Analysis of GPS-based Vehicle Mobility Data towards the Electrification of Transportation in Qatar

Usman Zafar
Qatar Environment and Energy Research Institute, Hamad Bin Khalifa Uni.
Doha, Qatar
uzafar@hbku.edu.qa

I. Safak Bayram
Dept. of Electronic and Electrical Eng. University of Strathclyde Glasgow, UK
safak.bayram@strath.ac.uk

Sertac Bayhan
Qatar Environment and Energy Research Institute Hamad Bin Khalifa Uni.
Doha, Qatar
sbayhan@hbku.edu.qa

Raka Jovanovic
Qatar Environment and Energy Research Institute Hamad Bin Khalifa Uni.
Doha, Qatar
rjovanovic@hbku.edu.qa

***Abstract*—Vehicle mobility analysis is a critical input to support mass adoption of electric vehicles (EVs), design required charging infrastructures, and forecast the impacts and risks of this transition.**

***Index Terms*—electric vehicles, GPS data, mobility, data analytics**

I. Introduction

The pressing need to decarbonise the transportation sector require the phasing out of petrol and diesel vehicles [1]. Electric vehicles (EVs) are the most promising technology option to offer low-carbon ground transportation [2]. To support mainstream EV adoption, most members of the G7 countries including the UK, the USA, and the EU, have introduced bold, time-bound, and front-loaded plans in the form of subsidies, tax breaks, and ban on the sale of petrol cars [3]. Despite significant improvements in the battery technology [4], the adoption rates are hindered by range anxiety, lack of sufficient charging infrastructure, and relatively high procurement cost.

Transition to EVs in the Gulf Cooperation Council (GCC) regionis particularly challenging due to highly subsidised petrol prices and extremely hot desert climates which degrades EV batteries. Ambient temperatures in summer could easily exceed 45 °C between May - October period and high demand for air conditioning reduces all electric ranges [5]. Therefore, EV charging infrastructure planning is a critical stage to meet the mobility needs of the drivers and to enable seamless transition towards electrified transportation.

Accurate estimation of EV energy consumption and recharging needs is a complex task due to the impacts of multiple non linear parameters such as ambient temperature, road type, traffic congestion, trip distance and drivers' aggressiveness [6], [7]. Field studies reveal that energy consumption rage (Wh/km) is often higher than the manufacturer's specification [6]. In some cases, national travel surveys are used as a basis to simulate EV mobility patterns and recharging needs [8], [9] by developing probabilistic demand analysis. Although such models provide useful insights, they fail to capture the impacts of aforementioned factors and rare-events (e.g. local demand peaks) affecting power grids. In addition, national travel surveys exist in a handful of countries (EU, UK, USA, etc.) and do not exist in the GCC region. Therefore, GPS traces are critical inputs to examine the temporal alignment of EVs electricity and mobility demands [10].

Due to lack of EVs in the GCC region, GPS traces collected from petrol cars could be used as inputs to EV mobility simulators [11] to examine the performance of EVs and assess the battery packs needed to complete real-world trips made by petrol cars. To that end, contributions of this paper can be listed as follows. First, we collected high resolution GPS-based mobility data from seven cars (six petrol cars one EV) using telematics devices for one year in Qatar. Our datasets include critical information such as speed, location, acceleration, idle period. Second, the gathered data is processed using machine learning based clustering algorithms to reveal location analysis to analyze daily activities, trip patterns, and impacts of weather on fuel efficiency. Third, case studies are presented to assess the suitability of existing EV models in Qatar and provide evidence-driven recommendations on possible EV charging options. To the best of author's knowledge, this is the first study conducted in the GCC region.

II. Literature Review

Over the last years, there has been a growing body of literature on the analysis of electric vehicle mobility patterns. The first group of papers present data analysis and collection from EVs and petrol cars using data loggers and telematics devices [7], [12], [13]. In [12], mobility traces of 40 different EVs were collected and analysed. The primary focus of the study is to understand spatio-temporal recharging habits and it is found that most EV drivers use their cars in a very risk-averse way due to range anxiety. This finding aligns with our analyses presented in Section IV that the daily milage of EV drivers is significantly lower than that of petrol car milage due to range anxiety. In [13], GPS mobility dataset was created from 982 drivers of conventional vehicles in Italy

This is a peer-reviewed, accepted author manuscript of the following article: Bayram, I. S., Zafar, U., Bayhan, S., & Jovanovic, R. (2022). Analysis of GPS-based high resolution vehicle mobility data towards the electrification of transportation in Qatar. Paper presented at 48th Annual Conference of the Industrial Electronics Society.

for over 2 years. The primary goal of the study is to develop a methodology to identify car driver segments to show segment-specific usability and impacts on the power grid. In [?], energy consumption of a Nissan Leaf is measured for four years in the UK. It was found that ambient temperatures significantly reduce the driving range, and short trips below 16 km consume nearly 10% more energy than longer trips.

The second group of studies examine EV mobility patterns using national travel surveys [8]–[10]. In [8], the travel surveys are used to create EV mobility patterns and corresponding charging loads. A similar probabilistic approach is developed in [9] using the travel survey of the UK and corresponding charge demand is estimated using Monte Carlo simulation techniques. In [10], an EV mobility analysis is carried by combining survey data with mobile phone activity. in California.

A number of studies investigate the factors energy consumption of EVs in real-world driving conditions [6], [14]. Ref. [6] shows that energy consumption rate of EVs is highly non-linear and depends on weather and driving patterns. It is reported that the driving ranges of EVs can reduce by 25% during non-ideal weather conditions. In [14], a regression analysis is carried out to analyse power and energy demand agains speed and acceleration of eight popular EV models. The results show that there is almost a linear relationship between speed and power demand.

III. GPS-based Mobility Dataset

A. Vehicle Types

The data was collected from seven volunteers who reside in Doha, Qatar. City of Doha is a metropolitan city with more than two million inhabitants and its transportation practices are shaped by the hot and arid harsh desert climate [5]. Temperatures in the summer can reach up to 50 °C with humidity levels reaching up to 70 percent. The environment is made bearable by extensive use of indoor and outdoor air-conditioning. As a consequence, most residents prefer to use vehicles with high engine volumes (SUVs with 3.0+ engines) and shift towards electric vehicles is particularly challenging to provide the same level comfort to drivers [15]. In Table I, details of the cars are presented. It can be seen that most cars are SUVs, while there is an actual EV used in this study. Moreover, driver types further makes an impact. For instance, cars operated by private drivers are quite common in large families and their mobility pattern is different the rest as the private drive could drop off one family member to workplace and then drive another family member to a shopping mall. Similar affect could be seen in shared drivers as a partners sharing the car are likely to have significantly different behaviour than solo drivers. Due to limited volunteer pool, only seven cars participated. Nevertheless, since this is the first study if its kind, the presented results provide important insights of mobility patterns in the region.

TABLE I
Details of the cars used in the study.

ID	Driver Gender	Make	Model	Type	Driver Type
1	Male	Dodge	Durango	SUV	Shared
2	Female	Toyota	Rav 4	SUV	Solo
3	Male	Lexus	RC350	Sedan	Solo
4	Male	Kia	Optima	Sedan	Solo
5	Male	Volkswagen	eGolf	Sedan (EV)	Solo
6	Male	Nissan	Patrol	SUV	Private Driver
7	Male	Honda	CRV	SUV	Solo

B. Data Acquisition

High resolution GPS data is collected from aforementioned seven vehicles using Teltonika Fm3001 on-board diagnostic (OBD) device [16]. Data presented in this paper covers June 2021 to February 2022 and delivers data every ten-seconds to an on-premise server and averaged at every five minutes for better representation. The telematic device is capable of collecting and delivering variety of information such as speed, acceleration, GPS location, idling detection, trip length, and the entire list of attributes can be found on the products datasheet [16]. Further information such as vehicle fuel level was possible to collect from the first volunteer as this specific car allows the OBD device to read this information. Therefore, the data collected from this car is further investigated and presented in the next section.

The schematic overview of the data acquisition system is presented in Fig. 1. The telematics data is transmitted every 10 seconds. Teltonika devices work on a specific TCP based protocol called Codec-8. This protocol is defined in detail in the OBD device's documentation [16]. A software server for accepting connections from teltonika devices and ingesting them in a time-series database was designed. More specifically, Influxdb [17] is used as time-series database. Influxdb is able to handle 100,000 writes per second on a regular 2-core machine. This amount of write speed is sufficient for acquiring data from all telematic devices. All database instances and code run on docker containers for ease of use and flexibility. The telematics is inserted into the time-series database in binary format. which is not yet parsed by the server. We write scripts to extract data from the binary format using the Codec-8 and insert into a spatial databases with the latitude-longitude pairs of the coordinates defining the geometry of each point.

C. Data Processing

As the volume of the data gathered from telematic devices are large, the following data processing procedures are carried out. A trip is identified by difference in duration between consecutive timestamps. If this difference is less than 10 seconds, then we deduce that the car is moving. If this difference is more than 10 seconds, it is assumed that the car is parked and the engine is off. These settings are configured within the telematics devices to optimise data transfer. Moreover, to improve the accuracy of data flow, it was made sure that the spatial coordinates are within the country boundary for each car. When the GPS device loses its 3G signal, the coordinates

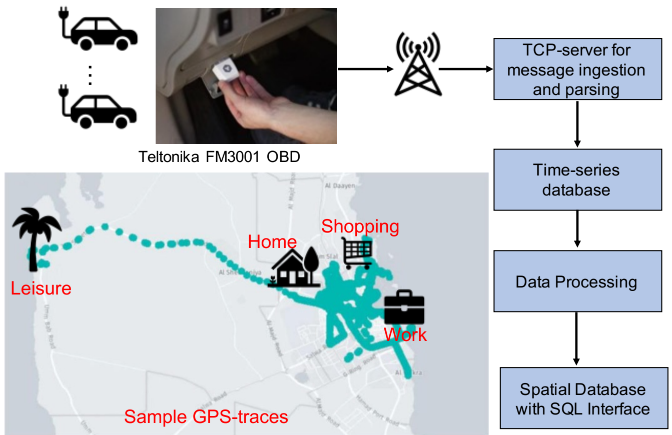

Fig. 1. Schematic Overview of the Data Acquisition System .

tend to point to non-calibrated location outside of the state of Qatar. The second condition is that it is assumed that the trip only starts when the difference between consecutive timestamps is more 10 seconds.

D. Location Analysis

Location analysis is one of the most critical parts of this study to cluster the locations of the vehicles. Most buildings and parking infrastructures in Qatar have large physical spaces (e.g. malls, beaches, etc.) and the gathered data need to be accurately mapped to the correct location. To begin our location analysis, we need to be able to plot data points on a map. We extract all points from our dataset for a single car where the trip field is "-1" i.e. the car is stationary. For accuracy purposes, we change the Coordinate Reference System(CRS) of the points to a UTM-based CRS so that distance between the points can be measured in meters. Our purpose is to automatically identify points that are close together while making sure points which are not nearby are not clustered and ignored. These"ignored" points could represent anomalies that could possibly be caused by a faulty gps reading. We also want to avoid creating large unnatural clusters that do not properly segment the locations.

To cluster data points, density-based spatial clustering (DBSCAN) algorithm is adopted to find arbitrary shaped clusters [18]. In this algorithm, a point is assigned to a cluster (e.g. home) if it is close to many other points from that cluster. DBSCAN algorithm has two input parameters ϵ (eps) and the minPts that shows the minimum number of points required to form a dense region. The DB-scan algorithm works as follows:

1) Random D points are selected within the dataset;
2) All points within the ϵ radius to D points are identified as core points;
3) The number of core points keeps increasing till a border condition is reached;
4) The border condition identifies border points which have at least one core point within its ϵ radius but does not have the required number of minPts to extend the cluster. The border points act as boundaries of the cluster;
5) All points which are not designated as core or border are designated as noise.

To tune these parameters in an automatic manner grid search method is adopted. For evaluation purposes, we need a metric that can inform us about the quality of clustering. In our experiments, we use the Silhouette Score metric [19] . This metric ranges from -1 to 1. A score close to boundaries represent that the clusters are well distinguished, while a score that is close to 0 means the clustering is indifferent. The Silhouette Score is given by

$$\text{Silhouette Score} = \frac{b-a}{\max(a,b)} \tag{1}$$

where a denotes the mean intra-cluster distance and b denotes the inter cluster distance. For instance, for the first volunteer's home location, the ϵ is calculated as 55 meters and minPnts is found as 9.

IV. Results

Our initial analysis investigates the car locations and activities during an average day. Since driver's mobility patterns depend day of the week, weekend and weekday results are presented separately. As shown in Figs. 2 and 3 location-based daily activities are grouped into six categories as home, work, leisure (beaches, sport events, gym etc.), school (to pick up kids), shopping (a major attraction point in the region), and others. Since vast majority of the time the cars are park at home at night, only activities between 6am to 10pm are shown. This results are critical to evaluate potential locations for EV drivers to recharge their EVs without major disruptions to their daily routines. For instance, in both weekend and weekday cases home charging is the most convenient charging option the probability of being parked at home is significantly higher than other cases. This is mainly because vast majority of the population in resides in the capital Doha, and there is limited activities and events in the rest of the country.

Workplace charging during weekdays also represent a good option for drivers who do not have home charging. As shown in Fig. 2, volunteers who drive to work and park have a charging window of nearly 8 hours (7am to 4pm) which is sufficient to charge a typical EV with a level 2 charger. Parking at shopping malls are relatively shorter than other locations, therefore, fast charging nodes are better suited.

Our second analysis examines the daily trips (in km) made by volunteers to estimate number of recharging needed during the day. As a first step, days with no trips are removed. This is because there is limited or no mobility during extended holiday season and days with Covid19 lockdown. To that end, 1100 daily trips of all volunteers are analyzed and empirical cumulative distribution function of all cars and driver #5 are calculated and presented in Fig. 4. From this results, it can be seen that nearly 80% of daily trips are less than 100 km and only 3% of the trips are higher than 200 km. The case for the EV is considerable different than the rest of the cars. The

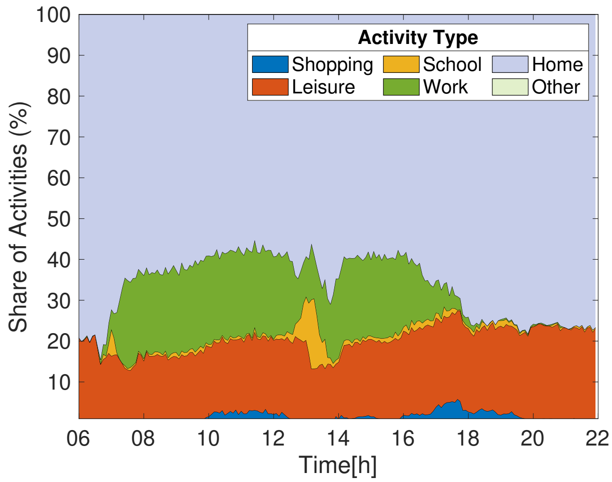

Fig. 2. Car locations and activities during an average weekday.

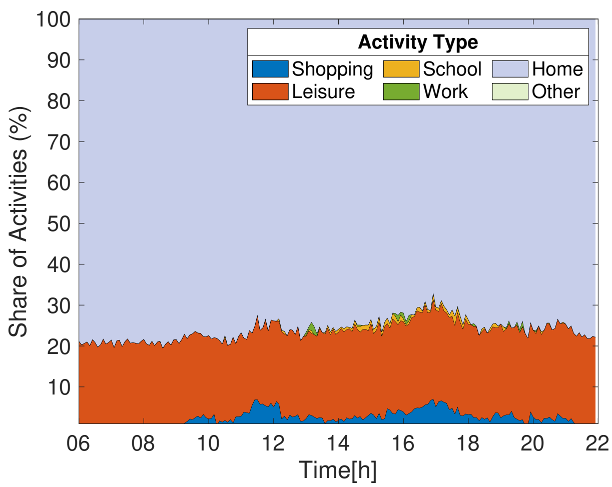

Fig. 3. Car locations and activities during an average weekend.

average daily trip for the EV is about 19 km, 80% of the trips are less than 26 km and the highest daily trip is 47 km. This is because due to the lack of charging infrastructure in the country, most trips took place between home and work, where recharging mostly occurred. As a consequence, the owner of the EV used another car (petrol) to complete other trips. The results reveal that the country's infrastructure require upgrades to support mainstream EV acceptance.

Detailed results for each car on the trip statistics are presented in Table II. It can be seen that there is a wide variation among median trip lengths and the distance between home and work is the primary factor determining the trip lengths. For instance, drivers #2 and #3 live in the same area and work at the same university. As a result, their trip statistics are close to each other. This is also the case for drivers #1 and #4 who has similar daily trip patterns. Vehicle #6 has the highest milage as this vehicle is driven by a professional driver and serves multiple people from the same household.

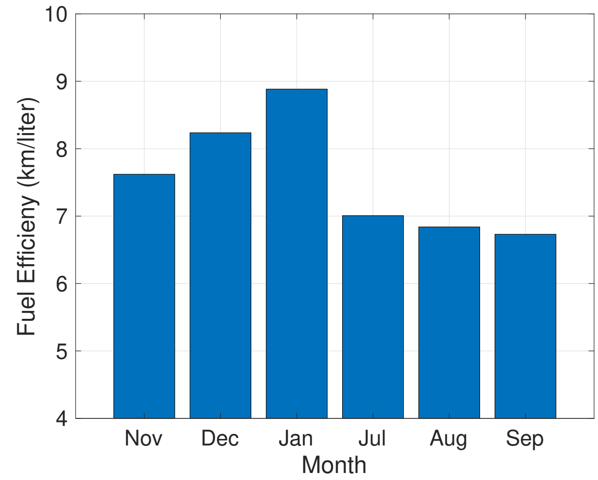

Fig. 4. Cumulative distribution function of daily trips of EV and all cars.

TABLE II
TRIP STATISTICS FOR THE STUDY MEDIAN, AND QUARTILES.

	Daily Trip Length (km)				
ID	Median	25%	75%	Total Distance (km)	Total Trips
1	38.5	26.5	59.8	11224.3	12461
2	75.0	30.2	113.8	7526.0	22341
3	71.7	46.2	100.9	10085.0	21131
4	32.2	26.3	45.3	7072.5	23593
5	20.8	8.1	23.9	784.5	19207
6	76.2	45.1	117.6	20712.9	17446
7	51.3	25.3	68.7	10096.5	19078

A significant challenge that residents in Qatar and the Gulf region face is hot desert climate which necessitate the use of air condition both indoors and cars. Therefore, during summer months vehicles consume extra energy for air conditioning and fuel efficiency of the card reduce. As Volunteer #1 owns the only car that allows our OBD device to read fuel levels, this vehicle's consumption levels are compared and presented in Fig. 5. It can be seen that the fuel efficiency (km/liter) reduces by 24% due to AC usage.

An intensified impact will be experienced by EV drivers as EV batteries are not designed to operate under hot arid climates. To improve the battery performance and ensure the safety, battery management systems use significant amount of stored energy to cool the battery [20]. Moreover, driver aggressiveness, a major issue in Qatar (and in Gulf region), increases energy demand of EVs and reduces daily driving range. In [21], an empirical study that shows the driving ranges of popular EVs under different temperatures and driving performances [22] are presented. Using popular EV models, EV trip completion rates using the statistics given in Fig. 5, are presented for winter and summer months. As presented in Table III, most of the daily trips in winter can be completed with small-size EVs. On the other hand, summer trips require significantly higher amount of battery packs.

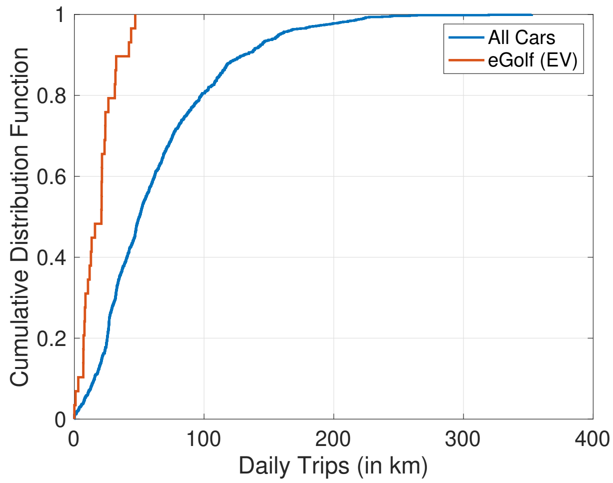

Fig. 5. Monthly fuel efficiency for Car #1.

TABLE III
ASSESSMENT OF POPULAR EV MODELS AND CORRESPONDING TRIP COMPLETION RATES.

EV Type	Battery Pack (kWh)	Trip Completion Rate (%) - Winter	Trip Completion Rate (%) - Summer
BMW i3	33.2	96	78
Hyundai Kona	64	100	97
Nissan Leaf	40	95	79
Kia Soul	64	100	97
Toyota Rav4	40	90	69

It is noteworthy that the real world impacts could be more significant due to the following reasons. First, data presented in [21] shows driving ranges up to 43 °C and the actual temperatures during the day could easily reach up to 50 °C in Qatar. Second, EVs need to be parked in shady areas and need to be connected to a charger to cool down the battery. If there is lack of chargers, then the stored energy will be used and daily driving ranges will further reduce. Third, to ensure safety battery charge-discharge levels are advised to be in 20%-80% state of charge. Therefore higher battery packs would be needed.

V. CONCLUSION

ACKNOWLEDGMENT

This publication was made possible by NPRP grant # NPRP12S-0214-190083 from the Qatar National Research Fund (a member of Qatar Foundation). The statements made herein are solely the responsibility of the authors.

REFERENCES

[1] T. Haasz, J. J. G. Vilchez, R. Kunze, P. Deane, D. Fraboulet, U. Fahl, and E. Mulholland, "Perspectives on decarbonizing the transport sector in the eu-28," *Energy strategy reviews*, vol. 20, pp. 124–132, 2018.

[2] N. Melton, J. Axsen, and D. Sperling, "Moving beyond alternative fuel hype to decarbonize transportation," *Nature Energy*, vol. 1, no. 3, pp. 1–10, 2016.

[3] I. S. Bayram, U. Zafar, and S. Bayhan, "Could petrol stations play a key role in transportation electrification? a gis-based coverage maximization of fast ev chargers in urban environment," *IEEE Access*, vol. 10, pp. 17 318–17 329, 2022.

[4] S. Feng and C. L. Magee, "Technological development of key domains in electric vehicles: Improvement rates, technology trajectories and key assignees," *Applied Energy*, vol. 260, p. 114264, 2020.

[5] İ. Ş. Bayram and M. Koç, "Demand side management for peak reduction and pv integration in qatar," in *2017 IEEE 14th International Conference on Networking, Sensing and Control (ICNSC)*. IEEE, 2017, pp. 251–256.

[6] G. M. Fetene, S. Kaplan, S. L. Mabit, A. F. Jensen, and C. G. Prato, "Harnessing big data for estimating the energy consumption and driving range of electric vehicles," *Transportation Research Part D: Transport and Environment*, vol. 54, pp. 1–11, 2017.

[7] Y. Al-Wreikat, C. Serrano, and J. R. Sodré, "Driving behaviour and trip condition effects on the energy consumption of an electric vehicle under real-world driving," *Applied Energy*, vol. 297, p. 117096, 2021.

[8] G. Pareschi, L. Küng, G. Georges, and K. Boulouchos, "Are travel surveys a good basis for ev models? validation of simulated charging profiles against empirical data," *Applied Energy*, vol. 275, p. 115318, 2020.

[9] Y. Wang and D. Infield, "Markov chain monte carlo simulation of electric vehicle use for network integration studies," *International Journal of Electrical Power & Energy Systems*, vol. 99, pp. 85–94, 2018.

[10] Y. Xu, S. Çolak, E. C. Kara, S. J. Moura, and M. C. González, "Planning for electric vehicle needs by coupling charging profiles with urban mobility," *Nature Energy*, vol. 3, no. 6, pp. 484–493, 2018.

[11] C. Gaete-Morales, H. Kramer, W.-P. Schill, and A. Zerrahn, "An open tool for creating battery-electric vehicle time series from empirical data, emobpy," *Scientific data*, vol. 8, no. 1, pp. 1–18, 2021.

[12] L. Kessler and K. Bogenberger, "Mobility patterns and charging behavior of bmw i3 customers," in *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2016, pp. 1994–1999.

[13] M. Sodenkamp, J. Wenig, F. Thiesse, and T. Staake, "Who can drive electric? segmentation of car drivers based on longitudinal gps travel data," *Energy policy*, vol. 130, pp. 111–129, 2019.

[14] R. Galvin, "Energy consumption effects of speed and acceleration in electric vehicles: Laboratory case studies and implications for drivers and policymakers," *Transportation Research Part D: Transport and Environment*, vol. 53, pp. 234–248, 2017.

[15] I. S. Bayram, "A stochastic simulation model to assess the impacts of electric vehicle charging on power generation: a case study for qatar," in *2019 IEEE Transportation Electrification Conference and Expo (ITEC)*. IEEE, 2019, pp. 1–5.

[16] Teltonika telematics fm3001. [Online]. Available: https://teltonika-gps.com/product/fm3001/

[17] Influx data time series platform. [Online]. Available: https://www.influxdata.com/

[18] K. Khan, S. U. Rehman, K. Aziz, S. Fong, and S. Sarasvady, "Dbscan: Past, present and future," in *The fifth international conference on the applications of digital information and web technologies*. IEEE, 2014, pp. 232–238.

[19] K. R. Shahapure and C. Nicholas, "Cluster quality analysis using silhouette score," in *2020 IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA)*. IEEE, 2020, pp. 747–748.

[20] I. S. Bayram, "Impacts of electric vehicle charging under cold weather on power networks," in *2021 56th International Universities Power Engineering Conference (UPEC)*. IEEE, 2021, pp. 1–6.

[21] "Understand how temperature affects how far your ev can go on a full battery." [Online]. Available: https://www.geotab.com/blog/ev-range/

[22] C. Dias, M. Kharbeche, D. Muley, A. Kashem, M. A. Fahed, S. V. Iral, and N. Abdelfattah, "Are professional drivers more aggressive than general drivers? a case study from doha, qatar," *Procedia Computer Science*, vol. 201, pp. 16–23, 2022.