A Comparison of Modelling Approaches for the Long-term Estimation of Origin Destination Matrices in Bike Sharing Systems

Ibai Laña*, Ignacio (Iñaki) Olabarrieta* and Javier Del Ser*,†

Abstract— **Micro-mobility services have gained popularity in the last years, becoming a relevant part of the transportation network in a plethora of cities. This has given rise to a fruitful research area, covering from the impact and relationships of these transportation modes with preexisting ones to the different ways for estimating the demand of such services in order to guarantee the quality of service. Within this domain, docked bike sharing systems constitute an interesting surrogate for understanding the mobility of the whole city, as origin-destination matrices can be obtained straightforward from the information available at the docking stations. This work elaborates on the characterization of such origin-destination matrices, providing an essential set of insights on how to estimate their behavior in the long-term. To do so, the main non-mobility features that affect mobility are studied and used to train different machine learning algorithms to produce viable mobility patterns. The case study performed over real data captured by the bike sharing system of Bilbao (Spain) reveals that, by virtue of a properly selected set of features and the adoption of specialized modeling algorithms, reliable long-term estimations of such origin-destination matrices can be effectively achieved.**

I. Introduction

Micro-mobility services like bike or scooter sharing systems have acquired a notable momentum in the last decade, with the outlook of having a central role in the future development of urban areas [1]. Environmental awareness has lead to a detriment of the private vehicle as individual mobility means within cities, and to the rise of new individual transportation means that are more efficient and sustainable. Such new mobility alternatives have been embraced by many city councils around the world, not only promoting the private selection of these transportation modes, but also fostering public services that provide access to networks of shared individual vehicles such as bikes, e-bikes, scooters [2] or even electric cars [3].

The upsurge of these new transport mode options unleashes a number of challenges, including the design of networks of protected paths [4], the estimation of the vehicle relocation needs for balancing the service, or the adequate match between the dimension and the demand of the service in different areas of the city [5], [6]. The latter is a main concern of docked micro-mobility services, namely, those in which vehicles are hired in fixed stations. The location and capacity of such stations must be estimated starting from initial knowledge of transportation needs, demographic and topographic features (e.g. slopes, availability of protected paths, proximity to educational buildings) and transport mode alternatives available in each district. However, the real usage of the deployed services allows improving these estimations and resizing, readjusting and optimizing the availability and balance of vehicles in each station [7], [8].

We herein set our focus on these docked micro-mobility services, in which Bike Sharing Systems (BSS) can be argued to be one of their major representatives. Characterizing the usage of BSS and their interaction with other city services [9] is a research area that has garnered the interest of many scholars [10], [11]. The analysis of Origin-Destination (OD) pairs among different stations or districts has been so far addressed mainly from a predictive modeling perspective, trying to estimate destinations given an origin [12], [13], or to estimate the future behavior of the system, usually in the short term [14], [15], [16]. From the managerial viewpoint, the long-term characterization of the demand in each station can be thought to be more actionable in practice, issuing availability (or demand) predictions that reach the 24 hour horizon [17]. This behavior estimation allows managers to proactively take actions that improve the availability of the service at each station.

This paper delves into this last issue, analyzing the main aspects that characterize the BSS-based mobility in the long term. Specifically, we propose and compare methods to obtain origin-destination matrices for any temporal interval in the future. The long-term estimation of mobility related time series is always a challenging topic [18], as the performance of typical modeling approaches degrades as the prediction horizon increases. When dealing with long prediction horizons, the task usually needs to be reduced to a characterization of patterns that can be used as future estimations [19]. In this study, essential pattern characterization features will be considered, and used to train different OD matrix long-term estimation models over real-world BSS usage data collected over the city of Bilbao (Spain). As a result, the most suitable modeling approach will be identified subject to the same set of input characteristics. The contributions of this work can be summarized as follows:

- We examine how to characterize the mobility patterns of a BSS, proposing a set of essential features that can define them in the long term.
- We explore different algorithmic approaches for such a characterization, from naïve statistical methods to advanced models that take into account spatio-temporal relationships between docking stations. A small selection of algorithms is chosen within each approach in order to

*Ibai Laña, Ignacio Olabarrieta and Javier Del Ser are with TECNALIA, Basque Research and Technology Alliance (BRTA), 48160 Derio, Bizkaia, Spain. Contact email: ibai.lana@tecnalia.com

†Javier Del Ser is also with the University of the Basque Country (UPV/EHU), 48013 Bilbao, Bizkaia, Spain.

contain the length of the paper.

- We evaluate and compare such methods to each other over a complete case study comprising real-world BSS data, towards analyzing the feasibility of the long-term OD estimation task.

The rest of this paper is structured as follows: Section II details the dataset, features and modeling approaches used in our study, whereas results are presented and discussed in Section III. Finally, Section IV concludes the paper with some remarks and a prospect of future research lines.

II. Materials and Methods

We begin by this section, which delves into the specifics of the dataset in use (Subsection II-A), the information that is used to characterize the mobility patterns in the long term (Subsection II-B) and the experimental setup designed to evaluate different estimation approaches departing from such information, together with the details of the approaches themselves (Subsection II-C).

A. Data and preprocessing

BSS data for experimentation have been obtained under the URBANITE Horizon 2020 project[1]. Trips among e-bike docking stations available over the city of Bilbao (Spain) have been recorded since the beginning of the current service in October 2018, storing duration, identifiers (ids) of origin and destination and air distance (straight line between both stations). The service started with 31 docking stations but has been continuously increasing its coverage to currently 41 stations, while some of them have been relocated. Figure 1 shows the district division of the city and the current location of the docking stations. The city has a central business district in a lower area (district 6), whereas some of the residential districts (2, 3, 4 and part of 5 and 7) are located in hills. Although the BSS provides e-bikes that help with slopes, it can be observed in the data that uppermost stations rarely receive trips that do not come from central districts. This lack of connectivity among certain OD pairs calls for a deeper analysis.

The number and location of stations in the period covered by this research study varied significantly. Together with the connectivity issues described above, this variability motivated an aggregation of data considering the districts of Bilbao and the hour of the day as the main spatial and temporal divisions, reformulating the case study as a *district-wise hourly* origin-destination matrix estimation. This aggregation allows maintaining a spatial coherence: in the long term, connectivity between districts is expected to remain stable. Thus, all trips from all stations of a district towards all stations of another have been aggregated, rendering a 8×8 OD matrix. After the initial aggregation, which spanned a total of around 1,400,000 trips during years 2019, 2020 and 2021, several preprocessing filters were applied to clean data. To begin with, the service is limited to 60 minutes, after which the bike must be returned. Leaving a safety margin, all trips with duration greater than 90 minutes were removed. Besides, in some cases, bikes undergo some kind of problem and are returned immediately. Thus, trips lasting less than 2 minutes and with the same origin and destination docking stations were also removed. Lastly, and as the dataset spans mostly dates within a COVID-19 pandemic period, the operating hours of the BSS have varied accordingly, from 24h service to different levels of service up to the current one, in which the bikes are not available at night. For this reason, and in order to ascertain that *0 trips* between two stations mean that no trips were produced and not that the service was actually down, a final filter was applied, removing from the dataset all hours for which the total sum of trips in the whole city was 0. After this preprocessing stage, the dataset is reduced to around 1,220,000 trips that occur during approximately 7,000 hours, which is around $1/4$ of the total hours of the dataset period, due to the aforementioned confinements and the reduced operation regime.

[1] URBANITE project website, `https://www.urbanite-project.eu`, accessed on March 14th, 2022.

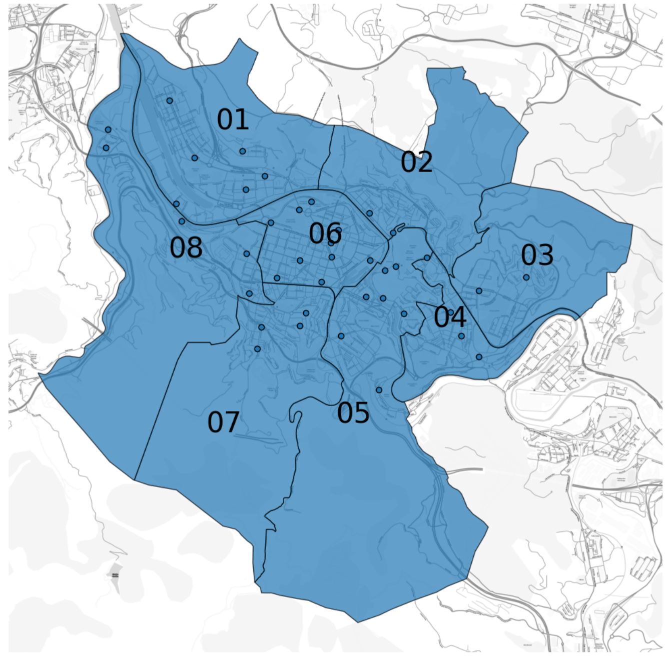

Fig. 1. Districts of the city of Bilbao (Spain), and location of the e-bike docking stations.

B. Calendar and Meteorological information

Our study is concerned with the characterization of mobility in origin destination pairs without a particular predictive horizon. Consequently, any model devised for this purpose must not include actual trip data, as these could not be available for future queries made to the model. Thereby, input data are used only to train data-based models in conjunction with calendar information, which can be obtained for any time in the future, and also with meteorological information, which is expected to be crucial for the characterization (e.g., rain can have a high impact in the usage of bikes). Other data sources could be considered with this same purpose upon their availability. For instance, sports or cultural events, demonstrations and parades, or other transportation modes timetables can have an effect on the way in which BSSs are used. However, and due to the data availability for our particular case study, for each hour of the available dataset, the following input data are used:

- Calendar data, which is realized by the hour of the day, the day of the week and two fields to represent whether the day is a public holiday or an academic holiday. These features permit to characterize the usage per hour and day of the week, but also consider aspects like the use of the bikes for leisure or as a transportation means to reach schools and universities.
- Meteorological data: hourly temperature and precipitation have been considered to compose the dataset. These data have been encoded into categorical features, leaving the temperature and rain as variables with four possible values: *extremely cold*, *cold*, *warm* and *hot* (for temperature); and *dry*, *light rain*, *rain* and *heavy rain* (for precipitation).

The assumption is that if heavy rain or very cold weather takes place, the amount of bike trips plummets, disregarding the occurrence of precipitation or the nuances provided by the Celsius degrees. Figure 2 shows that in fact, the amount of bike trips is drastically reduces when heavy rain occurs.

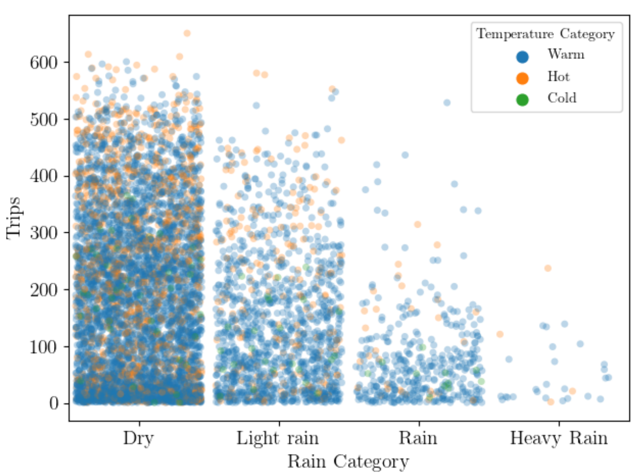

Fig. 2. Effect of precipitation categories and temperature in the bike usage.

With these data and the number of trips for a certain OD pair at a certain hour, three learning model are trained, aimed at predicting the future behavior of the whole origin destination matrix for the city. The prediction horizon is conceptually unlimited, but in practice it is limited to the availability of meteorological information, which will ultimately depend on weather forecasts. Thus, this prediction could be made for hours or even days into the future.

C. Experimental Framework and Estimation Approaches

An experimental framework has been designed in order to evaluate different modelling choices capable of providing long-term OD matrix estimations. In time series analysis, the long-term characterization of segments is a complex task that cannot be achieved under the same levels of accuracy as short-term forecasting. The latter relies mainly on previous observations of the time series, hence an increase of the predictive horizon usually ends up reducing the quality of predictions [18].

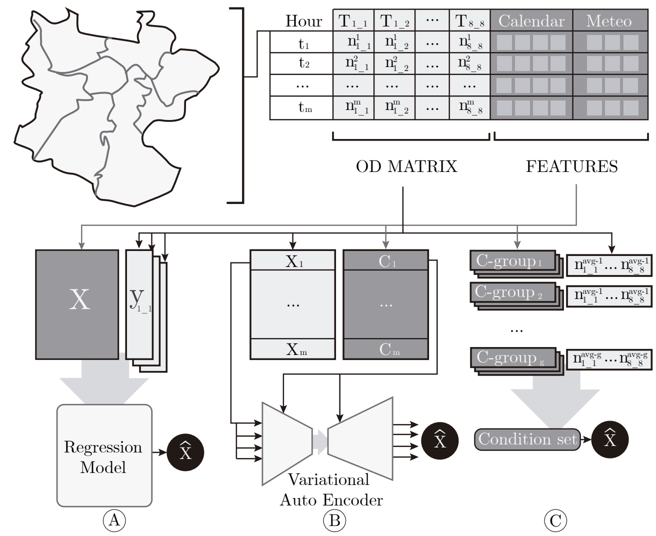

Fig. 3. Block diagram showing the different approaches and data flow used for OD estimation.

As a result, alternative approaches that allow feeding the model with variables that can be known for any time in the future need to be explored. We leave aside clustering approaches of training data [19] to concentrate on regression models, which can be trained to correlate the value of a target variable (i.e., bike trips estimated for an OD pair) with the set of input characteristics. It should be noted that since their value must be known for any future instant of time, in most cases input characteristics are coarsely defined and chosen to be discrete (weather and calendar data as exposed above support this statement). On the other hand, some learning algorithms such as autoencoders and generative adversarial networks allow *creating* or synthesizing output patterns that are consistent with the ones that were used to train them. All these notions are used to build three different approaches that are considered in this work, and whose general operation is illustrated in Figure 3:

Approach A: Supervised learning regression methods

A classical supervised machine learning framework is proposed as the first approach to the problem at hand. Under this framework, input variables $\mathbf{X}$ are conformed by the calendar and meteorological features, whereas the output variable $\mathbf{y}$ is a series defined by the trips registered for a particular OD pair at each time step, as described in Figure 3 A. This approximation requires training as many models as OD pairs (each OD pair produces a series of observations that acts as $\mathbf{y}$), or including the OD pair as a feature of a general model that uses this inserted feature to discriminate among OD pair. This first approach was selected as it is more suitable to generalize to other contexts where the number of considered areas (*districts*) can differ.

Based on this regression approach, different learning algorithms have been considered, taking into account that the limited variability of the input data will burden the performance of all of them. However, the different ways in which each algorithm represents the knowledge is expected to be decisive for their efficiency. The following algorithms have been used:

- Support Vector Regression (SVM, [20]), which extends the concept of maximum-margin hyperplane to regression problems, presenting a great generalization capability.
- Random Forest regressor (RF, [21]), namely, a bagging ensemble of decision trees that performs robustly in most modeling contexts.
- Catboost regressor (CBR, [22], [23]), i.e., a gradient boosting ensemble that supports categorical features, making it particularly suitable for the task at hand.

For the sake of fairness in the performed comparisons, hyper-parameters of all learners were tuned exhaustively by using off-line grid-search and cross-validation processes, so that results elicited by the best performing configuration of each model are reported.

Approach B: Conditional Variational Auto Encoder (CVAE)

The second approach evaluated in this work is an extension of the standard family of Variational Auto Encoders (VAEs), which are directed graphical models capable of generating new data instances by learning the parameters of the conditional distribution $P_{\mathbf{X}|\mathbf{Z}}(\mathbf{x}|\mathbf{z})$, where $\mathbf{X}$ is the input data and $\mathbf{z}$ is an unobserved random variable with prior distribution $P_{\mathbf{Z}}(\mathbf{z})$. A VAE [24] has two structural parts: 1) an encoder that learns how to map (*encode*) input $\mathbf{X}$ to the hidden embedding $\mathbf{Z}$ by learning a distribution $Q_{\boldsymbol{\theta}}(\mathbf{z}|\mathbf{x})$ defined by parameters $\boldsymbol{\theta}$; and 2) a decoder that learns a parametric approximation $P_{\boldsymbol{\phi}}(\mathbf{x}|\mathbf{z})$ of $P_{\mathbf{X}|\mathbf{Z}}(\mathbf{x}|\mathbf{z})$, allowing to reconstruct the input from the embedding space. To this end, a compound loss function is often used to account for the reconstruction error $\mathcal{L}_{rec}(\boldsymbol{\theta},\boldsymbol{\phi})$ from the latent distribution, as well as the divergence $\mathcal{L}_{div}(\boldsymbol{\theta})$ between the distribution $Q_{\boldsymbol{\theta}}(\mathbf{z}|\mathbf{x})$ learned by the encoder and the prior distribution $P_{\mathbf{Z}}(\mathbf{z})$ assumed for $\mathbf{Z}$ (often Gaussian). Once trained, new samples can be synthesized by feeding the encoder with realizations of $\mathbf{Z}$ drawn as per its assumed prior $P_{\mathbf{Z}}(\mathbf{z})$. Figure 4.a depicts the architecture of a typical VAE.

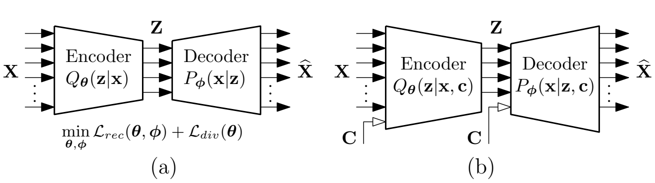

Fig. 4. Block diagram showing (a) a standard VAE; (b) a CVAE.

Unlike standard VAE, in CVAE the generative problem is reformulated to control the data generated by the model by means of an external supervisory signal $\mathbf{C}$ (*condition*) [25]. This signal can be the class of the image to be generated by the decoder module. In our case, $\mathbf{C}$ represents the conditions under which the origin-destination matrix is to be estimated: departing from a certain value of this signal $\mathbf{c}$, an estimated matrix is produced by drawing a sample from a conditioned prior distribution $P_{\mathbf{Z}|\mathbf{C}}(\mathbf{z}|\mathbf{c})$, whereas the output $\mathbf{x}$ results from sampling the distribution $P_{\mathbf{X}|\mathbf{C},\mathbf{Z}}(\mathbf{x}|\mathbf{c},\mathbf{z})$ learned by the CVAE decoder. The loss function is similar to that of VAE, the difference being that the reconstruction and divergence loss terms are now conditioned on $\mathbf{C}$. In other words, the CVAE, trained to reconstruct an input origin-destination matrix under different conditions, can easily produce new estimated matrices for any given value of $\mathbf{c}$, which encodes the weather and calendar features for which the query is done. This modified architecture is shown in Figure 4.b.

Approach C: Naïve baseline approach

Lastly, a naive approach is proposed as baseline to compare the performance of the other approaches. Since the selected weather and calendar features are discrete, there is a limited set of feature combinations that define all possible situations for that particular set of features. Thus, for each combination, the trips of all available OD pairs on hours that satisfy that combination are averaged. In prediction stage, a combination (set of calendar and meteorological conditions) is defined, so that the averages for each OD pair are used as the estimation produced for the query at hand. This approach takes no training time, and should establish the minimum achievable performance (*baseline*) for the problem under consideration.

III. Experiments and Results

Once data have been cleansed, initial three years of data were split into train, validation and test subsets, the latter including the late December days of 2020 and January and February 2021. This test subset contains samples of all kinds of days, with *cold*, diverse amounts of *rain*, bank holidays and academic holidays, namely, all circumstances that may affect the normal bike usage. However, it is relevant to highlight that training data include normal operation months of 2019, but also the most challenging months of the pandemic with very limited service. In terms of day similarity, samples within the test subset cannot be considered close to those during the normal operation in 2019, nor are they similar to those held during the pandemic scenario, when different levels of COVID-19 measures were in place and impacted on the service. This situation can burden the performance of the proposed approaches: although they can be compared to each other as they consider the same training and test data, performance metrics could be expected to improve if non-exceptional usage patterns are again in place.

Bearing this in mind, the performance of the models is assessed in terms of the Root Mean Squared Error (RMSE):

$$RMSE(\mathbf{y},\widehat{y}) = \sqrt{\frac{1}{N}\sum_{t\in\mathcal{T}_{test}}(y_t-\widehat{y}_t)^2}, \qquad (1)$$

where $RMSE(\mathbf{y}, \widehat{y})$ is the RMSE computed for a particular OD pair belonging to the test partition $\mathcal{T}_{test}$; y_t denotes the number of observed trips at test time t; $\widehat{y}_t$ the estimated number of trips at that same time; and N is the total number of considered query samples. This metric is obtained for each OD pair and model, resulting in a matrix of 64×5 RMSE values that are represented in Figure 5.

Focusing on this figure, RMSE values are plotted for every OD pair, labeled as O_D (with O and D integers denoting the index of the district). As it can be observed from the performances of the 5 approaches in the 64 OD pairs, in most situations all models behave very similarly, with slight error differences that do not justify solidly the use of advanced modelling approaches. However, for those OD pairs where mobility is higher, specially those with origin or destination in zone 6 (central business district), it is particularly noticeable that specialized modeling approaches provide better performances. Particularly CVAE (approach B) seems to leverage spatio-temporal relationships that are encoded in the network. It is the only approach that considers all trips at once during the training phase, so that an estimation is provided for all OD pairs together. In a closed system like a BSS, trips from one location to another increase the availability of bikes in the destination, augmenting the possibility of trips from it. The rest of approaches are unable to harness these interactions, what could explain the better performance of the CVAE approach.

Among the rest, CBR performs in general slightly better than the others, specifically in the most challenging scenarios with many trips. As it was anticipated previously, this superiority can be attributed to the specialization of CBR to deal with categorical features. Another interesting finding that emerges from these results is the extreme similarity of RF and the naive average (approach C). In a highly discrete scenario as the one presented here, it seems reasonable that the decision trees ensemble ends up averaging values and providing averages, what explains the very similar performances. However, that slight error difference is relevant, for some specific cases the naive averaging mechanism does not have enough data to provide an average: the particular combination of features has not been observed in the training phase. A machine learning method as RF is able to provide at least an interpolated value for those areas, obtaining a final better result.

However, this representation of the error by observing the individual RMSE values for each OD pair does not portray all the nuances that can be found in Figure 6, where estimations for all approaches for a particular winter day with moderate rain are shown. This is one of the cases where naive average approach does not have enough previous information, producing a non valid estimation, while RF and CBR behave very similarly and CVAE produces a line closer to the real one. In the case of SVM, the produced estimation follows the underlying scheme of the travel behavior, but is not able to capture all variability. This is reflected in Figure 5, where SVM usually performs worse than other approaches, but in some cases this smooth result yields the best individual estimation, as it can be observed in Figure 6 at hours 9:00 and 11:00.

When it comes to the connectivity between OD pairs, Figure 7 depicts the normalized error of each method for a particular day and hour, obtaining very similar representations for CBR, RF and AVG, while SVM and CVAE render a different behavior. It is interesting to observe that CVAE outperforms the rest of approaches specially when the amount of trips is large, while for connections with a low amount of trips (districts 3 and 4 on the right of the map) are better represented with the rest of the methods. CVAE provides an advantage in those cases where trips in one OD pair affect other OD pairs. This circumstance is more likely to happen with higher amounts of trips, which can lead to a shortage of bikes in certain stations, or a better availability in others. These hidden relationships are certainly part of the generally best performance levels attained by CVAE. For consistently lower OD pair connectivity over time, the rest of methods provide a more straightforward approach to the problem.

IV. Conclusions and Future Research Lines

In this work, we have examined several machine learning approaches to the long-term estimation of the number of trips held among origin and destination pairs in a bike sharing system. The estimation is done with no prediction horizon beyond the one defined by the limit of the weather forecasts. Several conclusions can be drawn from experiments performed over real data from Bilbao (Spain). The most basic approach based on averaging data from days with the same features occurs to suffice for providing accurate estimations. The modeling capability of more elaborated machine learning approaches increases the performance and the ability to produce estimations for types of days never observed before.

The proposed methods have shown that long-term mobility patterns based on calendar and weather features hold a certain stability in time, even under considerable disruptions in the BSS service. These disruptions, however, may impact on this performance, as the accumulated usage is particularly low in some of the zones, making it easier to estimate their behavior. If the limited service and changing schedules discourage users from taking a bike in certain districts, it is relative easy to estimate a null usage of the service in those areas. Nonetheless, in the districts where the amount of trips is significant – particularly in zone 6 (central district) and the inner connections of stations within district 1 – the spatio-temporal relationships modeled by CVAE allows obtaining better estimations than with the rest of the models.

Modeling a continuous variable (number of OD trips) with discrete features is in principle limited on its own as per their alphabet and cardinality. Therefore, achieving reliable estimations in such a limited setup can be certainly challenging. However, when dealing with periodic data patterns, estimations in very long term intervals can be produced under this setting, which can be very useful for service managers. In this manuscript, discrete features

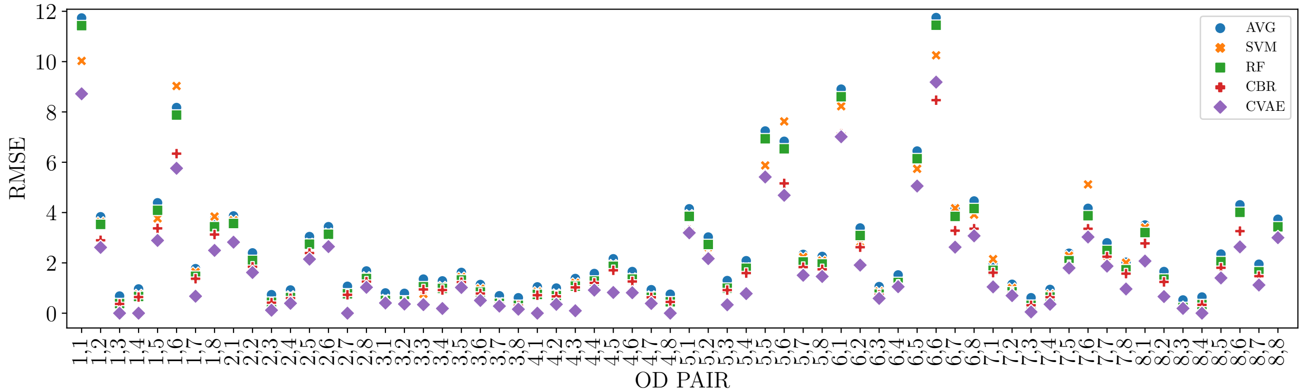

Fig. 5. RMSE results obtained by each model over the 64 OD pairs.

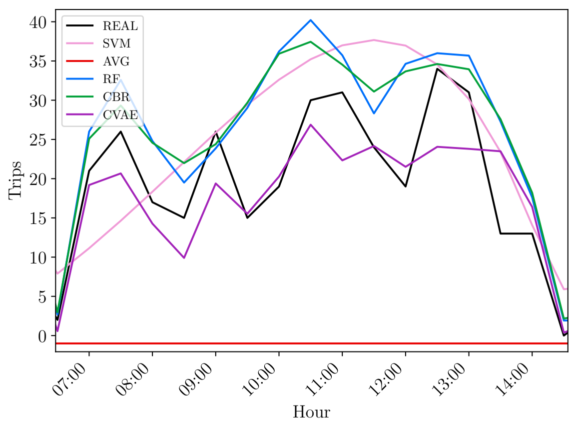

Fig. 6. Estimations for a winter day in the 6_7 OD pair.

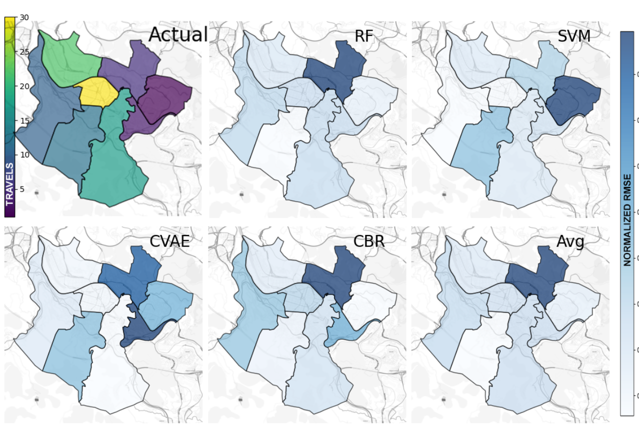

Fig. 7. Trips from zone 6 (central) on the midday of a rainy January day. Top left map represents the real amount of trips, whereas the rest of nested maps represent a normalized error of each method when estimating trips for each OD pair.

(calendar, weather) were restricted to those available in the scenario under study. However, we envision that incorporating more features such as events that can force docking stations to be provisionally closed (for instance, nearby stations are closed during football matches or street fairs) can help improve the performance of the developed models. Aspects related to the topography and the availability of cycling paths could be added as well as priors of the relationships between docking stations, potentially improving estimations for modeling choices that are sensitive to such interactions. Lastly, factors that are considered in short-term predictions, like the spillage to other stations or transport modes, could also be analyzed in the long term, towards achieve even better estimations.

Acknowledgments

This work has received funding support from the Basque Government (ELKARTEK program, ref. KK-2020/00049, and the Consolidated Research Group MATHMODE, ref. IT1294-19), and the European Union's Horizon 2020 programme URBANITE: Supporting the decision-making in urban transformation with the use of disruptive technologies (grant agreement 870338).

References

[1] R. L. Abduljabbar, S. Liyanage, and H. Dia, "The role of micro-mobility in shaping sustainable cities: A systematic literature review," *Transportation Research Part D: transport and environment*, vol. 92, p. 102734, 2021.

[2] D. J. Reck, H. Martin, and K. W. Axhausen, "Mode choice, substitution patterns and environmental impacts of shared and personal micro-mobility," *Transportation Research Part D: Transport and Environment*, vol. 102, p. 103134, 2022.

[3] R. Mounce and J. D. Nelson, "On the potential for one-way electric vehicle car-sharing in future mobility systems," *Transportation Research Part A: Policy and Practice*, vol. 120, pp. 17–30, 2019.

[4] S. P. Wall, D. C. Lee, S. G. Frangos, M. Sethi, J. H. Heyer, P. Ayoung-Chee, and C. J. DiMaggio, "The effect of sharrows, painted bicycle lanes and physically protected paths on the severity of bicycle injuries caused by motor vehicles," *Safety*, vol. 2, no. 4, p. 26, 2016.

[5] D. Chemla, F. Meunier, T. Pradeau, R. W. Calvo, and H. Yahiaoui, "Self-service bike sharing systems: simulation, repositioning, pricing," 2013.

[6] N. Jian, D. Freund, H. M. Wiberg, and S. G. Henderson, "Simulation optimization for a large-scale bike-sharing system," in *Winter Simulation Conference (WSC)*, 2016, pp. 602–613.

[7] J. Liu, Q. Li, M. Qu, W. Chen, J. Yang, H. Xiong, H. Zhong, and Y. Fu, "Station site optimization in bike sharing systems," in *IEEE International Conference on Data Mining*, 2015, pp. 883–888.

[8] F. Chiariotti, C. Pielli, A. Zanella, and M. Zorzi, "A dynamic approach to rebalancing bike-sharing systems," *Sensors*, vol. 18, no. 2, p. 512, 2018.

[9] W. Tu and H. Liu, "Transfer probability prediction for traffic flow with bike sharing data: A deep learning approach," in *Science and Information Conference*, 2019, pp. 71–85.
[10] I. B. I. Purnama, N. Bergmann, R. Jurdak, and K. Zhao, "Characterising and predicting urban mobility dynamics by mining bike sharing system data," in *IEEE 12th International Conference on Ubiquitous Intelligence and Computing*, 2015, pp. 159–167.
[11] V. Lucas and A. R. Andrade, "Predicting hourly origin–destination demand in bike sharing systems using hurdle models: Lisbon case study," *Case Studies on Transport Policy*, vol. 9, no. 4, pp. 1836–1848, 2021.
[12] P. Dai, C. Song, H. Lin, P. Jia, and Z. Xu, "Cluster-based destination prediction in bike sharing system," in *Artificial Intelligence and Cloud Computing Conference*, 2018, pp. 1–8.
[13] J. Jiang, F. Lin, J. Fan, H. Lv, and J. Wu, "A destination prediction network based on spatiotemporal data for bike-sharing," *Complexity*, vol. 2019, Article ID 7643905, 14 pages, 2019.
[14] Y. Li, Y. Zheng, H. Zhang, and L. Chen, "Traffic prediction in a bike-sharing system," in *23rd SIGSPATIAL International Conference on Advances in Geographic Information Systems*, 2015, pp. 1–10.
[15] S. Sohrabi, R. Paleti, L. Balan, and M. Cetin, "Real-time prediction of public bike sharing system demand using generalized extreme value count model," *Transportation Research Part A: Policy and Practice*, vol. 133, pp. 325–336, 2020.
[16] L. Lin, Z. He, and S. Peeta, "Predicting station-level hourly demand in a large-scale bike-sharing network: A graph convolutional neural network approach," *Transportation Research Part C: Emerging Technologies*, vol. 97, pp. 258–276, 2018.
[17] D. Cenni, E. Collini, P. Nesi, G. Pantaleo, and I. Paoli, "Long-term prediction of bikes availability on bike-sharing stations," in *27th International DMS Conference on Visualization and Visual Languages*, 2021.
[18] I. Laña, J. Del Ser, M. Velez, and E. I. Vlahogianni, "Road traffic forecasting: recent advances and new challenges," *IEEE Intelligent Transportation Systems Magazine*, vol. 10, no. 2, pp. 93–109, 2018.
[19] I. Laña, J. L. Lobo, E. Capecci, J. Del Ser, and N. Kasabov, "Adaptive long-term traffic state estimation with evolving spiking neural networks," *Transportation Research Part C: Emerging Technologies*, vol. 101, pp. 126–144, 2019.
[20] M. Awad and R. Khanna, "Support vector regression," in *Efficient Learning Machines*, 2015, pp. 67–80.
[21] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
[22] A. V. Dorogush, V. Ershov, and A. Gulin, "Catboost: gradient boosting with categorical features support," *arXiv preprint arXiv:1810.11363*, 2018.
[23] S. González, S. García, J. Del Ser, L. Rokach, and F. Herrera, "A practical tutorial on bagging and boosting based ensembles for machine learning: Algorithms, software tools, performance study, practical perspectives and opportunities," *Information Fusion*, vol. 64, pp. 205–237, 2020.
[24] C. Doersch, "Tutorial on variational autoencoders," *arXiv preprint arXiv:1606.05908*, 2016.
[25] K. Sohn, H. Lee, and X. Yan, "Learning structured output representation using deep conditional generative models," *Advances in Neural Information Processing Systems*, vol. 28, 2015.