [12] Wang, Y., Yin, H., Chen, H., Wo, T., Xu, J., and Zheng, K. (2019). “Origin destination matrix prediction via graph convolution: A new perspective of passenger demand modeling.” *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery Data Mining, KDD '19*, New York, NY, USA, Association for Computing Machinery, 1227–1235 <https://doi.org/10.1145/3292500.3330877>
- [13] Zhang, J., Che, H., Chen, F., Ma, W., and He, Z. (2021). “Short-term origin-destination demand prediction in urban rail transit systems: A channel-wise attentive split-convolutional neural network method.
- [14] Jiang, W. Bike sharing usage prediction with deep learning: a survey. *Neural Comput & Applic* 34, 15369–15385 (2022). <https://doi.org/10.1007/s00521-022-07380-5>
- [15] J. Corcoran, T. Li, D. Rohde, E. Charles-Edwards, and D. Mateo-Babiano. Spatio-temporal patterns of a public bicycle sharing program: the effect of weather and calendar events. *Journal of Transport Geography*, 41:292–305, 2014.
- [16] S. Reiss and K. Bogenberger. Validation of a relocation strategy for munich’s bike sharing system. *Transportation Research Procedia*, 19:341–349, 2016. *Transforming Urban Mobility. mobil.TUM 2016. International Scientific Conference on Mobility and Transport. Conference Proceedings.*
- [17] F. Sun, P. Chen, and J. Jiao. Promoting public bike-sharing: A lesson from the unsuccessful pronto system. *Transportation Research Part D: Transport and Environment*, 63:533–547, 2018
- [18] J. Ke, X. Qin, H. Yang, Z. Zheng, Z. Zhu, and J. Ye. Predicting origin-destination ride-sourcing demand with a spatio-temporal encoder-decoder residual multi-graph convolutional network. *Transportation Research Part C: Emerging Technologies*, 2021
- [19] R. Rochas, A. Furno, and NE. El Faouzi. Empirical analysis of the forecasting accuracy of ST-ED-RMGC with bike-sharing data under atypical weather-related scenarios. *Proceedings of the 14th International conference on application of statistics and probability in civil engineering [to appear]*.