QUANTIFYING THE IMPACT OF TRAFFIC ON ELECTRIC VEHICLE EFFICIENCY

Tim Jones
University of Rhode Island, t_jonas@my.uri.edu

Follow this and additional works at: https://digitalcommons.uri.edu/theses

Recommended Citation
https://digitalcommons.uri.edu/theses/1496

This Thesis is brought to you for free and open access by DigitalCommons@URI. It has been accepted for inclusion in Open Access Master's Theses by an authorized administrator of DigitalCommons@URI. For more information, please contact digitalcommons@etal.uri.edu.
QUANTIFYING THE IMPACT OF TRAFFIC
ON ELECTRIC VEHICLE EFFICIENCY

BY

TIM JONAS

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
IN
SYSTEMS ENGINEERING

UNIVERSITY OF RHODE ISLAND

2019
Abstract

Battery electric vehicle (BEV) sales in the United States (US) are constantly growing since 2010, resulting in 238,000 units in 2017. While the impact of various factors on the energy efficiency of internal combustion engines vehicles (ICEVs) was vastly explored in the past, research is recently focusing on BEVs.

The primary, major restraint for electric vehicles is the range limitation. The range of current vehicles on the market vary between 93-315 miles per charge. Combined with the long charging times for batteries, the suitability for daily use is restricted, yet their range is a significant factor of daily requirements of users. Drivers of electric vehicles sometimes fear running out of power, a phenomenon called range anxiety. In order to extend the range of electric vehicles and because BEVs market share is growing globally and nationally, understanding the impact of different range impacting factors (e.g., traffic, temperature, driver behavior) is essential.

The impact of traffic should be understood carefully since it is impressionable by driver behavior, eco-driving strategies, and range management. However, the impact of traffic on the range and, thus, the efficiency of electric vehicles under real-world conditions has not yet entirely transitioned to include BEVs. To quantify the impact of traffic on electric vehicle efficiency, an empirical experiment was conducted with a 2017 eGolf on a predetermined test route. All 30 participants drove the test route twice: once with increased traffic congestion during the morning commute, and once in low traffic congestion during the day. Time was a controlling factor to distinguish two scenarios with different traffic
intensities. Finally, both data sets (i.e., traffic data and non-traffic data) were compared to quantify the impact of traffic on the efficiency of BEV under real-world conditions.

Different measures were investigated to provide evidence of differences in the intensity of traffic on the chosen test route based on daytime. Outcomes provide evidence that traffic influences the acceleration and change in acceleration on the test route. A multiple linear regression was applied to quantify the impact of traffic on the difference in state of charge per mile of BEVs. Additionally, driver, temperature and initial state of charge were included in the model and investigated for significance. Results show that among all considered factors, temperature has the highest impact on the energy consumption of BEVs. A stepwise regression was carried and based on the results of both regression models, the influence of traffic could be quantified as to increase the difference in state of charge per mile by up to 2.6% respectively 0.0066 kWh compared to a non-traffic scenario. A logistic regression model was applied to confirm the positive correlation between traffic congestion and BEV energy consumption affirming previous findings.

The investigations based on naturalistic driving data provide new findings about the consumption behavior of BEVs in traffic. These results can help drivers to overcome range anxiety and range limitations by adapting new eco-driving strategies when considering traffic. Additionally, transportation, as well as navigation systems, can be improved. Manufacturers can benefit by using the findings for the development and improvement of electric drive trains and batteries, as well as routing algorithms.
Acknowledgments

First and foremost, I would like to thank my major professor Dr. Gretchen A. Macht for her guidance and infinite support with this research. Her valuable feedback and enormous commitment helped this thesis to be finished. She was always there to aid with the progress of this research. Her dedication to science and research has always been an inspiration, as well as motivation for me. I look forward to hearing about her continued success as a professor at the University of Rhode Island and wish her all the best.

I would also like to thank my inside committee member Dr. Jyh-Hone Wang for giving advice in the statistical analyses of this work. His motivation throughout the thesis process and his valuable counseling for the experimental part in this research was very valuable. Furthermore, I would like to thank Dr. Christopher Hunter, who served as my outside committee member. He has made great additions to this thesis that make the research even more relevant in the realm of transportation.

Additionally, I would like to thank my colleagues in the SIS Lab for their relentless support. I would particularly like to thank Taris Wilde for his participation in the process of data collection and his support and comments throughout the whole research. Advice and comments given by Nicholas Bernardo have been a great help in the process of reviewing the thesis.

Lastly, I would like to express my deep gratitude to Bettina Jonas and my grandparents Ursula and Heinz-Dieter Kampfhofer for their support during my studies and the given possibility to participate in the graduate program at the University of Rhode Island. Without their enormous support, I would not be here today.
Table of Contents

ABSTRACT ........................................................................................................................................... II

ACKNOWLEDGMENTS ...................................................................................................................... IV

LIST OF FIGURES ............................................................................................................................. VII

LIST OF TABLES ............................................................................................................................... IX

LIST OF EQUATIONS ......................................................................................................................... X

LIST OF ABBREVIATIONS ................................................................................................................. XI

1 INTRODUCTION .......................................................................................................................... 1

1.1 BACKGROUND .......................................................................................................................... 1

1.2 RESEARCH GOALS .................................................................................................................. 5

2 REVIEW OF LITERATURE .............................................................................................................. 7

3 METHODOLOGY .......................................................................................................................... 15

3.1 EXPERIMENTAL DESIGN ....................................................................................................... 15

3.1.1 Test Drivers ......................................................................................................................... 15

3.1.2 Test Equipment .................................................................................................................. 17

3.1.3 Test Route .......................................................................................................................... 18

3.1.4 Test Time .......................................................................................................................... 22

3.1.5 Auxiliary Systems ............................................................................................................... 26

3.2 DEPENDENT AND INDEPENDENT VARIABLES ................................................................. 29

3.3 METHODS OF DATA COLLECTION AND HANDLING ....................................................... 30
List of Figures

Figure 1: PEV sales trend in the US 2012-2018.................................................................2

Figure 2: Causal diagram for the energy consumption of BEVs.........................................4

Figure 3: Literature grouped by subjects of investigation....................................................9

Figure 4: Speed-Energy consumption relationship for electric vehicles.............................13

Figure 5: Speed-Energy consumption relationship for internal combustion vehicles......13

Figure 6: Gender of participants........................................................................................17

Figure 7: Test route (a) and road classification (b) for South Kingstown, Rhode Island..20

Figure 8: Altitude along the Test Route .................................................................................21

Figure 9: Temperature at the Beginning of each Test Drive.................................................25

Figure 10: Method of data collection and calculation............................................................31

Figure 11: Starting point of test route (a) and ending point of test route (b) with coordinates ........................................................................................................................................37

Figure 12: Stretch of the route with construction.................................................................38

Figure 13: Speed profile a) before and b) during construction.............................................38

Figure 14: Comparison of Speeds on route stretch with construction...............................39

Figure 15: ΔSOC/mile for every driver and scenario............................................................47
Figure 16: Normality plots for different traffic measures ........................................52

Figure 17: Boxplots for different traffic measures ....................................................53

Figure 18: Normality plots of average consumption in scenarios a) traffic b) no traffic ..56

Figure 19: Boxplots of average energy consumption for both scenarios ..................56

Figure 20: Histogram SOC/mile ..............................................................................58

Figure 21: Average energy consumption over driver and scenario .........................59

Figure 22: Residual vs. fit for whole model ..............................................................63

Figure 23: Scale-location plot for whole model .........................................................64

Figure 24: Normality plot for whole model ..............................................................64

Figure 25: a) Residual vs. fit and b) scale-location plot for reduced model ...............65

Figure 26: Increase in ΔSOC for different driver types and models .......................69

Figure 27: Differences in range between scenarios for different drivers and models ....70

Figure 28: ROC curve for initial logistic regression model ......................................72

Figure 29: ROC curve for logistic Regression model considering ΔSOC/mile ..........74
List of Tables

Table 1: Vehicle classes with respective engine types and characteristics ........................................11

Table 2: Technical features of the 2017 eGolf ..................................................................................18

Table 3: Road type classification Rhode Island ..................................................................................19

Table 4: Factors and corresponding types of statistical data ..................................................................29

Table 5: Sample of raw data as downloaded from the cloud service of Fleetcarma .........................32

Table 6: Cutting procedure for test runs ............................................................................................41

Table 7: Results of the Shapiro-Wilk-Normality-Tests for traffic measures .......................................54

Table 8: Two-sided paired t-test results for different traffic measures .................................................55

Table 9: Results of the Shapiro-Wilks-Normality-Test for average consumption ............................57

Table 10: Two-sided paired t-test results for different average consumptions ....................................60

Table 11: ANOVA Table for full Model ...............................................................................................62

Table 12: ANOVA table for reduced model .........................................................................................66

Table 13: Results of initial logistic regression .....................................................................................72

Table 14: Results of logistic regression with coefficient ΔSOC/mile .....................................................73
List of Equations

Equation 1: Transformation equation from SOC in % to SOC in kWh ...................... 33

Equation 2: Power-law ........................................................................................................ 33

Equation 3: Electrical energy formula........................................................................... 34

Equation 4: Electrical energy for every second measured ............................................. 34

Equation 5: Total electrical energy used in a drive ......................................................... 34

Equation 6: Transformation from velocity [km/h] into distance [miles] ....................... 35

Equation 7: Calculation of total distance driven .............................................................. 36

Equation 8: Calculation of average energy consumption for every test drive .......... 36

Equation 9: Target function for detecting cutting points ............................................... 40

Equation 10: Calculation of jerk ....................................................................................... 44

Equation 11: Linear multiple regression equation ......................................................... 46

Equation 12: Regression equation for setting "No Traffic" ............................................. 66

Equation 13: Regression equation for setting "Traffic" ................................................... 66
# List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A/C</td>
<td>Air conditioning</td>
</tr>
<tr>
<td>ANOVA</td>
<td>Analysis of variance</td>
</tr>
<tr>
<td>AUC</td>
<td>Area under the curve</td>
</tr>
<tr>
<td>BEV</td>
<td>Battery electric vehicle</td>
</tr>
<tr>
<td>ECU</td>
<td>Electronic control unit</td>
</tr>
<tr>
<td>EPA</td>
<td>United States Environmental Protection Agency</td>
</tr>
<tr>
<td>EV</td>
<td>Electric vehicle</td>
</tr>
<tr>
<td>HVAC</td>
<td>Heating, ventilation, and air conditioning</td>
</tr>
<tr>
<td>ICE</td>
<td>Internal combustion engine</td>
</tr>
<tr>
<td>ICEV</td>
<td>Internal combustion engine vehicle</td>
</tr>
<tr>
<td>IQR</td>
<td>Interquartile range</td>
</tr>
<tr>
<td>ITS</td>
<td>Intelligent transportation system</td>
</tr>
<tr>
<td>URI</td>
<td>University of Rhode Island</td>
</tr>
<tr>
<td>Lbf-ft</td>
<td>Pound-foot</td>
</tr>
<tr>
<td>NGSIM</td>
<td>Next generation simulation</td>
</tr>
<tr>
<td>OAT</td>
<td>Outside air temperature</td>
</tr>
<tr>
<td>PEV</td>
<td>Plug-in electric vehicles</td>
</tr>
<tr>
<td>PHEV</td>
<td>Plug-in hybrid electric vehicle</td>
</tr>
<tr>
<td>RI</td>
<td>Rhode Island</td>
</tr>
<tr>
<td>ROC</td>
<td>Receiver operating characteristics</td>
</tr>
<tr>
<td>SOC</td>
<td>State of charge</td>
</tr>
<tr>
<td>US</td>
<td>United States</td>
</tr>
<tr>
<td>UTC</td>
<td>Universal time coordinated</td>
</tr>
<tr>
<td>VBA</td>
<td>Visual Basic for Applications</td>
</tr>
</tbody>
</table>
1 Introduction

1.1 Background

Electric vehicles (EVs) have significantly increased in their importance on the national and international stage due to market share growth. Changes in both customer demand and government regulations are influencing the increasing trend in sales of EV drive trains.

Within the last five years, the worldwide total number of battery electric vehicles (BEVs) on the road increased nearly fifteen-fold from 0.23 to 3.29 million units (International Energy Agency, 2019).

By 2030, it is estimated that only 52% of the vehicles sold worldwide will be powered by an internal combustion engine (ICE) (Mosquet et al., 2018). As of today, the highest market share among registered Plug-In Electric Vehicles (PEVs) can be found in Norway with 39.2% in 2017. There is a significant gap between Norway and the country with the second largest market share, Sweden with 6.3%. The United States (US) has the seventh largest PEV market share worldwide with respect to national vehicle registrations with 1.2%. (International Energy Agency, 2018). Figure 1 shows the sales trend of PEVs in the United States between 2012 and 2018. PEV sales are constantly increasing and reached 361,000 units in 2018; only in 2015 PEV sales slightly decreased compared to 2014. The share of PEVs on newly registered cars in the US was steadily increasing in the past and reached a 2.1% incline in 2018, ending in a total of 361,000 units. The BEV share of PEVs increased
from 53% in 2017 to 66% in 2018, demonstrating that BEVs, in particular, are becoming progressively popular throughout the US (Irle, 2018).

Figure 1: PEV sales trend in the US 2012-2018
(Data adapted from Irle, 2018)

Especially the high contribution of air pollution from internal combustion engines plays a major role in the automotive outlook. According to the US Environmental Protection Agency (EPA), transportation in the economic sector caused 28% of the 2016 national greenhouse gas emissions (GHGs) (EPA, 2016). BEVs promise a reduction in the dependency on petroleum, improved energy efficiency, and enhanced sustainability potential (Agrawal, Zheng, Peeta, & Kumar, 2016). Survey results show, 47% of respondents in the US are currently worried a "great deal" about air pollution (Jones &
Saad, 2017). Already 16% of American citizens choose their car, make and model, under the consideration of making their driving behavior more eco-friendly (Statista, 2018). The potential of BEVs with regard to a reduction of GHGs is fortified by the fact that the energy consumption of BEVs is more sensitive to eco-driving behavior compared to ICEs (Miyatake, Kuriyama, & Yamamoto, 2011; Romm & Frank, 2006). Due to the recuperation feature which allows the generation of energy while breaking or even during times when the car is simply rolling (Wu, Freese, Cabrera, & Kitch, 2015), BEVs have an eco-driving saving potential of 30% (Walsh, Carroll, Eastlake, & Blythe, 2010).

Eco-driving is defined as “driver behavior targeted towards increased energy efficiency” (Arend & Franke, 2019). According to Arend and Franke (2019), the energy consumption of vehicles is influenced by a driver’s vehicle selection, route selection, and eco-driving strategies. Eco-driving strategies consist of a set of different factors that are considered and chosen by a driver (Sivak & Schoettle, 2012). Traffic, as part of eco-driving strategies, can intentionally be avoided by drivers to reduce energy consumption. Figure 2 illustrates the mentioned interactions between driving behavior, eco-driving, traffic, and energy consumption.
As of today, 37% of survey participants avoid traffic to become more eco-friendly drivers (Statista, 2018). Understanding BEV efficiency influencing factors will help to improve eco-driving, as well as the electrification of global transportation.

However, BEVs aptitude for daily use is debated and generally restricted by their limited range and their long charging time (Wager, Whale, & Braunl, 2016). As of September 2017, the highest range of EVs could be reached with a Tesla Model S with nearly 300 miles, while other manufacturers such as Volkswagen reach a range of just 125 miles with the 2017 eGolf (fleetcarma.com, 2018). The term “range anxiety” is often used in this context and describes the user’s fear of ending up with a discharged battery before reaching their destination (Tate, Harpster, & Savagian, 2008). In order to overcome range anxiety, the impact of different factors on their efficiency must be quantifiably understood and validated in a naturalistic driving environment.

Vast research was conducted in the past that aimed to investigate different factors influencing the efficiency of internal combustion engine vehicles (ICEVs) and electric

Figure 2: Causal diagram for the energy consumption of BEVs

As of today, 37% of survey participants avoid traffic to become more eco-friendly drivers (Statista, 2018). Understanding BEV efficiency influencing factors will help to improve eco-driving, as well as the electrification of global transportation.

However, BEVs aptitude for daily use is debated and generally restricted by their limited range and their long charging time (Wager, Whale, & Braunl, 2016). As of September 2017, the highest range of EVs could be reached with a Tesla Model S with nearly 300 miles, while other manufacturers such as Volkswagen reach a range of just 125 miles with the 2017 eGolf (fleetcarma.com, 2018). The term “range anxiety” is often used in this context and describes the user’s fear of ending up with a discharged battery before reaching their destination (Tate, Harpster, & Savagian, 2008). In order to overcome range anxiety, the impact of different factors on their efficiency must be quantifiably understood and validated in a naturalistic driving environment.

Vast research was conducted in the past that aimed to investigate different factors influencing the efficiency of internal combustion engine vehicles (ICEVs) and electric
vehicles (EVs). Due to the unique characteristics of BEVs, the previous findings of the efficiency of ICEs do not necessarily apply to BEVs. Separate investigations focusing solely on BEVs are essential to understand their consumption behavior, especially the factor of traffic as it is related to specific acceleration and speed profiles as these relationships differ between BEVs and ICEVs. Yet, the impact of traffic on the energy consumption of BEVs is not investigated thoroughly.

Understanding how different traffic conditions influence energy efficiency contributes to improvements in intelligent transportation systems (ITS), and BEV efficiency. Since driving behavior is an interaction between driver and vehicle, the results of this research aim to lead to a better understanding of driver strategies towards an eco-friendly driving behavior and thus, prolonged range limitations and increased usability of BEVs. The testing in a naturalistic driving setting will lead to new empirical findings that can advance the suitability of BEVs in daily use for all users.

1.2 Research Goals

The contribution of this research will serve to quantify the influence of traffic on the energy consumption of BEVs and, thus, its in-use range. An experiment was set up through a rigorous experimental design and statistical analyses applied to quantify the impact of traffic on energy consumption. Data were collected on 30 drivers from a 27.6 miles long route through South Kingstown, in Washington County, Rhode Island. Electronic sensors connected to the car’s On-Board Diagnostics (OBD) were used to extract raw data from the test runs. Each driver drove the route twice, once with increased congestion (in the following referred to as “Traffic” scenario) and once in a low congestion situation (in the
following referred to as “No Traffic” scenario), as predicted by the time of day for each test run. In order to get comparable data, the route chosen was held constant throughout all test drives and test scenarios. Both data sets were analyzed for evidence of differences in congestion and energy consumption in order to quantify the impact of traffic on the energy consumption of BEVs in a naturalistic environment.

A review of the literature is presented in Chapter 2, encompassing the influence of different factors on BEV energy consumption and the impact of traffic on the fuel consumption of vehicles in general. A discussion of why traffic as a factor for energy consumption has to be investigated separately for BEVs is given.

Chapter 3 provides a description of the applied methodology of data collection. Therein the design of experiment and calculations to obtain necessary measures, as well as different methods used for the statistical analyses, will be explained.

In Chapter 4 the results from the different statistical and mathematical analyses are presented and discussed. A multiple linear regression is applied as an initial model. A stepwise regression was used to improve the model fit. An a posteriori analysis in form of a logistic regression model was then carried out to confirm previous assumptions.

The last chapter, Chapter 5, concludes the results obtained from this research and compared to previous findings in literature. Furthermore, limitations of this work are explained, recommendations given, and potential for future work discussed.
2 Review of Literature

ICEVs and BEVs have been extensively studied for energy consumption impacting factors, such as:

- **road type** (Agrawal et al., 2016; Fontaras, Zacharof, & Ciuffo, 2017; Sivak & Schoettle, 2012; Walsh et al., 2010; Wang & Boggio-Marzet, 2018; Wu et al., 2015; Yao, Yang, Song, & Zuo, 2013),

- **speed** (Agrawal et al., 2016; Badin et al., 2013; Fontaras, Pistikopoulos, & Samaras, 2008; Fontaras et al., 2017; Grubwinkler, Brunner, & Lienkamp, 2014; Treiber, Kesting, & Thiemann, 2008; Wager et al., 2016; Wu et al., 2015),

- **acceleration** (Bingham, Walsh, & Carroll, 2012; Treiber et al., 2008; Wai, Rong, & Morris, 2015; Wu et al., 2015),

- **driving behavior** (Agrawal et al., 2016; Arend & Franke, 2019; Badin et al., 2013; Bingham et al., 2012; Fontaras et al., 2017; Franke, Schmalfuß, & Rauh, 2018; Gonder, Markel, Simpson, & Thornton, 2007; Jägerbrand & Sjöbergh, 2016; Neumann, Franke, Cocron, Bühler, & Krems, 2015; Pearre, Kempton, Guensler, & Elango, 2011; Sivak & Schoettle, 2012; van der Zwaag et al., 2013; Wai et al., 2015; Walsh et al., 2010; Wang & Boggio-Marzet, 2018; Yuksel & Michalek, 2015),

- **outside temperature** (Alvarez & Weilenmann, 2012; Bartels, Kowalsky, Jonas, & Macht, 2019; Fontaras et al., 2008, 2017; Jägerbrand & Sjöbergh, 2016; Kambly & Bradley, 2014; Yuksel & Michalek, 2015), auxiliaries (Badin et al., 2013; Bingham et al., 2012; Fontaras et al., 2017; Haworth & Symmons, 2001; Johnson, 2010; Kambly & Bradley, 2014; Nemry, Leduc, Mongelli, & Uihlein, 2008; Yuksel &
Michalek, 2015), and traffic (Bigazzi & Clifton, 2015; Cole, 2016; Fiori et al., 2019; Fontaras et al., 2017; Grubwinkler et al., 2014; Haworth & Symmons, 2001; Jereb, Kumperšcak, & Bratina, 2018; Kessler & Bogenberger, 2015; Miyatake et al., 2011; Sivak & Schoettle, 2012; Treiber et al., 2008; Wai et al., 2015).

Comprehensively, these investigations indicate that different factors can influence the energy consumption of vehicles in general. Figure 3 illustrates the amount of literature grouped by subjects of investigation in proportion to the size of the area. Overlapping areas show that several works investigated multiple factors with respect to the energy consumption of vehicles because of the correlations between their areas and the resulting complexity. Traffic, in particular, is correlated with speed, acceleration, and outside factors (e.g., temperature, precipitation, and light), and driving behavior.
Traffic is defined as the “level of service” (National Research Council, 2000), which is a qualitative measure based on metrics such as speed, travel time, traffic interruptions, comfort, and convenience. Traffic density can also be described as a function of speed and flow (Barth & Boriboonsomsin, 2008; Grubwinkler et al., 2014), while these factors interact with each other. Hence, traffic congestion itself influences the speed profile (e.g., average travel velocity and acceleration), as well as the driver’s behavior (i.e., route selection) (Fontaras et al., 2017) and ultimately the energy consumption of vehicles (Bigazzi & Clifton, 2015).

Various literature exists articulating the relationship between traffic and energy consumption of ICEVs (Fiori et al., 2019; Jereb et al., 2018; Treiber et al., 2008). Treiber
et al. (2008) used velocity and acceleration data as input for Next Generation Simulation (NGSIM) to measure the impact of traffic on ICE fuel consumption. Their results indicate that congestion can increase fuel utilization by 80%.

Jereb et al. (2018) conducted a real-field experiment on a 3.2 km section of road while obtaining traffic counting data. Their results show a possible increase in fuel consumption in a worst-case scenario of approximately 60% compared to the best-case scenario. These findings were supported by Fontaras et al. (2017), Haworth & Symmons (2001), and Sivak & Schoettle (2012). Sivak and Schoettle (2021) stated that fuel consumption could increase by 20-40% due to congestion based on simulation results. Hence, the literature has evidently shown that traffic has a significant influence on the fuel consumption of ICEs.

However, findings of the consumption behavior of ICEs in congested scenarios are not necessarily known to be transformable to EVs. Although some factors count equally for BEVs and ICEs (e.g., drag coefficient, wind resistance), BEVs have specific features that make separate investigations necessary. The latter are equipped with regenerative braking system (Agrawal et al., 2016; Miyatake et al., 2011; Romm & Frank, 2006), making the energy flow bidirectional (Franke et al., 2018). Thus, the impacts of different factors on BEV efficiency must be quantifiably understood and validated in a naturalistic driving environment. With regard to congestion, Romm and Frank (2006) point out that plug-in hybrid electric vehicle (PHEVs) could potentially even increase the mileage in stop-and-go traffic.

For propulsion, hybrid vehicles are powered by both, an electric motor and an internal combustion engine (ICE) by burning diesel or gasoline (Curtin, Shrago, & Mikkelsen,
BEVs instead are simply equipped with a battery and an electric engine (Agrawal et al., 2016). Table 1 groups the different layouts of EVs and shows their characteristics.

*Table 1: Vehicle classes with respective engine types and characteristics*

<table>
<thead>
<tr>
<th>Class</th>
<th>Engine Group</th>
<th>Engine Type</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>EV</td>
<td>Fully electric vehicles</td>
<td>Battery electric vehicles (BEVs)*</td>
<td>Only powered by electric engine; charged externally</td>
</tr>
<tr>
<td></td>
<td>Hybrid vehicles</td>
<td>Plug-in hybrid electric vehicles (PHEVs)*</td>
<td>Powered by petrol and electricity; charged externally or through ICE</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hybrid electric vehicles (HEVs)</td>
<td>Start driving electric and switch to petrol engine as speed rises; charged using regenerative breaking</td>
</tr>
<tr>
<td>ICEV</td>
<td>Internal combustion vehicles</td>
<td>Internal combustion engines (ICEs)</td>
<td>Powered only by petrol; no electric engine</td>
</tr>
</tbody>
</table>

* Plug-in electric vehicles (PEVs)

Simulation-based comparisons between ICEVs and BEVs have proven the assumptions from Romm and Frank (2016) (Agrawal et al., 2016; Fiori et al., 2019). Agrawal et al. (2016) used a simulation model to investigate the consumption behavior of the two in different vehicle congestion intensities, showing that BEVs, in contrast to ICEVs, use less energy in lower speeds and, thus, less energy in higher traffic congestion. Wager et al. (2016) confirmed the speed-consumption relationship for BEVs in their work.
Later, these findings could be supported in detail, when Fiori et al. (2019) used different trajectory data as input for different simulation software comparing the energy consumption of EVs and ICEVs. The results showed that low average speed in congested scenarios could lead to reduced consumption for EVs and, hence, have opposite characteristics compared to ICEVs. Figure 4 shows the relationship of speed [km/h] and energy consumption [kWh/mile] for EVs. Compared to the findings focusing on the energy consumption of ICEVs in speed (see Figure 5), BEVs are more efficient at lower speed. While the efficiency of BEVs decreases with higher speed, showed by the convex shape of the curve, ICEs reach an optimum at a higher speed. It should be mentioned that the results of Fiori et al. (2019) about fuel consumption over speed for ICEVs do not show a minimum and seem to be constantly decreasing. However, this minimum fuel consumption is reached at a speed of approximately 37 mph (Haworth & Symmons, 2001; Treiber et al., 2008). Although the findings of Fiori et al. (2019) are solely based on the factor speed, they conclude that EVs use indeed less energy in congested scenarios.

Wu et al. (2015) collected free-field data, comparing energy efficiency on in-city driving versus freeway driving. Their findings support the power-velocity relationship of BEVs as identified by Fiori et al. (2019). Wu et al. (2015) point out that EVs are more efficient on urban routes compared to freeway routes and that a BEV’s energy consumption has a negative correlation with acceleration. However, data for this experiment was only used by a single driver, that was a member of the faculty. Additional research has shown that driver’s aggressiveness has an influence on the energy consumption of a BEV (Badin et al., 2013). Thus, a potential driver specific behavior as well as a road type specific consumption behavior cannot be neglected.
As of today, research analyzing the energy consumption of BEVs based on congestion is few and far between. Previous research provided results based on velocity and speed profiles alone using simulations or small, non-representative samples. Additionally, it can be questioned whether investigating velocity alone, excluding acceleration/deceleration...
and change in acceleration, leads to a valid result that can be transitioned to congestion effects.

There is a gap in the literature investigating the influence of traffic on the energy consumption of BEVs based on a sophisticated and naturalistic experiment using diverse, more representative drivers. It is hypothesized that the route shows significant differences in congestion based on time of the day in a naturalistic driving environment. By testing this hypothesis and thus, addressing the gap in the literature, this research aims to be an important contribution in optimizing eco-driving strategies and understanding BEV efficiency.
3 Methodology

In order to analyze the influence of traffic on the energy consumption of BEVs, data was collected, and various statistical methods used to analyze the data. Within this chapter, the methodology of investigation is described and explained. The chapter begins with a detailed description of the experimental design before the process of data collection will be discussed. Afterward, preliminary calculations will be outlined and finally, the statistical methods chosen for data analysis covered.

3.1 Experimental Design

Experiments held under natural conditions involve multiple, uncontrollable factors. Limited controllability can lead to inaccurate conclusions and should be considered carefully. In order to avoid this and obtain consistent data, a precise, sophisticated experimental design is essential. Nonetheless, total control of all factors is not possible under natural conditions. The impact of not entirely controllable factors was minimized to every extent possible. The chosen experiment will be outlined, covering the following aspects: driver, equipment, test route, test time, and use of auxiliaries. The design and documentation were approved by the University of Rhode Island’s Institutional Review Board (IRB) (HU1617-055).

3.1.1 Test Drivers

Participant recruitment for the experiment was done through flyers and promotional materials at the University of Rhode Island (URI). Flyers were handed out to interested
persons in the URI Memorial Union between 9:00 AM and 11:30 AM and asked directly for their interest in participating. Advertising was also shown on the screens in the Union building and presented at various lectures. Additionally, the flyers were sent out to all graduate students via the URI Grad School. The recruiting material was approved by the IRB and can be found in Appendix A. Interested participants could contact the researchers via e-mail or directly sign up on a list. In order to take part in the experiment, drivers had to fulfill certain criteria: participants had to be at least 18 years old and in possession of a driver’s license, valid for driving in Rhode Island (RI).

Since research has shown that EV driving experience influences the efficiency of a driver (Cocron et al., 2013), it was required that participants had not driven an EV before. In order to arrange the test drives with the participants, a standard e-mail (see Appendix B) was sent out prior to the data collection. People, who were still interested were then scheduled personally via email or phone. A reminder was sent out one day before each scheduled drive (see Appendix C).

Every participant had to drive the route twice, once in the morning with assumed increased traffic congestion and once during low congestion. Participants were asked to drive normally, as they would do as if it was their own vehicle. Thus, they were not specifically dictated to keep the speed limits. All 30 drivers participated on a voluntary basis. The age of the drivers was between 19 and 30 years old with 23.73 years as the average, while 40% of participants were female and 60% male (see Figure 6).

Subjects who were willing to participate as test drivers in the research completed a consent form (see Appendix D) in compliance with the IRB’s rules for experimentation involving human subjects.
3.1.2 Test Equipment

The vehicle used for collecting the data was a 2017 Volkswagen eGolf. A vehicle’s characteristics (e.g., weight, drag coefficient, power) have an influence on the average energy or fuel consumption. Vehicle specifications should not vary over the time of data collection to ensure comparability of data. Thus, the test vehicle was the same for all test drives, keeping the following vehicle specific characteristics constant.

The eGolf has a weight of 1,455 lbs and is equipped with a permanent-magnet synchronous 100 kW electric motor which provides 134 hp (Volkswagen.de, 2019) with a torque output of 214 lbf-ft (Volkswagen.de, 2017). The engine obtains its energy from a 35.8 kWh battery that can achieve a range of 124 miles on average as specified by the manufacturer (Cole, 2016). The drag coefficient \([C_d]\) is given as 0.27. Table 2 summarizes the technical features of the 2017 eGolf that was used for the experiment.
Table 2: Technical features of the 2017 eGolf

<table>
<thead>
<tr>
<th>Component</th>
<th>Specification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mileage (City/Highway)</td>
<td>126/111 (124 on average)</td>
</tr>
<tr>
<td>Weight (lbs)</td>
<td>3,455</td>
</tr>
<tr>
<td>$C_d$</td>
<td>0.27</td>
</tr>
<tr>
<td>Hp</td>
<td>134</td>
</tr>
<tr>
<td>Battery Capacity (kWh)</td>
<td>35.8</td>
</tr>
<tr>
<td>Battery Power (kW)</td>
<td>100</td>
</tr>
<tr>
<td>Recuperation Modes</td>
<td>D1,D2,D3,B</td>
</tr>
</tbody>
</table>

An On-Board Diagnostics (OBD) system was connected to the vehicle’s Electronic Control Unit (ECU) and extracted the vehicle data during the test drives. The used OBD-II for data extraction was provided by Fleetcarma, a technology company headquartered in Waterloo, Canada that provides various solutions to manage and accelerate the transition to EVs (Fleetcarma, 2019). The tracked data included GPS, timestamps, duration, latitude, longitude, altitude, distance, and speed. The data could be exported from the personalized Fleetcarma application and used for data analysis. Additionally, the GPS was tracked using the application “GPS Track” as a backup in order to prevent any data loss. The App was installed on the investigator’s iPhones and run during the drives.

3.1.3 Test Route

The test route used for the data collection had a distance of 27.6 miles in total and covered several road types. Road types can be distinguished through various classification criteria.
The underlying criteria for this research were adopted from the Rhode Island Division of Statewide Planning that prepares and maintains plans for the physical, economic, and social development of the state. Table 3 shows the different road classes and road types covered by the test route and the associated distances and ratios to the total length. The basic functions of the road types are access to property and travel mobility, whereas these functions serve as differentiation criteria between the different road classes and types. Mobility is defined as “the ability of the road to move traffic” (Division of Statewide Planning Rhode Island, 2013). From Freeways to Locals, the different road types can be ranked increasing in accessibility while decreasing in mobility.

Table 3: Road type classification Rhode Island
(Division of Statewide Planning Rhode Island, 2013)

<table>
<thead>
<tr>
<th>Road Class</th>
<th>Road Type</th>
<th>Distance</th>
<th>Road Type Ratio</th>
<th>Road Class Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Principal Arterial</td>
<td>Freeways</td>
<td>5.6</td>
<td>20%</td>
<td>59%</td>
</tr>
<tr>
<td></td>
<td>Principal Arterial</td>
<td>10.6</td>
<td>38%</td>
<td></td>
</tr>
<tr>
<td>Minor Arterial</td>
<td>Minor Arterial</td>
<td>1.9</td>
<td>7%</td>
<td>7%</td>
</tr>
<tr>
<td>Collector</td>
<td>Major Collector</td>
<td>4.2</td>
<td>15%</td>
<td>34%</td>
</tr>
<tr>
<td></td>
<td>Minor Collector</td>
<td>4.5</td>
<td>16%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Local</td>
<td>0.8</td>
<td>3%</td>
<td></td>
</tr>
</tbody>
</table>

Since different road types have an influence on the energy efficiency of BEVs (Bigazzi & Clifton, 2015; Wu et al., 2015), excluding the influence of road types had to be ensured.
The influence of different road types is neutralized through holding the route constant over all conducted test drives and will not be explored further in this work.

The blue line in Figure 7.a depicts the chosen route that covered all different road classes and types except for Interstates. The latter does not exist in the region that the test drives took place. Figure 7.b shows the respective map section with the respective road types.

**Figure 7:** Test route (a) and road classification (b) for South Kingstown, Rhode Island (Division of Statewide Planning Rhode Island, 2013)

All drives started at the URI Fine Arts Parking Lot and headed down Flag Road and Plains Road. After proceeding to Route 138, the route proceeded southbound on Route 2. Route 2 was exited on Shannock Road, proceeding along Worden Pond Road, Ministerial Road, and Route 1 towards Wakefield. From Wakefield, the route continued northbound on Woodruff Avenue, Columbia Street, and North Road. Finally, the route headed on
Saugatucket Road, Curtis Corner Road, and South Road. The experimental route ended at the intersection of Kingstown Road and Upper College Road.

Since the experiment was conducted on natural roads, the altitude varied along the test route. Figure 8 shows the altitude for every point on the test route. The maximum altitude of 79.099 meters above sea level (illustrated by the grey line in Figure 8) was reached at mile 9.05 after the start. The deepest point of the route was reached after 18.23 miles with -3.099 meters below sea level. Since the test route was held constant, the influence of different altitudes on the results canceled each other out. This insured the comparability of the data.

![Figure 8: Altitude along the Test Route](image)
To obtain the data, a total distance of approximately 1,600 miles was driven in total, taking up 60 hours.

3.1.4 Test Time

For the collection of driving data in a naturalistic setting, the time of day for the test drives plays a central role. Rush-hours, or high road traffic conditions, caused by commuters occur at certain times during the day.

In order to analyze the influence of traffic on the efficiency of BEVs, two different test scenarios are necessary: one with increased traffic congestion and one with reduced traffic congestion. The difference between increased traffic and low traffic is not consistent. Zhang and Batterman (2013) pointed out that a strict definition of congestion is not necessary since traffic can be treated as a continuous variable. Thus, daytime is a key variable to control for traffic, as much as possible. The setting of daytime was used to differentiate and compare different traffic situations.

Since previous research was mostly focused on the effect of other factors (e.g., driving behavior, road type, velocity), the influence of traffic was intentionally constant throughout data collection. Works mostly controlled for traffic by varying the daytime of data (Bartels et al., 2019; Jägerbrand & Sjöbergh, 2016; Kowalsky, 2017; Schwertner, 2018). In order to avoid morning rush hours and darkness, Jägerbrand & Sjöbergh (2016) excluded the time period between 11:00 PM and 10:00 AM.

Kowalsky (2017) pointed out that due to the missing availability of traffic data, traffic could not be controlled completely. Thus, his data was not collected in the early morning
hours, where high traffic was assumed. Schwertner (2018) conducted test runs in the period between 10:00 AM and 5:00 PM to avoid peaks in traffic.

The assumptions of the mentioned authors were supported by other investigations focusing on energy usage and range of electric vehicles and plug-in hybrid electric (Gonder et al., 2007; Pearre et al., 2011). The authors of both works state, that commuters leave for work between 5:00-7:45 AM and arrive at work between 7:45-9:00 AM. Additionally, it was found out, that 99% of the investigated BEV fleet was parked during the hours of 12:10-5:50 AM (Pearre et al., 2011). Based on these results it can be summarized that commuters drive to work mainly between 5:00 AM and 9:00 AM (Pearre et al., 2011).

Furthermore, this driver behavioral pattern is supported by research focusing on air pollution due to traffic. Peaks in pollution due to vehicular emissions were found between 7:00 AM and 8:00 AM in the US (Liu, Chen, & Xue, 2017) and 7:30 AM and 8:30 AM in India (Bapna, Sunder Raman, Ramachandran, & Rajesh, 2013). If commuter patterns are similar in different regions, both findings could count for driving behavior in RI. Based on the outlined findings of commuter patterns and air pollution, data with increased traffic was collected between 7:30 AM and 9:00 AM.

Since commuting patterns do not stop and start at a certain time, but weaken and increase slowly in intensity, a gap of one hour was intentionally held between data collection for the settings. Furthermore, daytime correlates with light condition which is significant for the behavior of drivers in combination with weather conditions (Jägerbrand & Sjöbergh, 2016). Hence, data collection during dusk and nighttime were excluded. Therefore, data collection for the reduced congestion scenario was collected between 10:00 AM and 4:00 PM. All drives were conducted during workdays. Because traffic patterns at weekends and
national holidays (Pearre et al., 2011) differ from those at weekdays, these days were excluded from data collection.

Data were collected during the period of eight weeks, starting on April 9, 2019, and ending on June 7, 2019. The relatively short period of time was intentionally chosen because of the significance of different temperatures with regard to the energy consumption of EVs. In cold temperatures, the internal resistance of batteries increases and with this the efficiency, discharge capability, and available energy decreases (Yuksel & Michalek, 2015). The effect of ambient temperature was confirmed in different studies, investigating the energy consumption of PHEVs (Alvarez & Weilenmann, 2012; Fontaras et al., 2017). Although the period of data collection was held as tightly as possible and all drives were conducted in clear and dry weather, the average difference in temperature between the scenarios for each driver was 4.75°F. This average difference is mainly caused by the fact that daytime served as a variable to distinguish between the scenarios. The temperature was measured at the beginning of every drive with a maximum of 73.4 °F and a minimum of 37.4 °F. Figure 9 shows the temperatures measured for each drive.
Figure 9: Temperature at the Beginning of each Test Drive

Test drives were scheduled depending on the availability of the test participants. Therefore, test drives were planned with each participant personally. Drives were observed and accompanied by one of the two investigators, depending on their availability. It was ensured, that both investigators participated nearly in the same amount of drives. Both investigators had the same weight of approximately 161-165 pounds. Hence, the weight of the co-driver was relatively constant for each test run.

In addition to the experimental route, other conditions, such as weather, were held constant to the maximum extent possible. No drives were scheduled on days with forecasted rain because wet roads reduce the skid resistance drastically and result in a reduced coefficient of friction (Singh et al., 2019). The occurred slip, meaning the difference between the wheel velocities of driven and non-driven wheels can lead to higher energy consumption. Additionally, travel speed itself is influenced by light and weather conditions (Jägerbrand & Sjöbergh, 2016; Kowalsky, 2017; Schwertner, 2018). Therefore, in the case of
unexpected rain or wet roads, drives were rescheduled for another day. In the unlikely event of unexpected rain during a test run, data collected for that run had to be excluded from analyses and repeated at another time.

On a single day, a maximum of four test drives was possible. This constraint was caused by a lack of charging station availability and range limitations. Since the eGolf switches automatically in an energy saving mode when a range of 30 miles is reached, the investigators had to ensure that this critical edge was never undershot.

3.1.5 Auxiliary Systems

Auxiliaries of a car include all onboard systems that improve the driving safety and comfort of a car (i.e., air conditioning systems, heating systems, steering assist systems, and other electrical consumers and auxiliaries) (Fontaras et al., 2017). While in use, every system is consuming energy and, accordingly, influencing the efficiency of vehicles, regardless of if the car is currently moving or not.

Badin et al. (2013) investigated the influence of auxiliaries by simulating three different cases of power accessories (i.e., 250W, 500W, and 1000W). Their investigation led to the result that auxiliaries are significant for the efficiency of the BEV, especially in lower speed (Badin et al., 2013). The significance of auxiliary systems was supported in other research (Fontaras et al., 2017). Especially heating, ventilation, and air conditioning (HVAC) systems can reduce the range of BEVs significantly (Kambly & Bradley, 2014). Estimations of fuel consumption caused by the use of A/C range from 3.2 % (Nemry et al., 2008) to 15% (Haworth & Symmons, 2001). In the US, the A/C fuel consumption is
equivalent to 6% of domestic petroleum consumption or 10% of oil imports (Johnson, 2010).

Especially for BEVs, the ratio of a battery’s energy used for A/Cs is significant, since energy from the combustion process is missing compared to ICEs. Thus, waste in the form of heat that could be conducted into the cabin, needs to be produced using additional energy (Yuksel & Michalek, 2015). The intensity of A/C use often depends highly on the ambient temperature which varies regionally (Kambly & Bradley, 2014). For this reason, the settings for the HVAC system was kept constant with an A/C temperature of 68°F and fan power of level 2. Furthermore, settings for dashboard/screen lighting level were the same for each drive. Windows had to be closed during the drives to avoid an increase in aerodynamic drag that results from open windows (Haworth & Symmons, 2001). No additional loads (e.g., children’s car seats, bags, tools) were kept in the car during the drives except for the experimental equipment.

Cruise control can affect the energy economy of vehicles and save energy through increased constancy (Haworth & Symmons, 2001). Thus, cruise control was deactivated at all times. The navigation system was deactivated. Instructions regarding the route were given by the investigators. Prior to the drive, the instructors explained to the drivers that the use of auxiliaries is not allowed or had to remain constantly in the predetermined position during the experiment.

An equal playlist was played on the phone in each drive, making the impact of music equal for all drivers. The phone was connected to the sound system of the car via Bluetooth. The volume of the music was also held constant. Therefore, the phone volume was set to the maximum and the car volume to a level of 5.
A 2017 Volkswagen eGolf has three different recuperation levels, each allowing for different intensities of energy recuperation. The standard driving gear D offers four different brake energy recuperation levels, namely D (brake energy recuperation level 0), D₁ (brake energy recuperation level 1), D₂ (brake energy recuperation level 2), and D₃ (brake energy recuperation level 3). Additionally, the eGolf is equipped with a separate gear B, that offers a very high recuperation where a one-pedal driving (Cocron et al., 2013) is possible because the car decelerates strongly when lifting the foot from the gas (Volkswagen AG, 2017). In order to collect comparable data and exclude the influence of various recuperation intensities, the mode had to be equal for and during all drives. Recuperation level D (brake energy recuperation level 0) was chosen which enables recuperation just during the event of braking. This eGolf driving mode was assumed to be the closest to ICE vehicles and enabled fast adoption of participants.

In addition to the different recuperation levels, the test vehicle offers different driving modes available (i.e., “Normal”, “Eco”, “EcoPlus”). For the test drives, the driving mode was constantly chosen as “Normal” which means that the HVAC runs in normal mode, the full power of the electric motor is available, and the adaptive cruise control is in normal mode. The maximum speed in this mode is 85 mph. In contrast to “Normal”, the driving modes “Eco” and “EcoPlus” include consumption optimizing settings for HVAC, acceleration, breaking, motor power, and maximum speed in order to reduce the energy consumption (Volkswagen AG, 2017).

A sheet was placed in the glove compartment that covered all the information and details about the car settings and driver instructions. This test manual ensured a homogeneous experiment and prevented an unwanted difference in the scenarios of each drive.
3.2 Dependent and Independent Variables

In order to build a model, dependent and independent variables need to be defined. In general, the dependent variable is the result of the independent variables, which are manipulated by the investigators within the possible extend (Lane et al., 1998).

The dependent variable is the average consumption per mile, respectively the difference in the state of charge per mile [$\Delta$SOC/mile]. The $\Delta$SOC/mile is calculated as the used energy per unit distance [kWh/mile] and a continuous variable considered as $y$. The effects of different independent variables on the dependent variable are measured. The independent variables can be seen in Table 4. The scenario based on daytime as the mainly investigated variable is binary (0 for low congestion and 1 for increased congestion). As outlined previously, individual driving behavior has an impact on the energy consumption of electric vehicles and therefore, cannot be neglected. The factor driver is a categorical variable for all 30 drivers.

Since vast literature has shown the significance of temperature on the energy consumption of BEVs, this variable was considered as a continuous variable. Additionally, the initial state of charge was considered for investigations and treated as a continuous variable as well.

Table 4: Factors and corresponding types of statistical data

<table>
<thead>
<tr>
<th>Factor</th>
<th>Scenario</th>
<th>Driver</th>
<th>Temperature</th>
<th>Initial SOC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
<td>Categorical</td>
<td>Categorical</td>
<td>Continuous</td>
<td>Continuous</td>
</tr>
</tbody>
</table>
3.3 Methods of Data Collection and Handling

Multiple steps are necessary to get valid conclusions out of the driving raw data. Within this chapter, the raw data as extracted through the OBD II device will be described. Afterward, preliminary calculations for the transition of raw data to target figures (e.g., ΔSOC/mile, distance per drive) will be outlined. Additionally, applied methods for data cleaning will be explained.

3.3.1 BEV Raw Driving Data

Driving data during the test runs were extracted by a device provided by fleetcarma that was connected to the OBD II port of the research vehicle. The OBD is a male plug that can be used to connect computers or devices for system analysis or data extraction. Extracted data were automatically uploaded to the Fleetcarma server. The collected data (e.g., GPS, state of charge, timestamp) could be accessed with a Fleetcarma account and downloaded as a file from the Fleetcarma website. The raw data could be opened in Microsoft (MS) Excel for further calculations. Figure 10 illustrates the method of data collection from the collection of driving data to the final results.
Table 5 shows a sample extract of the raw data as downloaded from the Fleetcarma website. The vehicle data was measured every second and contains the following information: timestamp [ms], altitude [m], battery current [A], battery SOC [%], battery voltage [V], driving status [binary], latitude [º], longitude [º], outside air temperature (OAT) [ºC], vehicle Speed [km/h]. Additionally, each data sheet contains information about the starting date and daytime of the drive given in universal coordinated time (UTC). The time between starting the vehicle and data tracking through the device varied. In the shown example the first assessment happened after 1,400 ms and was then repeated every second until the ignition was switched off.
Table 5: Sample of raw data as downloaded from the cloud service of Fleetcarma

<table>
<thead>
<tr>
<th>Time</th>
<th>Altitude</th>
<th>Battery Current</th>
<th>Battery SOC</th>
<th>Battery Voltage</th>
<th>Driving</th>
<th>Latitude</th>
<th>Longitude</th>
<th>OAT</th>
<th>Speed</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>400</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1400</td>
<td>85.9</td>
<td>-4.5</td>
<td>79</td>
<td>343.5</td>
<td>1</td>
<td>41.48928278</td>
<td>71.52183028</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>2400</td>
<td>85.9</td>
<td>-4.5</td>
<td>79</td>
<td>343.5</td>
<td>1</td>
<td>41.48928278</td>
<td>71.52183028</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>3400</td>
<td>85.9</td>
<td>-4.5</td>
<td>79</td>
<td>343.5</td>
<td>1</td>
<td>41.48928278</td>
<td>71.52183028</td>
<td>10</td>
<td>0</td>
</tr>
</tbody>
</table>

3.3.2 Preliminary Calculations

The target result of each drive was the energy consumption per mile. The difference in the state of charge per mile driven $[\Delta SOC/mile]$ is a commonly used measure for calculating the consumption of a BEV (Bartels et al., 2019; Bingham et al., 2012) and was used for this investigation. In order to get to the target measure, the total energy consumed divided by the total distance had to be calculated for each test drive.
As obtained in the raw data, the SOC was given as a percentage of the charge capacity in every second. By multiplying with the total battery capacity of 35.8 kWh, the percentage could be converted into kWh.

**Equation 1: Transformation equation from SOC in % to SOC in kWh**

\[
SOC \ [kWh] = SOC \ [%] \times 35.8 \ [kWh]
\]

The raw data contained the SOC in 0.5% steps, thus, a difference in SOC \(\Delta SOC\) could not directly be calculated for every second. Several steps were necessary to get to the SOC per second by using interpolation. Instead, the application of the power-law was the promising alternative to calculate the electrical power \(P\). The power-law describes the relationship between electrical power [kW], voltage [V], and current [A]. It is defined by the following equation:

**Equation 2: Power-law**

\[
P \ [kW] = \frac{V \ [V] \times I \ [A]}{1000}
\]

with:

- \(P\) = Power in kW
- \(V\) = Voltage in V
- \(I\) = Amperage in A
In order to attain the electrical energy (E) used while driving from the power, another transformation was necessary. Electrical energy [kWh] can be calculated by multiplying the electrical power used with the total time the energy is used in hours:

Equation 3: Electrical energy formula

\[ E \text{ [kWh]} = P \text{ [kW]} \times t \text{ [h]} \]

with:
- \( E \) = Energy in kWh
- \( P \) = Power in kW
- \( t \) = time in h

In this case, the electrical energy had to be calculated per second, since current and voltage were given for every second. The following equation was used to get the energy in kWh for every entry of the table:

Equation 4: Electrical energy for every second measured

\[ \frac{E \text{ [kWh]}}{s} = P \text{ [kW]} \times \frac{1}{3600} \]

Furthermore, the total electrical energy used over the whole test drive was calculated as the sum of electrical energy used per second of the drive:

Equation 5: Total electrical energy used in a drive

\[ E_{\text{total}} \text{ [kWh]} = \sum \frac{E \text{ [kWh]}}{s} \]

After calculating the electrical energy used during the experiment, the exact driven distance was needed to calculate the target figure \( \Delta \text{SOC/mile} \). The Haversine formula seemed to be
a promising method to determine the mileage distance between two points from latitude and longitude, considering the curvature of the earth (Cai & Guiffrida, 2016). However, the GPS coordinates were often given equally for several seconds although the car was moving. This inaccuracy in tracking the vehicle resulted in several rows with a traveled distance of 0 miles and was a limitation. As this was evidently not representing the real driving data, the traveled distance between two coordinates had to be calculated in a different, more accurate way.

Since speed was measured for every second, from speed [km/h] could be transformed to traveled distance [miles] to obtain the traveled distance between two timestamps. A vehicle’s velocity in km/h can be transformed to the velocity in mph by multiplying it with the factor 0.6213711922. Additionally, a transformation was necessary to obtain the traveled distance for every second by dividing by 3600 [s/h]. Equation 6 shows the transformation necessary to obtain the traveled distance in miles for every data point.

Equation 6: Transformation from velocity [km/h] into distance [miles]

\[
d[miles per second] = \frac{v[km/h] \times 0.6213711922[miles/km]}{3600[s/h]}
\]

with:

\[d\] = Distance in miles
\[v\] = velocity in [km/h]

The total distance of each drive could now be calculated by summing up the traveled distances of every second (see Equation 7).
Equation 7: Calculation of total distance driven

\[ d_{\text{total}} \text{[miles]} = \sum \frac{d \text{[miles]}}{s} \]

By dividing the total electrical energy used by the total distance driven, the average energy consumption as the target figure could be calculated for every test drive.

Equation 8: Calculation of average energy consumption for every test drive

\[ \frac{\Delta SOC}{\text{mile}} \text{[kWh/mile]} = \frac{E_{\text{total}} \text{[kWh]}}{d_{\text{total}} \text{[miles]}} \]

with:

\[ \Delta SOC/\text{mile} = \text{difference in SOC per mile} \]

In addition to the average consumption, different metrics were calculated and later used for providing evidence of differences in traffic intensities between the two test scenarios. These metrics consisted of the total travel time, the average speed, the standard deviation of speed, the standard deviation of acceleration as the derivative of speed, and the standard deviation of jerk as the derivative of acceleration were calculated for each drive. Jerk, especially, as defined as the variation of speed per variation of time can be used to measure traffic intensity.

3.3.3 Data Cleaning

In order to have ensured comparability of different data sets, every drive required equal start and ending locations and, thus, an equal driven distance. Therefore, drive data needed to be cleaned. In order to do so, starting and ending points were determined in advance
using Google Maps (see Figure 11) with the starting coordinates 41.489207 latitude and -71.521673 longitude and ending coordinates 41.480337 latitude and -71.525495 longitude.

Additionally, road work on Worden Pond Road started unexpectedly during the period of data collection. The respective stretch began with the first sign that announced road works and ended with the end of the construction site itself. Thus, a total distance of 2.1 miles in total needed to be excluded from data analysis. The coordinates were detected using Google Maps and began at 41.432126 latitude and -71.607242 longitude and ended at 41.429276 latitude and -71.568301 longitude (see Figure 12).
Figure 12: Stretch of the route with construction

The road work influenced the speed profile in this section significantly. Figure 13.a and Figure 13.b show the speed profiles for the respective stretch as recorded by “GPS Track”. Green colored parts reflect higher speed than yellow and red colored sections. The speed profiles show that velocities before and during road works differed significantly.

Figure 13: Speed profile a) before and b) during construction

Figure 14 supports the mentioned speed differences by showing the different speeds at any point of the respective road stretch. The blue line shows the speeds before construction and
the orange line the speeds during construction. The averages showed a significant difference with 27.59 mph (before construction) and 16.09 mph (during construction).

Figure 14: Comparison of Speeds on route stretch with construction

The respective stretch had to be eliminated from investigations. The method used for detecting the cutting points in the data files was similar for every marker. Although longitude and latitude in the raw data included 13 digits after decimal point, a full match with the coordinates of the cutting markers was rarely achieved due to the infrequently given GPS coordinates in the raw data. Thus, a calculation similar to the Pythagorean Theorem was used to detect the minimum possible distance between the four different cutting coordinates and the point of the route. In order to do so, the squared difference between the longitude of the data point and the longitude of the cutting marker was calculated. The same was done with the latitude of each data point and the latitude of the
cutting marker. Both values were summed up for each entry, building the detecting value.

Equation 9 shows the formula used for detecting the cutting rows.

*Equation 9: Target function for detecting cutting points*

\[ \min d_{rc}^2 = (\varphi_r - \varphi_c)^2 + (\lambda_r - \lambda_c)^2 \]

with:

- \( d_{rc} = \text{distance measure between route coordinate in a second and cutting point} \)
- \( \varphi_r, \varphi_c = \text{the latitudes of a point on route (r) and cutting marker (c)} \)
- \( \lambda_r, \lambda_c = \text{the longitudes of a point on route (r) and cutting marker (c)} \).

The data was cut by deleting the rows before the starting point, after the ending point, and between the start and end markers of the road construction. Table 6 illustrates the cutting procedure. The red lines indicate where a cut is made to delete the grey filled rows that had to be excluded in order to obtain homogeneous data.
Table 6: Cutting procedure for test runs

<table>
<thead>
<tr>
<th>Timestamp (ms)</th>
<th>Latitude[deg]</th>
<th>Longitude[deg]</th>
<th>START CUT</th>
<th>END CUT</th>
<th>CON. START</th>
<th>CON. END</th>
</tr>
</thead>
<tbody>
<tr>
<td>20800</td>
<td>41.4891775</td>
<td>-71.521872</td>
<td>0.00022844</td>
<td>0.01246356</td>
<td>0.14242156</td>
<td>0.10633056</td>
</tr>
<tr>
<td>21800</td>
<td>41.4892631</td>
<td>-71.521674</td>
<td>5.7222E-05</td>
<td>0.01274689</td>
<td>0.14270489</td>
<td>0.10661389</td>
</tr>
<tr>
<td>993800</td>
<td>41.4325636</td>
<td>-71.607338</td>
<td>0.14230789</td>
<td>0.12961589</td>
<td>0.00053311</td>
<td>0.04232411</td>
</tr>
<tr>
<td>994800</td>
<td>41.4322214</td>
<td>-71.607216</td>
<td>0.14252872</td>
<td>0.12983672</td>
<td>0.00012612</td>
<td>0.04196051</td>
</tr>
<tr>
<td>1205800</td>
<td>41.4292828</td>
<td>-71.568063</td>
<td>0.10631372</td>
<td>0.09362172</td>
<td>0.04202272</td>
<td>0.00024528</td>
</tr>
<tr>
<td>1206800</td>
<td>41.42896</td>
<td>-71.566228</td>
<td>0.10480233</td>
<td>0.09211033</td>
<td>0.04417967</td>
<td>0.00238967</td>
</tr>
<tr>
<td>2731800</td>
<td>41.48041</td>
<td>-71.52537</td>
<td>0.012494</td>
<td>0.000198</td>
<td>0.130156</td>
<td>0.094065</td>
</tr>
<tr>
<td>2732800</td>
<td>41.4811111</td>
<td>-71.525482</td>
<td>0.01190456</td>
<td>0.00078744</td>
<td>0.13074544</td>
<td>0.09465444</td>
</tr>
</tbody>
</table>

The net distance of the route covered a distance between 24.3 and 24.4 miles in total. Little deviations occurred due to inaccuracies in the GPS and individual driving behavior.
All calculations and the data cleaning was automatically conducted using a “Visual Basic for Applications” (VBA) in MS Excel (see Appendix E). Additionally, a second code was programmed and used to create a final file with the summarized data of all 60 test drives (see Appendix G).

### 3.4 Statistical Analyses

The two scenarios were tested for differences in congestion intensity based on different traffic metrics. For simplification reasons, the words traffic and high congestion will be used synonymously to describe the scenario “Traffic” in the hours between 10:00 AM and 04:00 PM. The same holds for low traffic and decreased congestion accordingly for the scenario “No Traffic” in the morning hours. Furthermore, different factors were investigated for their significance on the average consumption of BEVs.

#### 3.4.1 Analyses for evidence of traffic

Since the intensity of traffic congestion is a function of speed and flow, the latter can be used to analyze the different test scenarios for traffic (Barth & Boriboonsomsin, 2008). Because there was no access to real-time traffic data for the chosen test route, measuring the flow for proof a difference in congestion intensity based on daytime was not possible. Alternative investigations using the available driving data were necessary.

In order to provide evidence of differences between the two test scenarios and analyze the stated hypothesis that the test route shows differences in traffic based on daytime, the
paired sample t-test was used to compare the data sets. The paired, two-sided t-test was chosen because both data subsets included data of the same 30 drivers and thus had to be considered as “paired” (Montgomery, 2013).

Testing for equal differences leads to the following hypotheses:

$$H_0: \mu_d = 0$$

$$H_1: \mu_d \neq 0$$

Some considered traffic measures were used in form of standard deviations. In order to apply the paired, two-sided t-test, the mean of the standard deviations of these measures over all drives had to be calculated.

Based on the significance level \(\alpha\), the null hypothesis can be rejected if \(p < \alpha\). The p-value is defined as the smallest level of significance that could lead to a rejection of \(H_0\) (Montgomery, 2013). Thus, a higher significance level can increase the likeliness of rejecting \(H_0\), although the hypothesis is true (Type I error) (Kim, 2015). The underlying significance level depends on the field of investigation and should be chosen under consideration of its decreasing function of sample size (Leamer, 1978). Following Leamer (1978), a total of 60 drives is a relatively small sample size and could justify the choice of a high significance level (\(p = 0.1\)). Nonetheless, it should be aimed to reduce the probability of a Type I error. Based on the relatively small sample size, hypothesis tests were performed based on a significance level of \(\alpha = 0.1\).

Various measures were compared for differences in mean using the two-sided paired t-test. All t-tests were conducted using R. Initially, the averages in speed were compared since
speed is part of the function of traffic. Because speed depends highly on road types (Division of Statewide Planning, 2018) an influence of the latter on the speed profile cannot be neglected. Therefore, both data sets were also analyzed for differences in the standard deviation of acceleration and the standard deviation of jerk. As a derivative of acceleration, jerk represents the changes in acceleration and has a high effect on traffic movement and vice versa (Ge, Zheng, Wang, & Cheng, 2015). Thus, the variability of jerk can serve as a measure to proof traffic for chosen test routes based on the available driving data retrospectively. Jerk can be calculated as follows:

\textit{Equation 10: Calculation of jerk}

\[ \text{Jerk} = \frac{\Delta \text{acc}}{\Delta t} \]

with:
\[ \text{acc} = \text{acceleration} \ [m/s^2] \]
\[ t = \text{time} \ [s] \]

A paired t-test assumes that differences between two data sets are normally distributed. Therefore, all tested measures had to be analyzed for normality prior to testing for differences in mean. Normality was tested both, graphically and mathematically. First, boxplots and normality plots for every considered traffic measure were plotted and investigated for outliers and abnormalities. Outliers are treated as such if they undercut the “minimum” or exceed the “maximum”. The “minimum” is determined by the first quartile minus 1.5 times the interquartile range (IQR). The “maximum” instead is calculated by adding 1.5 times the IQR to the value of the third quartile. Then, the Shapiro-Wilk-test was applied for the mathematical testing of normality. The test was chosen since the sample
size was relatively small (Shapiro & Wilk, 1965). The following hypotheses were tested using R:

\[ H_0: \text{The sample comes from a normally distributed population} \]

\[ H_1: \text{The sample does not come from a normally distributed population} \]

3.4.2 Analysis for difference in energy consumption for the two scenarios

After testing for differences in traffic metrics between the two test scenarios and, thus, possible differences in traffic congestion on the chosen test route, the average energy consumption was tested for differences. Again, the mean consumption per mile for both scenarios had to be tested for normality and outliers prior to the t-test. The procedure for normality testing was equal to the procedure applied for the different traffic metrics. After proving normality, the two-sided paired sample t-test was applied. All tests were conducted using R.

3.4.3 Multiple linear regression

Regression models are commonly used as a multivariate analysis technique to explain the relationship between a dependent variable \( (Y) \) and one or more independent variables \( (X) \). The choice of the right regression model depends highly on the assumptions that the different models meet. A multiple linear regression assumes linearity, reliability of measurement \( (R^2) \), homoscedasticity, and normality (Lane et al., 1998; Osborne & Waters, 2001). The response \( (Y) \) may be related to \( k \) regressor variables \( (X_j) \). The regression coefficients \( \beta_i \) display the expected change of the response variable per unit change \( X_j \).
(Montgomery 2013). Since models can never explain the response entirely, the error [$\varepsilon$] describes the remaining proportion of the response (Equation 11).

*Equation 11: Linear multiple regression equation*

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k + \varepsilon$$

with:

- $y = \text{Response}$
- $\beta_i = \text{Regression coefficient of factor } i$
- $X_i = \text{Value for factor } i$
- $\varepsilon = \text{Model error}$

Within this investigation, the impact of the four different dependent variables (i.e., scenario, driver, temperature, and initial SOC) on the average $\Delta$SOC/mile of BEVs is analyzed using the multiple linear regression. Since the experiment was conducted in a naturalistic setting, a statistical design of experiment analysis was not possible.

Since the coefficient driver was categorical with 30 different drivers, one driver was used as a reference. The referenced driver was chosen by the sum of deviations from the averages for both test drives. The smallest variation was detected for driver 30. Figure 15 shows the average consumption per driver and scenario and the overall averages for the scenario “Traffic” (gray line) and “No Traffic” (yellow line).
Figure 15: ΔSOC/mile for every driver and scenario

The remaining 29 drivers were treated as binary dummy variables (0 = Driver is not driving, 1 = Driver is driving). The binarity of the factors traffic and driver allowed the use of a multiple linear regression although the factors are non-numeric. The same approach was applied for the factor traffic. While the “No Traffic”-scenario served as a reference, “Traffic” was analyzed as a coefficient.

A linear model was calculated and tested for the reliability of measurement using the coefficient of determination (R²). R² is a measure of linear regression and measures how much of the total variability is explained by the model. While a low value for R² indicates a “poor model fit” and the chosen variables are not suited to predict the dependent variable, a high value for R² close to 1 means that the independent variables are well-suited to predict the dependent variables (“good model fit”). Because R² increases with the number of added
coefficients, for regression analyses with multiple coefficients, the adjusted value for $R^2$ is used as an indicator for the model fit (Montgomery, 2013). In order to test for the model for homoscedasticity, the residuals were plotted against fitted values of the model. A model can be considered as homoscedastic if the residuals appear to be homogeneously distributed among the fitted values and the red line is approximately horizontal. Additionally, the scale-location plot was used to confirm homoscedasticity. The scale-location plots the standardized residuals instead of the residuals and simplifies the homoscedasticity test. Again, the values appear homoscedastic if they are homogeneously spread around an approximately horizontal red line.

In order to have a quantifiable comparison between different driver types, a multiple linear regression was applied with two further drivers as a reference. The results of driver 30 as a reference were compared to the results for driver 4 (“most aggressive”) and driver 9 (“least aggressive”). The terms “most aggressive” and “least aggressive” are used based on the average consumption of the drivers in both scenarios. Driver 4 had the highest $\Delta$SOC/mile among all drivers, while driver 9 had the lowest.

The regression analysis was conducted in R and covered an analysis of variance (ANOVA). Based on the ANOVA, differences between two of more means can be tested by analyzing their variances (Lane et al., 1998). Based on the p-values as given in the ANOVA table within the linear regression, significant factors for the response could be detected.
3.4.4 Methods for improving the model fit

In the case of a relatively low model fit (R² adjusted below 60%), a stepwise regression was conducted and used to find a model with a better fit. A stepwise regression contains the procedures of forward selection and backward elimination. Both processes follow different steps to form a model that takes significant factors into account and has increased reliability compared to the initial model. This means that not all possible factors are necessarily taken into account. The forward selection and the backward elimination differ in their approaches while aiming for a better model. The forward selection adds different factors stepwise one at a time based on their significance. In each step, the most significant factor with a p-value below the significance level α is added to the model. Within each step, the remaining set of variables is considered for being included in the model (Derksen & Keselman, 1992). The procedure ends, when no additional significant factors can be detected. The backward elimination instead starts with the initial model that considers all factors for inclusion. In each step, the most insignificant factor with p > α is eliminated from the model. The backward selection stops when no insignificant factors are left (Derksen & Keselman, 1992). For both procedures, the new, potentially partial model had to be tested for homoscedasticity and compared to the initial model based on the values for R². The stepwise regression was applied considering an α of 0.1 to enter or remain in the model.
4 Results and Discussion

4.1 Evidence of Traffic

In order to investigate the impact of traffic on the energy consumption of BEVs and obtain an answer to the stated research question, it was essential to provide evidence that the chosen test route shows a significant difference in traffic congestion based on daytime. After testing for normality, different traffic metrics were analyzed for differences in means between the two test scenarios over the 30 test drivers.

Figure 16 shows the different normality plots as they were used for graphical normality analyses for the considered traffic measures (i.e., average speed, average duration, standard deviation of speed, standard deviation of acceleration, and standard deviation of jerk) and test scenarios (i.e., “traffic” and “no traffic”). The closer the different data points are to a theoretical normality line, the more likely the measures are normally distributed. The normality plots for average speed and all standard deviation measures follow the line and, thus, seem to be normally distributed. Solely the normality graph for the driving duration, especially in an increased congestion scenario shows some deviations since the data points are less close to the normality line and should further be investigated carefully.
<table>
<thead>
<tr>
<th>Measure Scenario</th>
<th>Traffic</th>
<th>No Traffic</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Average Speed</strong></td>
<td><img src="image1" alt="Graph" /></td>
<td><img src="image2" alt="Graph" /></td>
</tr>
<tr>
<td><strong>Average Duration</strong></td>
<td><img src="image3" alt="Graph" /></td>
<td><img src="image4" alt="Graph" /></td>
</tr>
<tr>
<td><strong>Standard Deviation Speed</strong></td>
<td><img src="image5" alt="Graph" /></td>
<td><img src="image6" alt="Graph" /></td>
</tr>
<tr>
<td><strong>Standard Deviation Acceleration</strong></td>
<td><img src="image7" alt="Graph" /></td>
<td><img src="image8" alt="Graph" /></td>
</tr>
</tbody>
</table>
Figure 16: Normality plots for different traffic measures

Figure 17 displays boxplots for the different obtained traffic measures. The left boxplots illustrate the scenario “Traffic” and the plots on the right-hand side display the setting “No Traffic”. Except for the average duration, none of the other measures shows outliers. Outliers are illustrated by single data points that are not in between the minimum and maximum.

The boxplot for duration shows outliers that lie above the “maximum” and confirms the findings of the graphical analysis of the normality plots. In addition to serving as a graphical test for outliers, boxplots can be used as a first indicator for equal medians. While the medians, illustrated by the thick black lines in the interquartile range, do not seem to be significantly different for the two scenarios for average speed, average duration, and standard deviation of speed, their difference could be significant for the standard deviation of acceleration and standard deviation of jerk. The significance of the differences between the two test scenarios will be investigated later through two-sided paired t-tests for all measures separately.
<table>
<thead>
<tr>
<th>Measure</th>
<th>Boxplot</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Speed</td>
<td><img src="image" alt="Boxplot" /></td>
</tr>
<tr>
<td>Average Duration</td>
<td><img src="image" alt="Boxplot" /></td>
</tr>
<tr>
<td>Standard Deviation Speed</td>
<td><img src="image" alt="Boxplot" /></td>
</tr>
<tr>
<td>Standard Deviation Acceleration</td>
<td><img src="image" alt="Boxplot" /></td>
</tr>
<tr>
<td>Standard Deviation Jerk</td>
<td><img src="image" alt="Boxplot" /></td>
</tr>
</tbody>
</table>

Figure 17: Boxplots for different traffic measures
The Shapiro-Wilk-test for the mathematical normality investigation was applied after the graphical analyses. The results are displayed in Table 7. A significance level of \( \alpha = 0.1 \) was taken as a basis for the hypothesis tests which approve the findings from the graphical tests.

For every considered traffic measures, the null hypothesis cannot be rejected. Abnormalities found for the measure duration during the graphical analysis (Figure 16) are not confirmed by the Shapiro-Wilk-test. Therefore, all measures are distributed normally.

Table 7: Results of the Shapiro-Wilk-Normality-Tests for traffic measures

<table>
<thead>
<tr>
<th>Measure</th>
<th>Average Speed</th>
<th>Average Duration</th>
<th>Std. Dev. Speed</th>
<th>Std. Dev. Acceleration</th>
<th>Std. Dev. Jerk</th>
</tr>
</thead>
<tbody>
<tr>
<td>p-Value (Traffic)</td>
<td>0.5031</td>
<td>0.09572</td>
<td>0.6332</td>
<td>0.8434</td>
<td>0.4634</td>
</tr>
<tr>
<td>p-Value (No Traffic)</td>
<td>0.2094</td>
<td>0.08707</td>
<td>0.4629</td>
<td>0.4124</td>
<td>0.1399</td>
</tr>
</tbody>
</table>

Note: \( p \)-value \( < 0.001 \) ‘***’ \( < 0.01 \) ‘**’ \( < 0.05 \) ‘*’ \( < 0.1 \) ‘.’

After the evidence for normality was confirmed by the normality tests, the scenarios could be tested for differences between the two scenarios. The results of the two-sided paired t-test for all measures can be found in Table 8. A significance level of \( \alpha = 0.1 \) was considered. The results show that the scenarios “Traffic” and “No Traffic” are significantly different for variation in acceleration \( (p = 0.0565) \) and the derivative of acceleration, jerk \( (p = 0.0609) \). This confirms the findings from the graphical boxplot analysis that showed differences in medians. Only the average standard deviations of acceleration and jerk proof that there is a difference in traffic intensity between different daytimes based on the driving data. Although real-time traffic data were not available, the findings qualify the test route.
to be used as a route for data collection for different traffic intensities based on daytime. Still, big differences in traffic would lead to differences in driving duration and average speed. The test route did not show enough variation in commuting patterns to include differences for these measures.

Table 8: Two-sided paired t-test results for different traffic measures

<table>
<thead>
<tr>
<th>Measure</th>
<th>Average Speed (mph)</th>
<th>Average Duration (sec)</th>
<th>Std. Dev. Speed (mph)</th>
<th>Std. Dev. Acceleration (m/s²)</th>
<th>Std. Dev. Jerk (m/s³)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (Traffic)</td>
<td>35.3823</td>
<td>2,510.633</td>
<td>13.9253</td>
<td>0.7820</td>
<td>0.6078</td>
</tr>
<tr>
<td>Mean (No Traffic)</td>
<td>35.0513</td>
<td>2,531.7000</td>
<td>13.8417</td>
<td>0.7594</td>
<td>0.5794</td>
</tr>
<tr>
<td>p-Value (two-sided)</td>
<td>0.4357</td>
<td>0.48771</td>
<td>0.4998</td>
<td>0.0565 *</td>
<td>0.0609 *</td>
</tr>
</tbody>
</table>

Note: p-value < 0.001*** p < 0.01 ** p < 0.05 * p < 0.1

Thus, the differences in the energy consumption of BEVs based on the impact of traffic could be investigated further.

4.2 Analysis of Impact of Traffic on the Energy Consumption of BEVs

Before an actual comparison for differences in means of energy consumption for the two data sets could be conducted, both sets had to be tested for normality.

The normality plots of average consumption for the settings “traffic” and “no traffic” are displayed in Figure 18.a respectively Figure 18.b. While the data points for the first scenario follow the line accurately, the data points for the subset reduced congestion show
deviations from the normality line in the lower and upper quantiles. This is a first indicator for a lower p-value for the Shapiro-Wilk test.

Figure 18: Normality plots of average consumption in scenarios a) traffic b) no traffic

Figure 19: Boxplots of average energy consumption for both scenarios

The boxplots of the average energy consumption for both settings are shown in Figure 19. No outliers can be detected. However, both, the interquartile ranges and the medians differ
significantly with a lower median for the setting no traffic. This lower median can serve as a first indicator for different means for the subsets.

The Shapiro-Wilks-test for normality leads to the conclusion that the rejection of the null hypothesis failed. Although the p-value for the setting “no traffic” is lower as indicated previously by the graphical normality test, both subsets are normally distributed for all possible significance levels. This qualifies the data sets for an application of the paired t-test to test for differences in means.

*Table 9: Results of the Shapiro-Wilks-Normality-Test for average consumption*

<table>
<thead>
<tr>
<th>Scenario</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traffic</td>
<td>0.9491</td>
</tr>
<tr>
<td>No Traffic</td>
<td>0.2456</td>
</tr>
</tbody>
</table>

Note: p-value < 0.001: *** p < 0.01: ** p < 0.05: * p < 0.1: '

The histogram for the average consumption over all drives are displayed in Figure 20. It shows that the occurrences of the responses are distributed normally for the entire sample.
Figure 20: Histogram SOC/mile

The averages in energy consumption for all drivers in both settings are illustrated in Figure 21. The blue markers indicate average consumptions [kWh/mile] in “Traffic”, while orange markers indicate average consumptions for different drivers in the scenario “No Traffic”.
Figure 21: Average energy consumption over driver and scenario

The plotted consumptions indicate that there are differences in the average consumption for every driver between the two settings with a tendency to higher consumption in higher congestion. The visually detected difference in average consumption between the two scenarios could be confirmed by the two-sided paired t-test. The results of the paired sample t-test are summarized in Table 10. The average consumption over all drives was higher in the traffic setting (0.2342 kWh/mile) compared to the setting with reduced consumption (0.218 kWh/mile). Additionally, $p < \alpha$ counts for the significance levels 0.001, 0.05, and 0.1 and the null hypothesis can be rejected for all considered significance levels. It can be stated that the difference in means for average energy consumption between the two scenarios is highly significant.
Based on this, BEVs show a similar consumption behavior in traffic congestion to ICEs (Fiori et al., 2019; Jereb et al., 2018; Treiber et al., 2008). However, the impact of traffic on BEVs is significantly smaller.

**Table 10: Two-sided paired t-test results for different average consumptions**

<table>
<thead>
<tr>
<th>Setting</th>
<th>Traffic</th>
<th>No Traffic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.23422</td>
<td>0.21802</td>
</tr>
<tr>
<td>p-Value</td>
<td>0.00037428***</td>
<td></td>
</tr>
</tbody>
</table>

Note: *p-value < 0.001*** p < 0.01 ** p < 0.05 * p < 0.1’’

With evidence of differences in the means of consumption between the two test settings, factors that impact the response (ΔSOC/mile) significantly needed to be detected. As shown earlier, there was a high average temperature difference between the two scenarios that could solely explain the differences in energy consumption.

A multiple linear regression including all considered factors (i.e., Scenario, driver, temperature, and initial SOC) was applied to test the different factors for their impact on the predictor difference in SOC per mile. Based on the ANOVA table, significant factors could be identified.

Table 11 shows the results of the ANOVA for the full model, considering the coefficients traffic (X₁), driver (X₂), outside temperature (X₃) and initial SOC (X₄).

The overall model under consideration of all four factors has coefficient of determination (R² adjusted) of 53.23%. Although this model fit cannot be considered as high, data
collection in a naturalistic setting with multiple uncontrollable factors makes a high model fit unlikely.

The results show that the intercept respectively the constant ($\beta_0$) has the highest estimator with 0.2678 and is significant ($p = 0.0000$). The regression coefficient of the constant indicates what the basic consumption of a BEV in a scenario without traffic on the chosen test route is. The only further significant factors considering a significance level of $\alpha = 0.1$ for this initial model is temperature ($p = 0.0031$) with a contribution of -0.0012 and driver 25 ($p = 0.0238$). An increase in temperature of 10ºF would lead to a decrease in the average BEV energy consumption of 0.0122 kWh/mile and shows its high significance.

Traffic has a positive value for the regression coefficient (0.0068) and indicates that the average consumption in a setting with congestion would be 0.0068 kWh/mile or 2.5244% higher compared to the “No Traffic” scenario. However, the factor is not considered to be significant for the response ($p = 0.12659$). The contribution of the factor “Initial SOC” is negligibly small ($\beta_4 = 0.0002$) and does not show significance ($p = 0.1173$). Considering all factors for a multiple linear regression model, traffic cannot be identified as having a significant impact on the average energy consumption of BEVs.

The plot for residuals vs. fit (Figure 22) does not show any abnormalities. Residuals seem to occur randomly over the different fitted values making the model homoscedastic.
Table 11: ANOVA Table for full Model

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t-Value</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>0.2847</td>
<td>0.0264</td>
<td>10.7650</td>
<td>0.0000***</td>
</tr>
<tr>
<td>Traffic</td>
<td>0.0068</td>
<td>0.0043</td>
<td>1.5760</td>
<td>0.1266</td>
</tr>
<tr>
<td>Driver 1</td>
<td>-0.0170</td>
<td>0.0127</td>
<td>-1.3350</td>
<td>0.1931</td>
</tr>
<tr>
<td>Driver 2</td>
<td>-0.0116</td>
<td>0.0132</td>
<td>-0.8760</td>
<td>0.3886</td>
</tr>
<tr>
<td>Driver 3</td>
<td>-0.0182</td>
<td>0.0136</td>
<td>-1.3420</td>
<td>0.1909</td>
</tr>
<tr>
<td>Driver 4</td>
<td>0.0021</td>
<td>0.0131</td>
<td>0.1610</td>
<td>0.8734</td>
</tr>
<tr>
<td>Driver 5</td>
<td>-0.0090</td>
<td>0.0124</td>
<td>-0.7270</td>
<td>0.4737</td>
</tr>
<tr>
<td>Driver 6</td>
<td>0.0018</td>
<td>0.0122</td>
<td>0.1510</td>
<td>0.8814</td>
</tr>
<tr>
<td>Driver 7</td>
<td>-0.0120</td>
<td>0.0119</td>
<td>-1.0090</td>
<td>0.3219</td>
</tr>
<tr>
<td>Driver 8</td>
<td>-0.0125</td>
<td>0.0127</td>
<td>-0.9850</td>
<td>0.3334</td>
</tr>
<tr>
<td>Driver 9</td>
<td>-0.0270</td>
<td>0.0124</td>
<td>-2.1780</td>
<td>0.0383</td>
</tr>
<tr>
<td>Driver 10</td>
<td>0.0125</td>
<td>0.0122</td>
<td>-1.0250</td>
<td>0.3146</td>
</tr>
<tr>
<td>Driver 11</td>
<td>-0.0090</td>
<td>0.0114</td>
<td>-0.7870</td>
<td>0.4379</td>
</tr>
<tr>
<td>Driver 12</td>
<td>-0.0047</td>
<td>0.0112</td>
<td>-0.4160</td>
<td>0.6804</td>
</tr>
<tr>
<td>Driver 13</td>
<td>0.0001</td>
<td>0.0127</td>
<td>0.0040</td>
<td>0.9966</td>
</tr>
<tr>
<td>Driver 14</td>
<td>-0.0141</td>
<td>0.0117</td>
<td>-1.2080</td>
<td>0.2377</td>
</tr>
<tr>
<td>Driver 15</td>
<td>-0.0308</td>
<td>0.0121</td>
<td>-2.5370</td>
<td>0.0173*</td>
</tr>
<tr>
<td>Driver 16</td>
<td>-0.0230</td>
<td>0.0121</td>
<td>-1.8930</td>
<td>0.0691</td>
</tr>
<tr>
<td>Driver 17</td>
<td>-0.0058</td>
<td>0.0120</td>
<td>-0.4810</td>
<td>0.6346</td>
</tr>
<tr>
<td>Driver 18</td>
<td>-0.0048</td>
<td>0.0116</td>
<td>-0.4150</td>
<td>0.6813</td>
</tr>
<tr>
<td>Driver 19</td>
<td>-0.0007</td>
<td>0.0117</td>
<td>-0.0610</td>
<td>0.9519</td>
</tr>
<tr>
<td>Driver 20</td>
<td>-0.0132</td>
<td>0.0118</td>
<td>-1.1180</td>
<td>0.2734</td>
</tr>
<tr>
<td>Driver 21</td>
<td>-0.0201</td>
<td>0.0130</td>
<td>-1.5420</td>
<td>0.1347</td>
</tr>
<tr>
<td>Driver 22</td>
<td>-0.0220</td>
<td>0.0123</td>
<td>-1.7900</td>
<td>0.0847*</td>
</tr>
<tr>
<td>Driver 23</td>
<td>-0.0062</td>
<td>0.0118</td>
<td>-0.5270</td>
<td>0.6022</td>
</tr>
<tr>
<td>Driver 24</td>
<td>-0.0114</td>
<td>0.0114</td>
<td>-1.0070</td>
<td>0.3229</td>
</tr>
<tr>
<td>Driver 25</td>
<td>0.0116</td>
<td>0.0115</td>
<td>1.0080</td>
<td>0.3224</td>
</tr>
<tr>
<td>Driver 26</td>
<td>-0.0100</td>
<td>0.0114</td>
<td>-0.8820</td>
<td>0.3853</td>
</tr>
<tr>
<td>Driver 27</td>
<td>-0.0151</td>
<td>0.0113</td>
<td>-1.3380</td>
<td>0.1920</td>
</tr>
<tr>
<td>Driver 28</td>
<td>-0.0011</td>
<td>0.0122</td>
<td>-0.0930</td>
<td>0.9268</td>
</tr>
<tr>
<td>Driver 29</td>
<td>-0.0147</td>
<td>0.0112</td>
<td>-1.3060</td>
<td>0.2026*</td>
</tr>
<tr>
<td>Temperature</td>
<td>-0.0012</td>
<td>0.0004</td>
<td>-3.2490</td>
<td>0.0031**</td>
</tr>
<tr>
<td>Initial SOC</td>
<td>0.0002</td>
<td>0.0001</td>
<td>1.6180</td>
<td>0.1173</td>
</tr>
</tbody>
</table>

Note: p-value < 0.001 *** p < 0.01 ** p < 0.05 * p < 0.1 †
Figure 22: Residual vs. fit for whole model

The scale-location plot as shown in Figure 23 confirms the findings from the residual plot. The standardized residuals appear randomly over the fitted values. The light deviation between $y = 0.22$ and $y = 0.23$ does not show a significance heteroscedasticity.

Furthermore, the normality plot of all responses for the average consumption per mile seems to follow the theoretical normality line and thus, indicate a normal distribution over all responses (Figure 24). This confirms previous findings from the histogram.
Figure 23: Scale-location plot for whole model

Figure 24: Normality plot for whole model
Because of the space for improvement of a model fit, a stepwise regression was applied to eventually find a better fitting model. Both procedures, the forward selection and the backward elimination, were leading to the same reduced model. This new model consisted of the coefficients setting, temperature, and initial state of charge, eliminating the factor driver. This could result from the relatively homogenous demographic factor age among the drivers with an average of 23.73 years. The model fit for the reduced model slightly deteriorated with $R^2=0.4521$. This could be attributed to the fact that fewer factors are included and these reduced number of factors cannot explain the response entirely. In other words, the factor driver explained 8% of the response. However, the coefficient of determination shows that the model can explain the response with an accuracy of 45.21%, which is acceptable for an experiment in a naturalistic driving setting.

Figure 25.a shows residual plot for the model after the stepwise regression. There are no abnormalities detectable although the fitted line shows a slightly higher deviation from the fitted line compared to the full model. The scale-location plot in Figure 25.b shows a slight variation around a response of 0.23. However, these variations cannot be considered to be homoscedastic.

![Figure 25: a) Residual vs. fit and b) scale-location plot for reduced model](image)
The ANOVA table for the new model is displayed in Table 12. Both, the constant and the coefficient temperature are highly significant \( (p < 0.001) \). Taking a significance level of \( \alpha = 0.1 \) as a basis, the factors initial state of charge and traffic do also show significance. The significance of initial state of charge contradicts findings from Bartels et al. (2019) and should be treated carefully.

The basic consumption in a “No Traffic”-scenario is now 0.2571 kWh per mile. The regression coefficients for the three considered factors changed only marginally.

Table 12: ANOVA table for reduced model

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t-Value</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>0.2571</td>
<td>0.01390</td>
<td>18.4850</td>
<td>2e-16   ***</td>
</tr>
<tr>
<td>Traffic</td>
<td>0.0066</td>
<td>0.0038</td>
<td>1.7720</td>
<td>0.0818  '</td>
</tr>
<tr>
<td>Temperature</td>
<td>-0.0010</td>
<td>0.0002</td>
<td>-4.4250</td>
<td>4.5e-05 ***</td>
</tr>
<tr>
<td>Initial SOC</td>
<td>0.0003</td>
<td>0.0001</td>
<td>2.9200</td>
<td>0.0050  **</td>
</tr>
</tbody>
</table>

Note: p-value < 0.001 '***' p < 0.01 '**' p < 0.05 '*' p < 0.1''

Since the stepwise selected model has only one categorical respectively binary factor \( \text{(Traffic}=1, \text{No Traffic}=0) \) left, two separate regression equation can be formulated:

*Equation 12: Regression equation for setting "No Traffic"

\[
y = 0.2571 - 0.001001X_3 + 0.000269X_4
\]

*Equation 13: Regression equation for setting "Traffic"

\[
y = 0.2637 - 0.001001X_3 + 0.000269X_4
\]
The constants in both equations differ and indicate a slightly higher consumption of approximately 2.567% in a traffic scenario. The difference of 0.0066 kWh/mile is necessarily equal to the regression coefficient of traffic ($\beta_1$). The percentage change in average BEV energy consumption for this model is similar to the results of the full model. The impact of traffic is still low and 40% smaller than the influence of a 10°F change in temperature. The latter would lead to a decrease in average consumption of approximately 0.1 kWh/mile. This supports previous findings in the literature (Alvarez & Weilenmann, 2012; Fontaras et al., 2008; Yuksel & Michalek, 2015).

The stated positive correlation between congestion and difference in state of charge differs from the findings of Agrawal et al. (2016), Fiori et al. (2019), Wager et al. (2016), and Wu et al. (2015). However, a majority of these works were either based on road types (Wu et al., 2015) or speed profiles (Fiori et al., 2019; Wager et al., 2016). Although Agrawal et al. (2016) investigated the influence of traffic directly based on congestion, the authors collected data through simulations. This investigation is based on a naturalistic setting that investigates acceleration and jerk to provide evidence of traffic and shows different findings.

### 4.3 Quantification of Traffic Impact on Energy Consumption and Range of BEVs

After proving the impact of traffic on the energy consumption for an initial model with the most average driver and a reduced model, traffic’s influence for different driving behavior needs to be quantifiably understood. For this reason, two further multiple linear regression models were set up using different drivers (i.e., most aggressive and least aggressive) as a
reference. The most aggressive driver is referred to as the driver with the highest average energy consumption over both drives. While the regression coefficients and p-values remain the same independently from the driver reference, the basic consumption and, thus, the percentage impact of traffic changes.

As Figure 26 illustrates, the additional ΔSOC/mile varies up to 0.265% depending on the driving behavior. While the most aggressive driver used 2.357% more energy in traffic, the least aggressive driver had an additional consumption of 2.622%. If the factor driver is eliminated from consideration in the model, the increase in ΔSOC/mile is still 2.583%. Differences in percentage impacts of traffic on the energy consumption, especially between the least aggressive driver and the most aggressive driver result from the different basic consumption.
Figure 26: Increase in $\Delta$SOC for different driver types and models

Since the increased consumption in traffic was equal in absolute terms for all references and models, the meaning of traffic for the range of BEVs is necessary. Considering the chosen test route and the used test vehicle with a battery capacity of 35.8 kWh, none of the compared drivers and models would reach the declared range of 150 miles in any of the scenarios. Nevertheless, there are differences in terms of range among drivers as shown in Figure 27. While driving in the traffic scenario decreased the reachable range by 2.874 miles for the most aggressive driver, the range deteriorated by 3.550 miles for the least aggressive driver. In other words, the least energy consuming the driving behavior, the higher is the range extension potential for a BEV driver. Not taking the factor driver into account still results in a range extension of 3.507 miles. Driving a BEV with a bigger battery would lead to even bigger range differences between the two scenarios. For a
vehicle with a battery capacity of 100 kWh for example, the additional range in a “No Traffic”-scenario would be up to more than 10 miles. As part of eco-driving, traffic plays an important role and its avoidance can help drivers to overcome their range anxiety and extend the reachable range to the maximum possible.

![Figure 27: Differences in range between scenarios for different drivers and models](image)

**Figure 27: Differences in range between scenarios for different drivers and models**

### 4.4 A posteriori: Evidence of Traffic as a significant Factor based on Logistic Regression Model

The results of the t-tests for differences in means did not show significance for most three out of five traffic measures. Furthermore, an initial regression model taking all considered
factors into account did not prove significance of the factor traffic on the ΔSOC/mile. In a multiple linear regression model with categorical factors as dependent variables, these factors have to be treated as dummy variables. Additionally, the relationship between the dependent and the independent variable is assumed to be linear. These restrictions might lead to unsatisfying results.

In order to confirm the impact of traffic respectively the scenario on the energy consumption of BEVs, a logistic regression was carried out as a posteriori investigation. With a logistic regression model, the likelihood of the occurrence of a dichotomous outcome (0 for absence of outcome and 1 for presence of outcome) various numeric variables can be calculated (Davis & Offord, 1997). For this investigation, the dependency of ΔSOC/mile and the scenario (i.e., 1 for “Traffic” and 0 for “No Traffic”) was reversed. Given the values of the predictors (i.e., ΔSOC/mile and temperature) the probability of the outcome event (i.e., traffic) could be estimated. The results of the logistic regression are summarized in Table 13. Considering the factors average energy consumption and temperature, only the matter can be identified as significant (p-Value = 0.004). As shown earlier, indeed temperature is significant for the consumption of BEVs but not on the likelihood of the scenario driven. Therefore, the insignificance of temperature was expected. For the ΔSOC/mile, the percent change in odds is calculated for an increase of 0.01 kWh instead of 1 kWh for a more realistic interpretation. If the energy consumption is increased by 0.01 kWh, the odds for the event “Traffic” increase by 121%.
Table 13: Results of initial logistic regression

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Estimate</th>
<th>Percent Change in Odds</th>
<th>Std. Error</th>
<th>t-Value</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>-16.3681</td>
<td>-99.9999</td>
<td>7.5283</td>
<td>-2.174</td>
<td>0.0296  *</td>
</tr>
<tr>
<td>ΔSOC/mile</td>
<td>79.4855</td>
<td>121.4120</td>
<td>27.6414</td>
<td>2.876</td>
<td>0.0040  **</td>
</tr>
<tr>
<td>Temperature</td>
<td>-0.0286</td>
<td>-2.8195</td>
<td>0.0469</td>
<td>-0.610</td>
<td>0.5420</td>
</tr>
</tbody>
</table>

Note: p-value < 0.001*** p < 0.01 ** p < 0.05 * p < 0.1

The Receiver Operating Characteristics (ROC) diagram (Figure 28) illustrates the Area Under the Curve (AUC). The latter is a coefficient that quantifies the chance that the model will be able to distinguish between positive outcome (“Traffic”) and negative outcome (“No Traffic”) and was calculated as 79.11%. In other words, the accuracy of predicting the response correctly is 79.11%.

Figure 28: ROC curve for initial logistic regression model
Since the factor temperature was not significant for the logistic regression in its initial setup, a second iteration was carried out eliminating this factor. The results are summarized in Table 14. The only considered factor average energy consumption remained significant with an even smaller p-Value (0.009). With an additional consumption of 0.01 kWh, the odds for the event “Traffic” are even higher with 135%.

*Table 14: Results of logistic regression with coefficient ΔSOC/mile*

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Estimate</th>
<th>Percent Change in Odds</th>
<th>Std. Error</th>
<th>t-Value</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>-19.3400</td>
<td>-99.9999</td>
<td>5.8390</td>
<td>-3.312</td>
<td>0.0009 ** ***</td>
</tr>
<tr>
<td>ΔSOC/mile</td>
<td>85.6090</td>
<td>135.3939</td>
<td>25.8440</td>
<td>3.313</td>
<td>0.0009 ** ***</td>
</tr>
</tbody>
</table>

Note: p-value < 0.001 *** p < 0.01 ** p < 0.05 * p < 0.1

Additionally, the AOC for this second logistic regression model was marginally increased with 79.7778%. Hence, the chances that the model will be able to distinguish between the two scenarios is approximately 80%. The ROC diagram is illustrated in Figure 29.
The a posteriori analysis in form of the multiple linear regression helps to minimize the limitations that occurred in the multiple linear regression models. Although not all analyzed traffic measures differed significantly in both scenarios, the results of the logistic regression provide evidence of the relationship between the scenario and the consumption. The calculated percent changes in odds for the occurrence of the scenario “Traffic” when increasing the consumption by 0.01 kWh proves that increased traffic congestion leads to higher average energy consumption of BEVs. In other words, the chosen test route respectively the created scenarios based on daytime showed differences in ΔSOC/mile because of differences in traffic congestion.

Figure 29: ROC curve for logistic Regression model considering ΔSOC/mile
5 Conclusion and Future Work

This thesis aimed to determine whether traffic has an effect on the reachable range of BEVs and quantify the impact on the average energy consumption. To reach this goal, an experiment with 30 different test drivers was conducted on a predetermined test route in a naturalistic setting with two different traffic scenarios controlled by the time of the day. Driving data was collected directly from the OBD II of the test vehicle.

Since no real-time traffic data was available, actual differences in traffic on the test route were tested by investigating different traffic metrics. Based on the traffic measures average speed, travel duration, and standard deviation of speed, the test route did not show significant differences between the settings “Traffic” and “No Traffic”. However, differences between the two test scenarios “Traffic and “No Traffic” were detected for the metrics variation of acceleration and variation of jerk.

After providing evidence for differences in traffic congestion on the chosen test route based on daytime, a multiple linear regression served as a model to investigate different factors that could potentially influence the ΔSOC/mile for BEVs. The initial regression model included the factors traffic, driver, temperature, and initial SOC and was referred to as a full model. Based on the ANOVA, temperature was identified as the most significant factor. There was no significance measured for traffic although driving in the morning setting was related to a higher energy consumption. The coefficient of determination of the full model was aimed to be improved using stepwise regression. Both procedures, forward selection and backward elimination led to the same reduced model, eliminating the factor driver. While the full model could determine the response with an accuracy of 53.23%, the
reduced model was slightly less accurate (45.21%). Both coefficients of determination are acceptable for driving experiments in a naturalistic setting. In the new model, temperature remained highly significant while initial state of charge was significant for a significance level of 0.1. Additionally, traffic was detected as slightly significant with a significance level of 0.1 and therefore, led to a higher average difference in SOC per mile.

An a posteriori analysis was carried out to support the results of the multiple linear regression and minimize its limitations. Evidence could be provided that the scenario “Traffic” was correlated with a higher energy consumption while the model could predict the outcome with a probability of approximately 80%. Based on these results, it can be stated that traffic congestion has an increasing effect on the energy consumption of BEVs. In addition, focusing on speed alone is not sufficient to conclude about the effect of traffic on the ΔSOC/mile. Previous findings were based on the positive correlation between speed and energy consumption of BEVs alone (Agrawal et al., 2016; Fiori et al., 2019; Wager et al., 2016; Wu et al., 2015). Because of the fact that the average speed in the “Traffic”-scenario was even higher than in the “No Traffic”-scenario, the increasing consumption in traffic could be an indicator that variation in acceleration and variation in jerk play a major role when it comes to the energy consumption of BEVs. In other words, with higher differences in acceleration as they occur more likely in traffic, more energy is consumed even if there is a lower average speed. Future work could specifically focus on these two measures to distinguish between the impact of speed and the impact of variation in acceleration respectively jerk on the energy consumption of BEVs. While for ICEs the fuel consumption increases for all named measures, the single effect of these measures on the energy consumption of BEVs is not yet explored entirely.
Although evidence for the impact of traffic was provided using the two scenarios, the chosen test route did not show a significant difference for the traffic metrics average speed and average duration. Since higher traffic congestion is related to lower average speed respectively longer travel time (Barth & Boriboonsomsin, 2008; Ge et al., 2015), future work should conduct experiments on a test route with higher sensitivity towards changes in traffic congestion over daytime. Although jerk and acceleration patterns are promising measures, the evidence of traffic should additionally be confirmed based on real-time traffic data (Barth & Boriboonsomsin, 2008; Grubwinkler et al., 2014).

Ultimately, this research was able to quantify the impact of traffic on the difference in state of charge of electric vehicles while limitations in measuring traffic and driver diversity should be minded. Based on the results of the regression analyses, traffic can lead to an increase in average energy consumption of up to 2.6% compared to a scenario with no traffic depending on the driving behavior. In other words, there is a range extension potential of 3.55 miles when driving in a BEV with a battery capacity of 35.8 kWh. This potential increases with the battery capacity and, therefore, can be a crucial factor with regard to driving behavior and navigation.

By understanding the importance and quantifying the impact of traffic, intelligent transportation systems, as well as eco-driving strategies, can be developed and improved. A negotiation of range limitation and anxiety can be overcome and the transition to a more ecological transportation system and advanced sustainability of society can be relieved. This understanding can help to overcome the obstacles of range limitation and range anxiety to facilitate the transition to a more ecological transportation system and relieve an advanced sustainability of society.
6 Appendices

Appendix A – Recruitment Advertisements

DRIVING FOR SUSTAINABILITY

Interested in driving an electric vehicle? Contact us for participating in test runs!

Contact:

Time commitment: 3 runs â€œ 1 hour each
When? March-May 2019
Requirements: At least 18 years old, valid driver’s license

DRIVING FOR SUSTAINABILITY

Although electric vehicles are getting more popular, widespread adoption is hindered by range limits and long charging times. We are investigating the influence of traffic and road types on the efficiency of electric vehicles.

Interested in driving an electric vehicle? Contact us for participating in test runs!

Contact:

Time commitment: 3 runs â€œ 1 hour each
When? March-May 2019
Requirements: At least 18 years old, valid driver’s license
Appendix B – Test Drive Contact Email

Dear interested EV Driver,

Thank you for being interested in helping research in the field of electric vehicles. Test drives will start on Monday, April 1st and will be scheduled over a period of 10 weeks.

Each participant conducts three test drives on predetermined test routes starting at URI and supervised by one of the researchers. The duration of each drive will be between 35 and 50 minutes.

One test drive will be scheduled in the morning between 7:30 AM and 8:30 AM while the other two drives will be done between 10:00 AM and 4:00 PM. A maximum of two drives can happen on the same day. To participate you must possess a valid driver’s license and you must be at least 18 years old.

Collected data will be anonymized and exclusively used for the purpose of this research.

In order to schedule the runs, please send an E-mail with proposed dates to Taris Wilde (taris_wilde@my.uri.edu) or Tim Jonas (t_jonas@my.uri.edu).

Thank you and Best wishes,

Taris Wilde & Tim Jonas

_____________________________
Principle Investigator: Gretchen A. Macht, Ph. D.
This is URI research
This research has been approved by The University of Rhode Island Institutional Review Board

Appendix C – Test Drive Reminder Email

Dear participant,

This is a reminder for the test drive scheduled for you tomorrow.

Time: 10:00 AM
Location: Entrance of Kirk building
Investigator:

If for any reason you cannot participate, please contact one of the investigators directly.
Appendix D – Consent Form

THE UNIVERSITY OF RHODE ISLAND
Dr. Gretchen A. Macht
Department of Mechanical, Industrial and Systems Engineering
Electric Vehicle Driver Behavior

Consent Form for Research

Dear Participant,

You are being asked to take part in a research study. The purpose of this study is to investigate how driver behavior affects electricity used while driving electric vehicles. Please read the following before agreeing to be in the study. If you agree to be in this study, it will take you approximately 1-3 hours to complete this survey. Questions will be asked about electric vehicle driver behavior. There are no known risks, benefits or compensation. You must be at least 18 years old and be in possession of a valid American or international driver’s license to be in this research project.

Your responses will be strictly confidential. All collected information will be stored in Pastore 333 in a locked file cabinet and digital information will be stored in computers in Pastore 254 as locked files. Only the researchers will have access to these files. No audio or video file will be collected. All the information required will be collected directly from the car. The responses may be used in research papers and master theses.

The decision to participate in this study is entirely up to you. You may refuse to take part in the study at any time without affecting your relationship with the investigators of this study or the University of Rhode Island (URI). Your decision will not result in any loss of benefits to which you are otherwise entitled. You have the right not to answer any single question, as well as to withdraw completely from the survey at any point during the process; additionally, you have the right to request that the researchers not use any of your responses.

You have the right to ask questions about this research study and to have those questions answered by me before, during or after the research. If you have questions about the study, at any time feel free to contact Dr. Macht from the Mechanical, Industrial and Systems Engineering Department at (401)-874-2243.

Additionally, you may contact the URI Institutional Review Board (IRB) if you have questions regarding your rights as a research participant. Also contact the IRB if you have questions, complaints or concerns which you do not feel you can discuss with the investigator. The University of Rhode Island IRB may be reached by phone at (401) 874-4328 or by e-mail at researchintegrity@etal.uri.edu. You may also contact the URI Vice President for Research and Economic Development by phone at (401) 874-4576.
Appendix E – Macro for Calculations for Every Test Drive

Sub LoopFolderAndCopy()
    Dim sFolder As String
    Dim sFile As String
    Dim wbSource As Workbook
    Dim StartTime As Double
    Dim MinutesElapsed As String
    Dim lRow As Integer
    Dim i As Integer
    Dim j As Integer
    Dim k As Integer
    Dim l As Integer
    Dim m As Integer
    Dim n As Integer
    Dim o As Integer
    Dim p As Integer
    Dim q As Integer
    Dim r As Integer
    Dim lRowCut As Integer
    Dim MinColQ As Integer
    Dim MinColR As Integer
    Dim MinColS As Integer
    Dim MinColT As Integer
    Dim StartCut1 As Integer
    Dim EndCut1 As Integer
    Dim StartCutCon As Integer
    Dim EndCutCon As Integer
    Dim StartCut2 As Integer
    Dim EndCut2 As Integer

    sFolder = "\Users/timjonas/Desktop\tryfolder"

    ' On Error GoTo errHandler 'reset application setting on error
    Application.ScreenUpdating = False

    'Start Timer for entire process
    StartTime = Timer

    'loop through all excel.xlsm files in folder
    sFile = Dir(sFolder & "*.xlsm")
    Do Until sFile = ""


'open the source workbook
Set wbSource = Workbooks.Open(sFolder & sFile)

'activate and copy the first worksheet and rename it to: "CuttingEdges"
wbSource.Activate
'Acutivate and copy the first worksheet and rename it to: "CuttingEdges"
wbSource.ActiveSheet.Columns("A:L").AutoFit
'Copy Worksheet(1)
'Rename Worksheet(2)
ActiveSheet.Name = "CuttingEdges"

'Headers Columns
wbSource.ActiveSheet.Range("K6") = "VehicleSpeed(m/s)"
wbSource.ActiveSheet.Range("L6") = "Acceleration"
wbSource.ActiveSheet.Range("M6") = "Jerk"
wbSource.ActiveSheet.Range("N6") = "deltaDist_m/s(mi)"
wbSource.ActiveSheet.Range("O6") = "deltaSOC(kWh/sec)"
wbSource.ActiveSheet.Range("P6") = "deltaSOC/mi"
wbSource.ActiveSheet.Range("Q6") = "START CUT"
wbSource.ActiveSheet.Range("R6") = "END CUT"
wbSource.ActiveSheet.Range("S6") = "CON. START"
wbSource.ActiveSheet.Range("T6") = "CON. END"

'find the last row and the amount of rows
'MsgBox (lRow)

'Calculation of columns
For i = 7 To lRow
    wbSource.ActiveSheet.Range("K" & i).Formula = "=J" & i & "/3.6"
Next i

For j = 7 To lRow
    wbSource.ActiveSheet.Range("L" & j).Formula = ";IFERROR(K" & j & ";-K" & j - 1 & ",""")"
Next j

For k = 7 To lRow
    wbSource.ActiveSheet.Range("M" & k).Formula = ";IFERROR(L" & k & ";-L" & k - 1 & ",""")"
Next k
For l = 7 To lRow
    wbSource.ActiveSheet.Range("N" & l).Formula = "+=IF(K & l & "+=0,K & l & "+=/1609.344)"
Next l

For m = 7 To lRow
    wbSource.ActiveSheet.Range("O" & m).Formula = "=((C & m & "*E" & m & "+)/1000)/3600)"
Next m

For n = 7 To lRow
    wbSource.ActiveSheet.Range("P" & n).Formula = "=IF(OR(N & n & "+="",N & n & "+=0),"",O & n & "+/N & n & "+")"
Next n

For o = 7 To lRow
    wbSource.ActiveSheet.Range("Q" & o).Formula = "=SUM(SQRT((G & o & "+41.489207)^2),SQRT((H & o & "+71.521673)^2))"
Next o

For p = 7 To lRow
Next p

For q = 7 To lRow
    wbSource.ActiveSheet.Range("S" & q).Formula = "=SUM(SQRT((G & q & "+41.432126)^2),SQRT((H & q & "+71.607242)^2))"
Next q

For r = 7 To lRow
    wbSource.ActiveSheet.Range("T" & r).Formula = "=SUM(SQRT((G & r & "+41.429276)^2),SQRT((H & r & "+71.568301)^2))"
Next r

'Fit the width of all Columns
wbSource.ActiveSheet.Columns("A:AN").AutoFit

'Copy Worksheet(2)

'Rename Worksheet(3)
ActiveSheet.Name = "CuttedEdges"

'find the last row and the amount of rows before the cut
' MsgBox (MinColQ & Space(2) & MinColR & Space(2) & MinColS & Space(2) & MinColT)

' Cut of Start
StartCut1 = 7
EndCut1 = MinColQ - 1

Rows("" & StartCut1 & "," & EndCut1 & ",").EntireRow.Delete


' Cut of Construction site
StartCutCon = MinColS
EndCutCon = MinColT + 9

Rows("" & StartCutCon & "," & EndCutCon & ",").EntireRow.Delete

' What is now the last row in the Sheet?


StartCut2 = MinColR + 9
EndCut2 = lRowCut

Rows("" & StartCut2 & "," & EndCut2 & ",").EntireRow.Delete

' Format All Cells as Values
With ActiveSheet.UsedRange
 .Value = .Value
End With

Worksheets.Add(After:=Worksheets(3)).Name = "Output"
Worksheets("Output").Activate

' Name final cells
wbSource.ActiveSheet.Range("A1") = "Total consumption"
wbSource.ActiveSheet.Range("B1") = "Distance"
wbSource.ActiveSheet.Range("C1") = "Mean Consumption"
wbSource.ActiveSheet.Range("D1") = "Total Time (ms)"
wbSource.ActiveSheet.Range("E1") = "StdDev SOC/mile"
wbSource.ActiveSheet.Range("F1") = "Average Speed"
wbSource.ActiveSheet.Range("G1") = "StdDev Speed"
wbSource.ActiveSheet.Range("H1") = "Average Acceleration"
wbSource.ActiveSheet.Range("I1") = "StdDev Acceleration"
wbSource.ActiveSheet.Range("J1") = "Average Jerk"
wbSource.ActiveSheet.Range("K1") = "StdDev Jerk"

'Calculation of final values

wbSource.ActiveSheet.Range("A2").Formula = "+SUM(Z7:Z9999)*(-1)"
wbSource.ActiveSheet.Range("B2").Formula = "+SUM(Y7:Y9999)"
wbSource.ActiveSheet.Range("C2").Formula = "+(A2/B2)"
wbSource.ActiveSheet.Range("D2").Formula = "COUNT(L7:L9999)"
wbSource.ActiveSheet.Range("E2").Formula = "+STDEV.P(AA7:AA9999)"
wbSource.ActiveSheet.Range("F2").Formula = "+AVERAGE(U7:U9999)*0.621371"
wbSource.ActiveSheet.Range("G2").Formula = "+STDEV(U7:U9999)*0.621371"
wbSource.ActiveSheet.Range("H2").Formula = "+AVERAGE(W7:W9999)"
wbSource.ActiveSheet.Range("I2").Formula = "=STDEV.P(W7:W9999)"
wbSource.ActiveSheet.Range("J2").Formula = "=STDEV.P(X7:X9999)"
wbSource.ActiveSheet.Range("K2").Formula = "=STDEV.P(X7:X9999)"

'Headers of Road Type Calculations

wbSource.ActiveSheet.Range("A7") = "Interstate_I"
wbSource.ActiveSheet.Range("A8") = "Interstate_D"
wbSource.ActiveSheet.Range("A9") = "FreExp_I"
wbSource.ActiveSheet.Range("A10") = "FreExp_D"
wbSource.ActiveSheet.Range("A11") = "OthPrinArt_I"
wbSource.ActiveSheet.Range("A12") = "OthPrinArt_D"
wbSource.ActiveSheet.Range("A13") = "MinArt_
wbSource.ActiveSheet.Range("A14") = "MajCol_I"
wbSource.ActiveSheet.Range("A15") = "MajCol_D"
wbSource.ActiveSheet.Range("A16") = "MinCol_I"
wbSource.ActiveSheet.Range("A17") = "MinCol_D"
wbSource.ActiveSheet.Range("A18") = "Local_I"
wbSource.ActiveSheet.Range("A19") = "Local_D"

'Fill Color

'Text Color
wbSource.ActiveSheet.Range("A1:K1").Font.Color = RGB(250, 125, 0)

'Text Bold

'AutoFit
wbSource.ActiveSheet.Columns("A:AT").AutoFit

'Border of cells

wbSource.Close savechanges:=True

Application.CutCopyMode = False

'get the next file
sFile = Dir()
Loop

'Determine how many seconds code took to run
MinutesElapsed = Format((Timer - StartTime) / 86400, "hh:mm:ss")

'Notify user of time
MsgBox "The entire folder was successfully processed in incredible: " & MinutesElapsed & " hh:mm:ss", vbInformation

'tidy up
Set wbSource = Nothing
Appendix F – Macro for Extracting Data in Compact Sheet

Sub copydata()
    Dim FolderPathO As String, FilePath As String
    Dim TargetPath As String
    Dim wbSourceO As Workbook
    Dim wbDest As Workbook
    Dim wbName As String
    FolderPathO = "C:\Users\taris\Desktop\testData\"
    TargetPath = "C:\Users\taris\Desktop\CollectionTraffic.xlsm"
    FilePath = Dir(FolderPathO & "*.xlsm")
    Dim erow As Long, lastrow As Long, lastcolumn As Long

    Set wbDest = Workbooks.Open(TargetPath)
    wbDest.Worksheets(1).Activate
    Range("A1") = "Workbookname"
    Range("B1") = "Driver ID"
    Range("C1") = "Total consumption"
    Range("D1") = "Distance"
    Range("E1") = "Mean Consumption"
    Range("F1") = "Total Time (ms)"
    Range("G1") = "StdDev SOC/mile"
    Range("H1") = "Average Speed"
    Range("I1") = "Std Dev Speed"
    Range("J1") = "Average Acceleration"
    Range("K1") = "Std Dev Acceleration"
    Range("L1") = "Average Jerk"
    Range("M1") = "StdDev Jerk"

    'Fill Color
    'Text Color
    wbDest.ActiveSheet.Range("A1:M1").Font.Color = RGB(250, 125, 0)
    'Text Bold
    wbDest.ActiveSheet.Range("A2:B31").Font.Bold = True 'Bold First two columns

    'AutoFit
    wbDest.ActiveSheet.Columns("A:M").AutoFit

    'Border of cells


Do Until FilePath = ""

Set wbSourceO = Workbooks.Open(FolderPathO & FilePath)
    wbSourceO.Worksheets("Output").Activate
    wbName = ActiveWorkbook.Name

    Range(Cells(2, 1), Cells(2, 11)).Copy

    wbDest.Worksheets(1).Activate
    'Last filled row in Column 3(C)
    erow = ActiveSheet.Cells(Rows.Count, 3).End(xlUp).Offset(1, 0).Row
    'Paste the previously copied range into erow, starting in column C
    ActiveSheet.Cells(erow, 3).PasteSpecial Paste:=xlPasteValues
    'Insert the Workbook name into Column A
    Range("A" & erow & ").Value = wbName

    'AutoFit
    wbDest.Worksheets("A:M").AutoFit

    'Close Workbook, but dont save
    wbSourceO.Close savechanges:=False

    FilePath = Dir

Loop

    erow = ActiveSheet.Cells(Rows.Count, 1).End(xlUp).Offset(1, 0).Row
    Cells(erow, 1).Select

End Sub
7 Bibliography


https://doi.org/10.1016/j.atmosenv.2008.01.053


