32nd Electric Vehicle Symposium (EVS32)
Lyon, France, May 19 - 22, 2019

PREDICTING ELECTRIC VEHICLE CONSUMPTION : A PHYSICAL MODEL THAT FITS

Anthony Deschênes[1], Jonathan Gaudreault[1], Kim Rioux-Paradis[1]
[1]*CRISI Research Consortium for Industry 4.0 Systems Engineering, Université Laval, Québec City, Québec, CANADA Quebec, G1V 0A6, CANADA, anthony.deschenes.1@ulaval.ca, jonathan.gaudreault@ift.ulaval.ca, kim.rioux-paradis.1@ulaval.ca*

Abstract

Electric vehicles become more and more important in our society. Using them in a fleet to minimize energy cost is therefore a compelling opportunity for taxi companies. It is, however, crucial to develop efficient models that estimate energy consumption to travel from one point to another. Consumption can be estimated using a physical model, but such a model is too theoretical and fails to fit real world data, especially in taxi driving conditions. Therefore, we propose to learn from historical data in order to correct/improve the physical model. The same techniques can be used to estimate consumption for a new vehicle model, which can be useful for a company who wants to add a new type of car for which they know nothing except its physical parameters.

energy consumption, fleet, EV, efficiency

1 Introduction

TEO Taxi (Montreal, Canada) is a company that runs a fleet of 100 % electric taxis (approximately 170 cars). This leads to cost reduction and a reduction of greenhouse gas emissions [1]. However, electric vehicles used by the company, which are Nissan LEAF, Kia Soul EV, Tesla Model S and Tesla Model X, do not offer as much autonomy as conventional internal combustion engine cars. Since official specifications lacks precision [2], the need for an accurate energy consumption prediction model is a real preoccupation for the company. This is mandatory to allow optimized usage of each owned vehicle as well as for future acquisitions. TEO Taxi also wants to be able to use the developed model to allow them to predict the consumption for "new/unknown" vehicle models with a better efficiency early on.

Such models do exist [3][4], but they are not adapted to manage electric taxis. They both rely on a physical model, but do not take into account the specific taxi driving constraints such as frequent stops, doors being opened and closed frequently and intense urban driving. Another very important factor is the temperature in which the fleet evolves that can go from -30°C to 30 °C depending on the time of the year. Autonomy of any electric vehicle decline in cold temperature [5] and it can be even worse in the context of taxi driving where doors frequently open, thus leading to heat loss and increase in energy consumption to maintain a decent temperature in the cabin.

Inspired by [3], our main goal is to adapt a common physical model to take into consideration the various external factors that affect electric vehicle consumption. Two objectives were considered : (1) enhance the accuracy of the physical model using historical data from a known vehicle model in order to adjust it using linear regression, and (2) allow a better prediction for the physical model for a new/unknown vehicle model added to the fleet.

2 Dataset

The dataset was generated using historical data from all TEO Taxi vehicle models. It contains information such as distance driven, speed, elevation, temperature, trip duration and more. For Nissan LEAF, Kia Soul EV, Tesla Model S and Tesla Model X we respectively had access to 26269, 131866, 22082 and 13130 rides. Each ride was divided into steps of approximately 2-8 seconds containing information about the distance driven during the step, the elevation delta, the speed and the acceleration. Data derives mainly from a GPS signal and a data logger made by Fleet Carma company.

There is no interest in predicting the consumption for really small rides and potential errors of precision from the sensors could deceive the results hence we decided to remove it from the dataset. Ride with an average speed over 100 km/h are not normal in a taxi driving context and are also excluded. Since the data came from Montreal and the highest mountain is 233 meter tall, it is likely that a ride that greatly exceeds this have an error and should be excluded. A taxi ride rarely exceeds one hour meaning all ride too short or too long can also be excluded. Finally, a known error with the logs occurs when the vehicle passes through a tunnel, causing it to cross it in 2-8 seconds, causing the speed to be really quick (around 900 to 2000 km/h depending on the cases). We also considered that a driver in normal taxi driving circumstances would never go faster than 130 km/h, thus making any ride containing this an error probably due to the sensors. Therefore the following filters were applied :

- Keep all rides with distance driven greater or equal than 1 km (removed 15264 rides)
- Keep all rides with average speed lesser than 100 km/h (removed 53 rides)
- Keep all rides with gained and lost altitude smaller than 500 meters (removed 8 rides)
- Keep all rides with duration between 5 and 3600 seconds (removed 165 rides)
- Remove rides where the speed for a step is greater than 130 km/h (removed 12 942 rides)

Inspecting carefully the speed of consecutive steps of a given ride, we realized there were some inconsistencies. It was later established that the timestamps of the GPS points were not evenly spread in time as they should be. The timestamps were therefore corrected and a moving average was applied to correct what was identified as unexplained abnormalities.

3 Physical Model

3.1 Basic Physical Model

We started with the physical model from [3] :

$$E = \frac{1}{3600}(mg(f\cos\theta + \sin\theta) + \frac{1}{2}(\rho C_x A(\frac{v}{3.6})^2) + ma) \cdot d \tag{1}$$

E = Energy required to travel distance d (kWh)
m = Mass of vehicle (kg)
g = Gravitational acceleration (m/s^2)
f = Rolling resistance of vehicle (-)
θ = Road angle in radians (kg/m^3)
ρ = Air density (-)
C_x = Drag coefficient of vehicle (-)
A = Frontal area of vehicle (m^2)
v = Speed of vehicle (km/h)
a = Acceleration of vehicle (m/s^2)
d = Distance driven (km)

The first two terms can be seen as the energy required for ascension ($f\cos\theta$) and potential recovery in descent ($\sin(\theta)$). The third one ($\rho C_x A(\frac{v}{3.6})^2$) is the aerodynamics loss and the last one (ma) is the loss or regeneration due to acceleration.

3.2 Extending The Physical Model - Additional Terms

Some external factors are not taken into account by this basic model. Therefore, we added another term (T) to take into account air conditioning energy as a function of the outside temperature (in Celsius). The function is represented in Figure 1. It has been derived from historical data captured using Fleet Carma data logger. The function is divided into five segments and each choke-point coordinates has been determined by fitting the data using a solver to perform a least squared minimization. The logic behind this function is the following: the colder it is, the more energy will be needed to heat the cabin, but there is a point at which the heating component reaches its maximal energy consumption. When temperature rises, there is a point at which we used air conditioning, which takes less energy than heating [6].

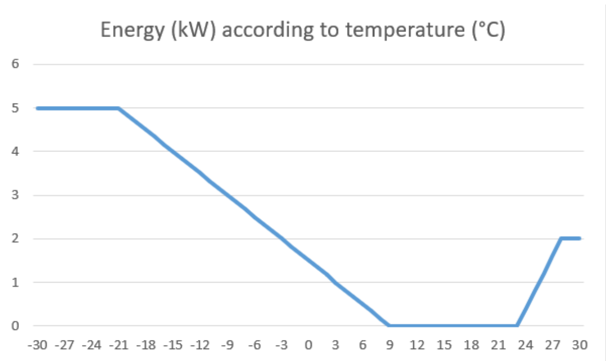

Figure 1: Air conditioning energy according to outside temperature

Similarly, we also extended the physical model by adding a term A_c(kWh) for accessories consumption. Again, the value was approximated using the Fleet Carma data logger.

$$A_c = 0.2 \cdot s/3600 \tag{2}$$

$$s = \text{ Time elapsed (seconds)}$$

The physical model after extending it with our terms is now the following :

$$E = \frac{1}{3600}(mg(f\cos\theta + \sin\theta) + \frac{1}{2}(\rho C_x A(\frac{v}{3.6})^2) + ma) \cdot d + T + A_c \tag{3}$$

3.3 Fitting The Physical Model

In order to correct/improve the physical model, a "weight" parameter was added to each term of the physical model. Using the historical database containing thousands of taxi rides, we were able to determine the best values for those weight parameters using linear regression. This way, we can correct the physical model to take into consideration external factors that were not initially considered by the model (that is, taxi specific conditions) and allowed the trained models to get a better prediction for future rides. Moreover, some factors, such as battery capacity [7] and rolling resistance [8], are not constant and vary with external elements such as temperature, but these variations can, in part, be corrected while training/adjusting the models.

4 Experiments

In this section, we compare the different models: Basic physical model, Extended physical model and Extended physical model fitted using a linear regression.

The models for each car were trained separately, allowing the learning models to fit at best with each type of car. For each vehicle model, a subset of the database (80 % of the taxi rides) was randomly chosen in order to define training set. The resulting trained model was tested using the 20% remaining taxi rides. We repeated this process 10 times for each model to create a 95 % confidence interval.

4.1 Results

Tables 1 to 3 show the results for the Basic physical model, the Extended physical model and the Fitted physical model using linear regression. We compared the result using the Mean Absolute Error (MAE):

Table 1: MAE (kWh) for the Basic physical model

Vehicle Model	0-5km	5-10km	10-15km	15-20km	20-25km	Average [1]
Kia Soul	0.33 ± 0.0052	0.55 ± 0.020	0.84 ± 0.18	1.019 ± 0.70		0.68 ± 0.081
Nissan LEAF	0.59 ± 0.016	0.88 ± 0.066	1.22 ± 0.47			0.90 ± 0.037
Tesla Model S	0.51 ± 0.015	0.70 ± 0.048	0.73 ± 0.056	1.014 ± 0.12	1.48 ± 1.23	0.89 ± 0.11
Tesla Model X	0.55 ± 0.030	0.73 ± 0.071	0.76 ± 0.057	1.067 ± 0.20	1.34 ± 0.72	0.89 ± 0.10
Weighted Average	0.40 ± 0.22	0.62 ± 0.25	0.87 ± 0.35	1.024 ± 0.55	1.41 ± 0.95	

[1] There was a different number of rides for each distance cluster thus we had to use poststratification [9] to calculate the confidence interval

Table 2: MAE (kWh) for the Extended physical model

Vehicle Model	**0-5km**	**5-10km**	**10-15km**	**15-20km**	**20-25km**	**Average**[1]
Kia Soul	0.22 ± 0.0043	0.50 ± 0.019	0.81 ± 0.15	0.99 ± 0.65		0.63 ± 0.076
Nissan LEAF	0.45 ± 0.013	0.72 ± 0.065	0.91 ± 0.32			0.69 ± 0.025
Tesla Model S	0.42 ± 0.018	0.63 ± 0.048	0.77 ± 0.063	1.18 ± 0.13	1.50 ± 1.09	0.90 ± 0.10
Tesla Model X	0.43 ± 0.024	0.64 ± 0.061	0.78 ± 0.049	1.18 ± 0.19	1.15 ± 0.57	0.83 ± 0.080
Weighted Average	0.29 ± 0.20	0.56 ± 0.17	0.82 ± 0.17	1.038 ± 0.54	1.33 ± 0.89	

Table 3: MAE (kWh) for the Fitted extended physical model using linear regression

Vehicle Model	**0-5km**	**5-10km**	**10-15km**	**15-20km**	**20-25km**	**Average**[1]
Kia Soul	0.17 ± 0.0044	0.38 ± 0.019	0.58 ± 0.12	0.62 ± 0.39		0.44 ± 0.045
Nissan LEAF	0.25 ± 0.014	0.41 ± 0.042	0.51 ± 0.27			0.39 ± 0.021
Tesla Model S	0.28 ± 0.014	0.42 ± 0.033	0.48 ± 0.039	0.58 ± 0.059	0.75 ± 0.62	0.50 ± 0.057
Tesla Model X	0.30 ± 0.02	0.46 ± 0.043	0.56 ± 0.041	0.66 ± 0.16	1.004 ± 0.45	0.60 ± 0.063
Weighted Average	0.21 ± 0.094	0.39 ± 0.056	0.56 ± 0.15	0.62 ± 0.31	0.87 ± 0.52	

Figure 2 summarizes the results. It reports the average MAE for each model and each vehicle. Those results corresponds to the last column of Tables 1, 2, and 3. Except for the Nissan LEAF, there is only a marginal improvement between the Basic and the Extended physical models. However the Fitted extended physical model is significantly better for each vehicle model. On average we observe, for linear regression, a reduction of the error by 36 % in comparison to the Extended physical model.

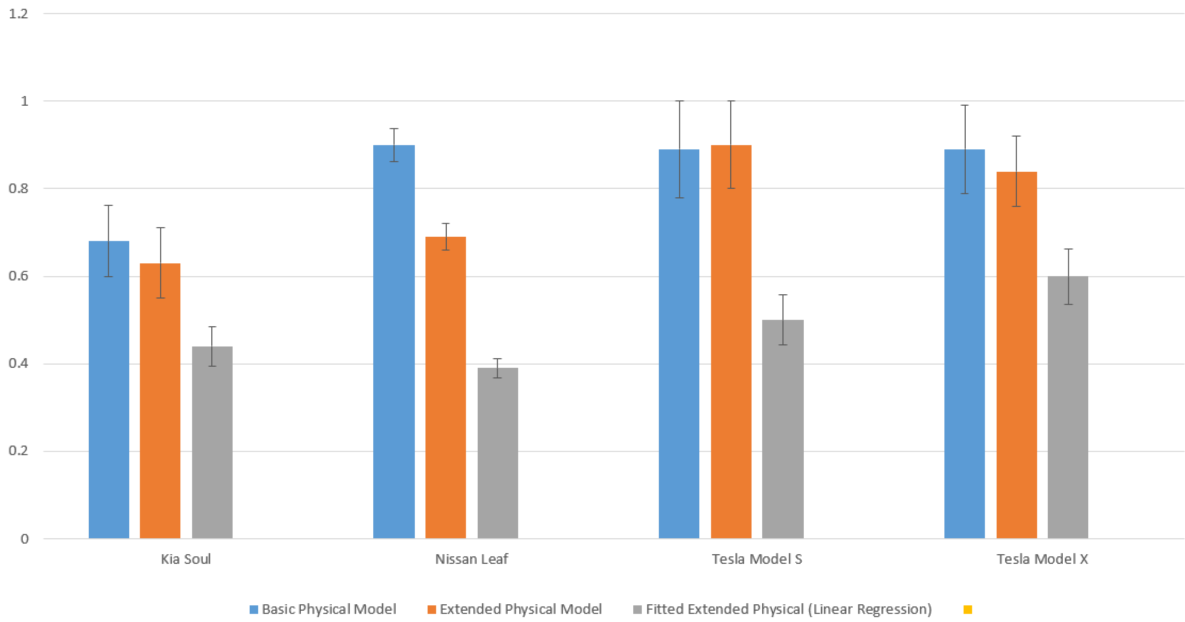

Figure 2: Average of MAE for each training model based on vehicle model

4.2 Predicting consumption for a new vehicle

Predicting the consumption for a new/unknown vehicle is crucial for companies that manage a fleet of electric vehicles. Physical models are interesting because they can easily adapt to new vehicle particularities. We could use a similar technique as previously, but we lack historical data to learn from. Therefore, we propose that, for each vehicle model, we evaluate if training using data from other known vehicles allows improvement of the Extended physical model for the new/unknown vehicle.

We used linear regression to learn from all vehicle except the one we consider as the new/unknown vehicle. As an example, for the Nissan LEAF, our training set contained all data from the Kia Soul, Tesla

Model S and Tesla Model X, and then we tested the resulting model on all Nissan LEAF rides. However, as we did not had access to data for driven distance larger than 25 km for the Nissan LEAF, we excluded data from the Nissan LEAF when training for other cars. Table 5 presents the results for all vehicles models.

Table 4: Comparing MAE (kWh) for new/unknown vehicle model (Nissan LEAF excluded from the training phase)

Vehicle Model	**Basic Physical Model**	**Extended Physical Model**	**Fitted Extended Physical Model (Linear Regression)**
Kia Soul	0.68	0.62	0.46
Nissan LEAF	0.85	0.66	0.63
Tesla Model S	0.91	0.93	0.59
Tesla Model X	0.88	0.83	0.61

Since each test set contains all data about a specific vehicle model, it is normal to have exact values without confidence intervals. For linear regression we observe on average a reduction of 23 % of the error in comparison to the Extended physical model.

5 Conclusion

Using historical data to train the Extended physical model with linear regression does improve its efficiency by as much as 44 %, which is huge from the point of view of the company. However, the historical data used contains a lot of variations due to the fact that it comes from physical sensors that are subjected to various errors and many uncontrolled parameters (e.g. how each driver drives the car). These variations suggest that the results are not as precise as they could be. Also, the fact that our dataset did not contain enough rides in the distance clusters larger than 25 km restricted precision for such distance thus not evaluating models for higher driven distance. As for future works, we are now working on a non linear model (relying on neural network) that includes parameters characterizing the different car and rides (e.g. age of the car, actual driver, road conditions, etc.)

Acknowledgments

The authors would like to acknowledge TEO-Taxi for providing us with the data.

References

[1] Maarten Messagie, Faycal-Siddikou Boureima, Thierry Coosemans, Cathy Macharis, and Joeri Mierlo. A range-based vehicle life cycle assessment incorporating variability in the environmental assessment of different vehicle technologies and fuels. 7(3):1467–1482.

[2] De C., Maarten M., Heyvaert S., Coosemans T., and Van J. Electric vehicle use and energy consumption based on realworld electric vehicle fleet trip and charge data and its impact on existing EV research models. 7(3):436–446.

[3] Cedric De Cauwer, Wouter Verbeke, Thierry Coosemans, Saphir Faid, and Joeri Van Mierlo. A data-driven method for energy consumption prediction and energy-efficient routing of electric vehicles in real-world conditions. 10(5):608.

[4] Green race 2.0. https://www.jurassictest.com/greenrace-2. Accessed on 2019-02-22.

[5] Juhani Laurikko, Robert Granstrom, and Arto Haakana. Realistic estimates of EV range based on extensive laboratory and field tests in nordic climate conditions. In *2013 World Electric Vehicle Symposium and Exhibition (EVS27)*, pages 1–12. IEEE.

[6] The truth about electric vehicles in cold weather. https://fr.slideshare.net/fleetcarma/the-truth-about-electric-vehicles-in-cold-weather. Accessed on 2018-07-17.

[7] O. Erdinc, B. Vural, and M. Uzunoglu. A dynamic lithium-ion battery model considering the effects of temperature and capacity fading. In *2009 International Conference on Clean Electrical Power*, pages 383–386. IEEE.

[8] Emma Arfa Grunditz and Torbjorn Thiringer. Performance analysis of current BEVs based on a comprehensive review of specifications. 2(3):270–289.

[9] John T. Kulas, David H. Robinson, Jeffrey A. Smith, and Donald Z. Kellar. Post-stratification weighting in organizational surveys: A cross-disciplinary tutorial. 57(2):419–436.

Authors

Anthony Deschênes is a bachelor student at Université Laval since 2017. He graduated in computer science from Cegep de Ste-Foy in 2017 and has been working with professor Jonathan Gaudreault on projects related to electric vehicle consumption prediction since 2018.

Jonathan Gaudreault is the director of CRISI, a research consortium dedicated to Industry 4.0 systems engineering. Over the last eighteen years he has led a number of key projects related to operations planning and scheduling, bringing together ideas from the artificial intelligence and operations research fields. He previously worked as a RD manager in the software industry and was awarded as ”One to watch” by the Federation de l'Informatique du Québec (OCTAS award). In 2012, Pr. Gaudreault and other researchers from the FORAC Research Consortium were recognized with the Brockhouse Canada Prize for Interdisciplinary Research in Science and Engineering, Canada's premier award for interdisciplinary research.

Kim Rioux-Paradis is a PhD student at Université Laval. She has a bachelor's degree in mathematics and computer science and a certificate in statistics at Université Laval. She did her master in computer science in constraint programming, specifically on the multiple travelling salesman problem also at Université Laval. She now works on charging plans for electric taxi fleets.