14th International Conference on Applications of Statistics and Probability in Civil Engineering, ICASP14
Dublin, Ireland, July 9-13, 2023

Empirical analysis of the forecasting accuracy of ST-ED-RMGC with bike-sharing data under atypical weather-related scenarios

Romain Rochas
Phd candidate, Dept. of Transportation, ENTPE, Univ. Gustave Eiffel, LICIT-ECO7 UMR_T9401, F-69675, Lyon, France

Angelo Furno
Professor, Dept. of Transportation, ENTPE, Univ. Gustave Eiffel, LICIT-ECO7 UMR_T9401, F-69675, Lyon, France

Nour-Eddin El Faouzi
Professor, Dept. of Transportation, ENTPE, Univ. Gustave Eiffel, LICIT-ECO7 UMR_T9401, F-69675, Lyon, France

ABSTRACT: The transport sector is one of the most greenhouse gas emitting sectors, and users waste a lot of time in transport due to congestion, with relevant socio-economic consequences. Understanding the spatio-temporal traffic patterns could help traffic forecasts and allow to adjust the transportation infrastructure more efficiently in order to cope with these challenges. Deep learning has had a huge success in recent years due to its ability to capture patterns to make predictions that were not able to be properly assimilated by older regression models. More recently, Graph Neural Networks (GNN) have emerged as the state of the art in capturing spatio-temporal traffic patterns thanks to their capability to model the structure of the traffic network and to capture its dynamics. The present study analyzes the impact of weather scenarios in the case of bike-sharing orgin-destination matrix prediction. In this study, a type of Graph Neural Network (ST-ED-RMGC) has been employed. We consider real world bike-sharing data related to urban areas. The results of the study show a strong impact of the weather on the quality of the prediction, whether it is an atypical temperature, a strong wind or the presence of rain. This suggests the need for introducing as input features contextual information about weather conditions as well as data from other sources, such as historical demand information on transport modes alternative to the one being predicted, in order to quickly identify the onset of atypical events

1. INTRODUCTION

With the advancement of neural network algorithms and, more recently, graph neural networks, traffic metrics prediction has been constantly improving in the last decade. However, these algorithms demonstrate limitations in predicting atypical events (An et al. (2019)), as they learn to recognize patterns and adapt their responses based on prior experiences. With climate change and other hazards such as the Covid crisis, atypical events are increasingly frequent, making it imperative to identify and address the challenges faced by forecasting models. To address these limitations in traditional neural-network forecasting solutions, one idea is to generate data associated with atypical scenarios using simulation tools, and integrate more targeted contextual information into our inputs. In this paper, we focus on the state-of-the-art Origin-Destination (OD)-based forecasting model proposed by Ke et al. (2021), which is one of the most recent and powerful OD-based forecasting models. Herein, we apply the model to bike-sharing

data of the city of Lyon, France, contrary to their study which refers to for-hire vehicles in Manhattan. The main objective of this study is to evaluate the performance of forecasting models in atypical scenarios, compared to typical reference scenarios.

The main contributions of this study are summarized as follows :

1. We introduce a custom spatial aggregation approach and describe briefly how the flows are distributed in the identified zones.
2. We define multiple weather-related atypical scenarios in order to evaluate the performance of the selected forecasting algorithm. We prove that rainfall is the factor with the strongest impact on prediction quality. Wind also affects the quality of the prediction, but to a lesser extent.

In Sec. 2, we present recent literature on the topic of traffic forecasting and performance evaluation of neural-network based solutions. In Sec. 3, we present the overall structure of the ST-ED-RMGC proposed by Ke et al. (2021), the differences we made in the spatial dependency modeling between OD pairs, our spatial aggregation and our test scenarios. In Sec. 4, we detail the results obtained on all our tested scenarios. Sec. 5 concludes the work and highlights future research directions.

2. RELATED WORK

Graph Neural Networks (GNNs)

Previous research has examined multiple aspects of traffic flow forecasting, and graph neural networks (GNNs) are considered the cutting-edge technology in traffic prediction due to their ability to capture spatial-temporal dependencies. The primary technique is graph convolution, which was first introduced by Bruna et al. (2014) and later applied to transportation forecasting by Yu et al. (2017). Derivatives such as the diffusion graph convolution have also been developed Li et al. (2017).

Concerning bike-sharing prediction, different approaches can be utilized, including predictions by station, as in Chai et al. (2018), or through dynamic station aggregation via clustering before predicting demand, as proposed by Chen et al. (2016). Region-based approaches involve using a deep neural network to predict inflow and outflow for each cell of the grid, as presented by Zhang et al. (2016).

Due to the lack of real-time data and the inherent complexity of OD prediction - with the destination being unknown until arrival - the forecasting of ODs is particularly challenging (see Ke et al. (2021), Zhang et al. (2021)). Furthermore, data sparsity represents a relevant challenge (Wang et al. (2019)), as the number of ODs is quadratic in the number of zones, along with the current limitations of traffic forecasting. As a result, limited research exists on the subject, especially in the context of bike-sharing prediction.

Taking into account contextual information

GNNs were used to capture spatial information, and then methods were developed to integrate external factors or inter-modal links. Geng et al. (2019a) use POIs data to add similarity relationship between zones and apply a multi graph convolution, while Zhu et al. (2021) proposed a neural network model that includes static information such as POIs, and dynamic contextual information such as weather. Ma et al. (2021) proposed a model that incorporates contextual information such as inter-day and intra-day traffic pattern by using day of the week, season, weather and holiday to enhance traffic flow prediction. However, the article treats the sensors independently by not using GNN. Generally these data are integrated in inputs by weighted matrices that extract relationships from graphs. Studies are therefore interested in multi-graph convolution. Geng et al. (2019b) proposed a model for heterogeneous spatial units, but did not emphasize the links between transport modes. On the contrary, Ye et al. (2019) have emphasized the relationships between two transport modes by proposing a co-prediction of bike-sharing and taxi pick-up and drop-off demand. Also, Liang et al. (2022b), and Liang et al. (2022a) addressed the issue of station-based demand prediction by involving multi-relational graph relations between bike and other modes.

Taking into account atypical values

Castro-Neto et al. (2009) developed a support vector regression-based model for forecasting flow under typical and atypical conditions, but results

were limited and the model did not use neural networks or contextual information.

3. METHODOLOGY

This section presents the operation of the ST-ED-RMGC in more detail, by focusing on our adaptation of the algorithm to the context of bike-sharing OD demand prediction. We also describe the performed spatial aggregation of the bike-sharing stations.

Problem statement

Let $\mathscr{Z}$ be the set of areas corresponding to a partitioning of the analyzed city and $OD_{\mathscr{Z}} = \{(z_i, z_j) \forall (z_i, z_j) \in \mathscr{Z}^2 | i \neq j\}$ be the set of all pairs of zones. $Y^t_{(z_i,z_j)} \in \mathbb{R}$ is the demand between (z_i, z_j) at t time-step. It is defined as the outflow demand for destination z_j that left from z_i at t. Let $X^t_{(z_i,z_j)} \in \mathbb{R}^S$ be the associated historical sequence which allows to predict $Y^t_{(z_i,z_j)}$ and S is the historical sequence length. Let $G_u(V, E, A_u)$ a graph associated to a relationship u and A_u its adjacency matrix which takes into account the spatial dependency. The temporal dependencies are instead taken into account thanks to historical demand and its periodicity. Let $Y^t = [Y^t_{(z_{i_1},z_{j_1})}, ..., Y^t_{(z_{i_N},z_{j_N})}]$, where N is the number of ODs and $X^t = [X^t_{(z_{i_1},z_{j_1})}, ..., X^t_{(z_{i_N},z_{j_N})}]$. The problem can be described as follow :

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

where F is the prediction function and $\ddot{A}$ is the concatenation of all adjacency matrices.

3.1. ST-ED-RMGC

We have chosen to work on Ke et al. (2021) algorithm as it is one of the few addressing the issue of forecasting by OD. 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 spatial structure of the OD graph, and a spatial encoder with a 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 in complex deep networks. The multi-graph convolution allow apply graph convolution on stacked weighted adjacency matrices.

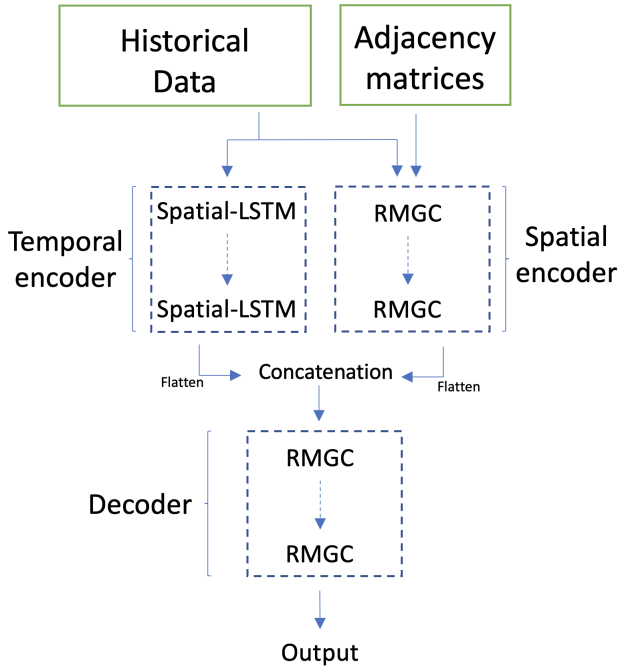

Figure 1: framework of the ST-ED-RMGC model, adaptated from Ke et al. (2021)

3.2. Modelling spatial dependencies between OD pairs

Adjacency matrices allow to take into account of the spatial dependency. Research papers usually use neighborhood relationship graph, similarity graph between zones, correlation graph or other kinds of graph built via other spatial relationship. Defining adjacency matrices that make sense in a OD approach is less intuitive. As for non symmetrical relationship, Ke et al. (2021) define origin-based and destination-based adjacency matrices to cope with this issue. In this paper, we use the same type of adjacency matrices, including: the *neighborhood relationship graph*, the *centroid-distance graph*, the *functional similarity graphs* and the *mobility pattern correlation graph*. Only the centroid distance graph and the similarity graph have been modified compared to Ke et al. (2021). Thus, these are the only ones we detail in the following.

3.2.1. Centroid distance graphs

Centroid distance graphs have been introduced to treat irregular areas, or areas that are close but not necessarily adjacent. Let $c(*)$ be the function extracting the centroid from the polygon describing the geometry of a generic area. Centroid distance graphs are built following Shuman et al. (2013) via a Gaussian kernel weighting function:

$$f(x,y) = exp^{(-\frac{\text{dist}(x,y)}{\delta^2})} \tag{2}$$

where 'dist' is the euclidean distance, x and y are positions vector, and δ is the standard deviation of the euclidean distance matrix, which are different between the origin-based one and the destination-based one. Let $k,l \in OD_{\mathscr{Z}}$ such that $k = (o_k, d_k)$ and $l = (o_l, d_l)$. A_d^O and A_d^D are respectively the origin-based and destination-based distance matrix:

$$[A_d^O]_{k,l} = f(c(o_k), c(o_l)) \tag{3}$$

$$[A_d^D]_{k,l} = f(c(d_k), c(d_l)) \tag{4}$$

3.2.2. Functional similarity relationship graph

Functional similarity represents the probability that traffic patterns are similar in two given areas a_i, a_j based on contextual information related to the two areas. The context features considered hereinafter are: population density, housing density, amount of train stations and road density. Population density is defined as the number of people per unit of area, while house density is defined as the number of housing units per unit of area. These two features can provide an indication about the transport demand. Road density is defined as the ratio of the total length of roads within a defined area to the total area, this can provide an indication of the accessibility and connectivity of the urban transport network. As train stations are transit areas, the amount of train stations can provide an indication about transport demand and public transport accessibility. We define the origin-based and destination-based functional similarity matrices as follows:

$$[A_f^O]_{i,j} = \left[\sqrt{(F_i^O - F_j^O)(F_i^O - F_j^O)^T}\right]^{-1} \tag{5}$$

$$[A_f^D]_{i,j} = \left[\sqrt{(F_i^D - F_j^D)(F_i^D - F_j^D)^T}\right]^{-1} \tag{6}$$

where F_i^O, F_j^O, F_i^D and F_j^D are vectors of functionalities.

3.3. Feature vector

Let $k \in OD_{\mathscr{Z}}$. The input feature vector at time t for the generic OD pair k, *i.e.,* X_k^t, includes the historical demand in the form of tendency, daily periodic pattern and weekly periodic pattern.

1. Tendency: the demand in the OD graph at the last two time intervals $Y_k^{(d,t-1)}, Y_k^{(d,t-2)}$ on the day 'd'.
2. Daily periodic feature: the demand in the OD graph at the same time interval in the last day: $Y_k^{(d-1,t)}$.
3. Weekly periodic feature: the demand in the OD graph at the same time interval in the last week: $Y_k^{(d-7,t)}$.

Formally:

$$X_k^t = [Y_k^{(d-7,t)}, Y_k^{(d-1,t)}, Y_k^{(d,t-1)}, Y_k^{(d,t-2)}] \tag{7}$$

$$X^t = [X_1^t, ..., X_N^t]^T \tag{8}$$

3.4. Spatial Aggregation

First, the bike-sharing stations have been aggregated according to the French IRIS segmentation, developed by the French Institute of Statistics. This segmentation divides the conurbation of Lyon into small geographical areas, called IRIS sectors, each grouping approximately 2000 inhabitants[1]. IRIS areas are thus used to group bike-sharing stations according to socio-economic criteria. Then, in order to reduce data sparsity, we decided to further aggregate pairs of IRIS zones according to an iterative procedure based on three criteria maintaining the spatial homogeneity of the division: the proximity between two IRIS zones, the sum of their area and the common perimeter between two adjacent areas. At each iteration, we are looking for the two IRIS zones i^* and j^* minimizing the following objective:

$$(i^*, j^*) = argmin_{(i,j)\in Z_{\text{adj}}}\left(\left\{\frac{1}{P_{i,j}}(s_i + s_j)\right\}\right) \tag{9}$$

[1] `https://www.insee.fr/fr/metadonnees/definition/c1523`

where $\mathscr{L}_{\text{adj}}$ 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.

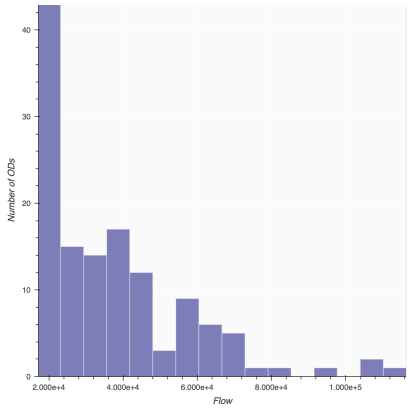

Figure 2: Histogram of the flow from the 130 selected ODs

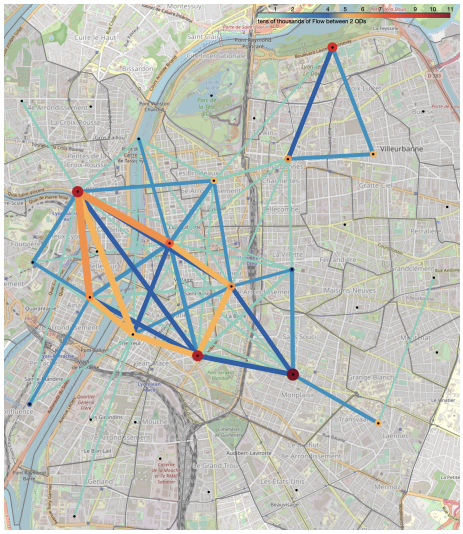

Figure 3: Bike-sharing Flow distribution between the 130 selected ODs in 2019

3.5. Scenarios

It is known that weather conditions may have strong impact on transport demand. In the following, we determine scenarios and combinations of scenarios related to weather conditions to experimentally evaluate their impact on the prediction quality of ST-ED-RMGC, used herein as a representative of state-of-the-art GNN-based approaches for OD traffic prediction with bike sharing data. Each of the considered scenarios will be presented and analyzed below. Among these scenarios, only the dates corresponding to a period from 7am to 9pm have been retained.

3.5.1. Scenarios related to atypical weather

To build our scenarios, we use the weather data detailed in table Table 1. We define *atypical weather values* the highest (or lowest) 3% of the value set of a weather variable. Specifically:

1. A temperature is atypically cold if it belongs to the set of the 3% coldest temperature recorded over the winter 2019-2020.
2. A wind is atypically strong if it belongs to the set of the 3% strongest winds over the winter 2019-2020.
3. A rainfall is atypically strong if among all the dates with actual rainfall, it belongs to the 3% strongest rainfall over the winter 2019-2020.

On a milder scale, we also define as *strong weather values* those in the 8% to 12% highest (or lowest) range of the value set of the considered weather variable (Table 1). Based on these considerations, we then defined scenarios and combination of scenarios for which we have evaluated their prediction accuracy from the ST-ED-RMGC model on Table 2.

Table 1: Weather variables

mark (unit)	meaning	strong values
T (C)	shade temperature	$T < 3$
w $(m.s^{-1})$	hourly wind speed	$w > 8$
r $(mm.h^{-1})$	hourly rainfall	$r > 0$
dr $(mm.h^{-1}$per day)	daily cumulative rainfall intensity per hour	$dr > 5$
dd ($h$per day)	daily cumulative rainfall duration per hour	$dd > 300$
dw $(m.s^{-1}$per day)	daily cumulative wind speed per hour	$dw > 100$

4. EXPERIMENTS

4.1. Dataset

We use multi-source data such as bike-sharing transport data provided by the operator JCDecaux[2], The French Institute of Statistics INSEE provides public socio-economic data[3], and weather data provided by Meteo France[4]. Two years of data (2019-2020) have been used. Since 2020 was quickly

[2]https://www.jcdecaux.fr
[3]`https://www.insee.fr/fr/statistiques`
[4]https://meteofrance.com

marked by the consequences of the COVID-19 pandemic, we are training the algorithm with bike-sharing data from 01/01/2019 to 09/11/2019, using the data between 09/11/2019 and 01/12/2019 for validation purposes, and the data from 01/12/2019 to 15/03/2020 as the testing dataset (on top of which our weather-related scenarios are defined). We selected and hourly aggregated demand data related to the bike sharing stations of 220 IRIS zones around the center of the city of Lyon (including its neighbouring area of Villeurbanne), which we further aggregated into 50 larger zones according to the iterative spatial aggregation procedure described in Sec. 3.4.

Volume of Bike sharing data

We noticed that more than 90% of the transits takes place between 7am and 9pm. In order to obtain ODs with an average of at least 3 exchanges per hour between 7am and 9pm, we decide to keep the ODs with at least 17 000 exchanges over the year. Figure 2 and Figure 3 present the distribution of flows of the 130 selected ODs.

Weather data

The weather data are collected hourly from 2 weather stations in the city of Lyon from January 1st, 2019 to December 31th, 2020. Weather scenarios from Table 2 are defined with respect to the mean value of the considered variable(s) from both stations.

4.2. Experimental results

Table 2 describes the set of considered weather scenarios (1:27) along with the related performance metrics. The testing set is split according to multiple criteria related to the considered weather variable(s) in order to allow for a fair comparison in the presence or not of an atypical or strong weather-related situation. For instance, considering the rain situation, we split the testing set into a subset where we observe extreme rain, *i.e., atypical rain* scenario; a second subset where *strong rain* (but not extreme) is observed and a final subset with no rain at all. The idea is to comparatively analyze accuracy metrics in such splits to determine the impact of the weather situation with respect to the baseline (no weather event). The table also reports the performance results when no atypical or strong weather situation is considered at all, and all time slots are therefore collapsed into one testing dataset (scenario 0). We first assess the impact of rain on the quality of the prediction (scenarios 1:3). We look at the Mean Squared Error (MSE) and Mean Absolute Error (MAE) which are absolute errors linked to the application. That is, they are generally greater when the demand is higher, but this does not necessarily indicate poor prediction quality. The Mean Absolute Percentage Errror (MAPE) is a relative error, so we will use it to compare the performance of the scenarios prediction toward the baseline.

Table 2: Accuracy prediction on test set 01/12 - 15/03.

n° 01/12 - 15/03	Scenario	MSE	MAE	MAPE tested	dates
0	test set	8.85	2.13	0.57	1485
1	Atypical r	9.71	2.37	1.49	4
2	r > 0	9.54	2.18	0.75	159
3	r = 0	8.76	2.13	0.55	1326
4	Atypical w with r=0	7.45	2.03	0.60	42
5	typical w with r = 0	8.81	2.13	0.55	1282
6	w > 8 among r = 0	8.74	2.16	0.57	159
7	w < 8 among r = 0	8.77	2.12	0.55	1167
8	T > 3 among r = 0	9,07	2,18	0,55	1161
9	T < 3 among r = 0	6.59	1.78	0.55	165
10	T < 3	6.63	1.78	0.55	166
11	Atypical T among r=0	8.52	2.02	0.60	31
12	dr = 0	8.83	2.13	0.54	930
13	dr = 0 and dw > 100	8.63	2.10	0.56	165
14	dr = 0 and dw < 100	8.87	2.13	0.54	765
15	dr > 0	8.88	2.14	0.62	555
16	dr > 0 and dw < 100	8.89	2.14	0.62	465
17	dr > 0 and dw > 100	8.85	2.15	0.62	90
18	0 > dr > 1	8.58	2.11	0.58	120
19	1 > dr > 3	9.12	2.17	0.60	255
20	3 > dr > 5	8.55	2.11	0.61	45
21	dr > 5	8.82	2.13	0.70	135
22	1 > dr > 3 and dd > 200	8.33	2.06	0.63	135
23	dd = 0	8.86	2.13	0.54	765
24	dd > 0	8.84	2.14	0.61	720
25	dd > 60	8.97	2.15	0.63	510
26	dd > 200	8.65	2.11	0.66	300
27	dd > 300	8.96	2.15	0.68	135

The results from the table shows that scenarios with rain are more difficult to predict. Those without rain but with strong wind or atypical temperature are also more difficult to predict.

Using scenario 3 as a baseline, we see from scenario 2 that the presence of rain is inversely related to the quality of the prediction with an MAPE 36% higher (*i.e.,* MAPE of 0.75 for scenario 2 versus 0.55 for scenario 3).

Atypical heavy rainfall scenarios have the worst

prediction accuracy, with an MAPE 170% higher than the baseline (scenario 3). In these scenarios, rainfall intensity is measured per hour, and scenarios (18:21) represent a gradation of increasing rainfall intensity over the day. Results found that the more rain that falls during the day, the worse the prediction quality is, with even light rain leading to a 7% increase in MAPE. As rainfall intensity increases, so does the MAPE, with scenario 21 leading to a 30% higher MAPE than the baseline. These findings suggest that the algorithm struggles to adapt to rainfall, even if historical demand shows an anomaly. Figure 4 highlights this phenomenon. The duration of rain is highly correlated with its cumulative intensity. Same conclusion can then be drawn on scenarios (23:27). Rainfall has then a significant impact on prediction accuracy, leading us to examine scenarios without rain. In the absence of rain, strong wind increases the MAPE by 4% (compare scenarios 13 and 14). Similarly, an atypical wind leads to an MAPE increase of 9% compared to the baseline. This suggests that wind has an impact on demand and prediction quality, especially with atypically strong winds in the absence of rain. When rain is present, scenarios (16:17) show no link between wind intensity and prediction accuracy when compared to the baseline scenario 15, indicating that rain is the dominant weather factor. In Figure 4 we have a window comparing the actual demand and the predicted demand on a day when it rains all day long. Compared to the other figures presented, the prediction is globally bad, whatever the rain intensity. It is possible that the behavior of bike users was strongly influenced by the expected rainfall intensity rather than by the actual rainfall intensity received. The comparison between scenario 22, which reflects a rainy day of low intensity, and scenario 19, which reflects all days with a low cumulative intensity, supports this hypothesis. Compared to the reference scenario 3, predicting bike demand in the atypical temperature scenario 11 was more challenging, with a 9% increase in MAPE. No difference was observed in the case of *cold temperatures*. It is possible that there is a temperature threshold at which bike sharing demand is significantly affected, but this could not be studied further in the context of this paper and will be matter of future investigation.

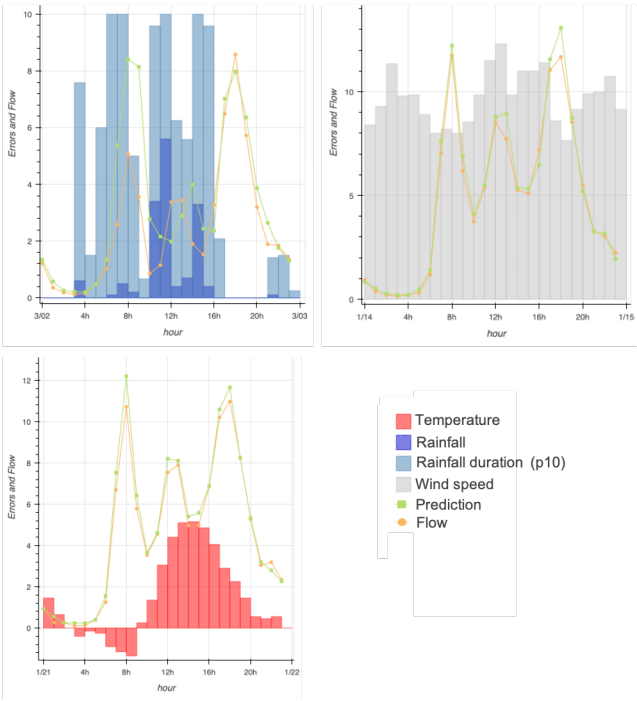

Figure 4: Bike-sharing demand and prediction over time on all ODs and on certain days. Duration is the rainfall duration expressed from 0 (0 min) to 10 (60min).

5. CONCLUSIONS AND FUTURE WORK

Our study has allowed us to highlight forecasting problems related to weather conditions when using a state-of-the-art deep learning framework with bike-sharing demand data. The presence of rain appears as the most challenging factor of error, with a variation of up to a 170% variation compared to a baseline scenario. Other factors such as an atypical temperature, an atypical wind, or a strong wind during the day appear as conditions deteriorating the quality of the prediction. Our study also emphasizes the impact on forecasting accuracy of the timing, the intensity and duration of the weather events.

It is worth to highlight that this study is a preliminary work with multiple limitations. First, each bike station presents a limited capacity and, consequently, bounded demand over time. It has been therefore necessary to aggregate stations together in order to have a statistically viable representation of the demand per ODs. We have chosen to work with

aggregate IRIS zones in order to integrate similarity graphs. However, the spatial aggregation performed groups small areas adjacent to the city center. Some of these areas have a large number of stations that are important transit points. The demand is thus not evenly distributed among the ODs. One direction of future work would be to propose another type of aggregation that would smooth the distribution of flows among ODs.

As another important research direction, we think of integrating contextual information with the help of attention mechanisms and hope for a net improvement of the accuracy. It would also be interesting to study the correlation with other modes (*e.g.,* public transport) as well as the interactions between bike sharing and road traffic to integrate them as inputs and improve the global accuracy. Noticing that the error is high even with the presence of rain during several time-steps, we also consider the possibility of integrating the previous error as an input feature, and thus allow it to correct the shot.

6. REFERENCES

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.

Bruna, J., Zaremba, W., Szlam, A., and Lecun, Y. (2014). "Spectral networks and locally connected networks on graphs." *International Conference on Learning Representations (ICLR2014), CBLS, April 2014*.

Castro-Neto, M., Jeong, Y.-S., Jeong, M.-K., and Han, L. D. (2009). "Online-svr for short-term traffic flow prediction under typical and atypical traffic conditions." *Expert Systems with Applications*, 36(3, Part 2), 6164–6173.

Chai, D., Wang, L., and Yang, Q. (2018). "Bike flow prediction with multi-graph convolutional networks." *CoRR*, abs/1807.10934.

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).

Geng, X., Li, Y., Wang, L., Zhang, L., Yang, Q., Ye, J., and Liu, Y. (2019a). "Spatiotemporal multi-graph convolution network for ride-hailing demand forecasting." *Proceedings of the AAAI Conference on Artificial Intelligence*, 33(01), 3656–3663.

Geng, X., Wu, X., Zhang, L., Yang, Q., Liu, Y., and Ye, J. (2019b). "Multi-modal graph interaction for multi-graph convolution network in urban spatiotemporal forecasting." *CoRR*, abs/1905.11395.

Ke, J., Qin, X., Yang, H., Zheng, Z., Zhu, Z., and Ye, J. (2021). "Predicting origin-destination ride-sourcing demand with a spatio-temporal encoder-decoder residual multi-graph convolutional network." *Transportation Research Part C: Emerging Technologies*, 122, 102858.

Li, Y., Yu, R., Shahabi, C., and Liu, Y. (2017). "Graph convolutional recurrent neural network: Data-driven traffic forecasting." *CoRR*, abs/1707.01926.

Liang, Y., Huang, G., and Zhao, Z. (2022a). "Bike sharing demand prediction based on knowledge sharing across modes: A graph-based deep learning approach, <https://arxiv.org/abs/2203.10961>.

Liang, Y., Huang, G., and Zhao, Z. (2022b). "Cross-mode knowledge adaptation for bike sharing demand prediction using domain-adversarial graph neural networks, <https://arxiv.org/abs/2211.08903>.

Ma, D., Song, X., and Li, P. (2021). "Daily traffic flow forecasting through a contextual convolutional recurrent neural network modeling inter- and intra-day traffic patterns." *IEEE Transactions on Intelligent Transportation Systems*, 22(5), 2627–2636.

Shuman, D. I., Narang, S. K., Frossard, P., Ortega, A., and Vandergheynst, P. (2013). "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains." *IEEE Signal Processing Magazine*, 30(3), 83–98.

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>.

Ye, J., Sun, L., Du, B., Fu, Y., Tong, X., and Xiong, H. (2019). "Co-prediction of multiple transportation demands based on deep spatio-temporal neural network." *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery Data Mining*, KDD '19, New York, NY, USA, Association for Computing Machinery, 305–313, <https://doi.org/10.1145/3292500.3330887>.

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.

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.

Zhang, J., Zheng, Y., and Qi, D. (2016). "Deep spatio-temporal residual networks for citywide crowd flows prediction." *CoRR*, abs/1610.00081.

Zhu, J., Wang, Q., Tao, C., Deng, H., Zhao, L., and Li, H. (2021). "Ast-gcn: Attribute-augmented spatiotemporal graph convolutional network for traffic forecasting." *IEEE Access*, 9, 35973–35983.