A Review of Statistical Models in Predicting Electric Vehicle Range

Yusef Mostafa

Abstract:

Recently, the automotive industry has been in the process of developing fully electric vehicles (EVs) to draw from clean energy and reduce the carbon footprint of vehicles. Using statistics regression methods, models of the predicted driving range can be generated. Whether by a dynamics energy equation or machine learning algorithm, EV driving data can be used in developing these models. Using a simple linear regression model, taking into account driving factors such as acceleration, wind resistance, and incline driving, regression models were successfully developed with a correlation coefficient (R^2) of 95.6%. Similarly, using a machine learning algorithm to develop multiple linear regressions (MLR) led to the development of multiple regression models, most notably a polynomial regression with R^2 of 97.3%. While these statistical models are limited in the number of factors chosen and their significance, using basic polynomial regressions have proven to yield competent EV range predictor models.

Nomenclature:

E_{ij}	Mechanical energy required
M_{ij}	Vehicle mass
M_f	fictive mass of rolling inertia
G	gravitational acceleration
F	Vehicle coefficient of rolling resistance
φ	Road gradient angle
ρ	Air density
C_x	Drag coefficient
A	Vehicle cross sectional area
V_{ij}	Vehicle speed between i and j
D_{ij}	Distance driven between i and j
B_i	Regression coefficient
V	Average speed
T	Time
S	Distance
H	elevation (positive/negative)
ε	error

Introduction:

In recent years, the automotive industry has undergone significant innovation in the electric vehicle (EV) space. Given the push for cleaner energy sources, the majority of auto makers have developed or are in the process of developing electric vehicles to transition from gas engines to renewable energy sources. While this change is essential for the transition to clean, renewable energy and decreasing the large carbon footprint of the automotive industry, the transition comes with its own challenges. One such challenge is the calculation and prediction of the range an electric car range has in between charges. In the past, the range of gas-powered cars was typically computed by the car's computer by the amount of gas remaining in the tank, the known average consumption rate of gas by the engine, and a fuel economy number set by the manufacturer. Similarly, electric car ranges can be determined by using statistical models of real world data in order to predict the remaining driving range of an electric vehicle. This study will review some of these statistical models and methods used by automotive manufacturers and researchers to accurately predict the driving range.

Currently, manufacturers such as Tesla compute battery range of the vehicle from fixed historical testing data [4]. While this yields a relatively accurate battery range, it is noted that there are uncontrolled variables that can significantly increase or decrease the driving range, such as high-speed interstate driving, uphill driving, inclement weather, and stop and go driving. Additionally, adding weight to the vehicle such as roof racks or heavy cargo, and even opening and closing the windows can cause the aerodynamic drag to increase or decrease. While individually, these uncontrolled variables may cause slight "noise" in the testing data, a combination of these factors can lead to a significant float in the vehicle driving range. Researchers propose unique statistical models, such as a regression analysis, the use of complex energy consumption models, and machine learning models to accurately predict the battery range of the EVs.

This study aims to assess these unique methods of EV driving range prediction, and to analyze the techniques used in developing these statistical models. The various models will be reviewed in hopes of gaining a better understanding of relevant real-world applications of basic statistics, and to better understand some of the intricacies and considerations that go into EV design. By reviewing these methods, the effectiveness of the use of differing statistical prediction methods in achieving similar EV driving range predictions will be analyzed, and a better understanding of statistics applications will be developed as well.

Methods:

In *Energy Consumption prediction for Electric Vehicles Based on Real-World Data* by De Cauwer *et al*, the authors suggest a mathematical energy model based on real world measurements [2]. By developing a formula to describe the total required mechanical energy output at the wheels of an EV, a vehicle dynamics equation can be expressed (Eq. 1).

$$E_{ij} = \frac{1}{13600}[m_{ij}g(f cos\varphi + sin\varphi) + 0.038(\rho C_x A {v_{ij}}^2) + (m_{ij} + m_f)\frac{dv}{dt}]d_{ij}, \qquad (1)$$

Using the energy consumption equation, between distances i and j, the energy consumption of the EV can be calculated. As a result, a linear regression for the energy consumption over time can be approximated (Eq. 2).

$$\Delta E = B_1 * \Delta s + B_2 * v^2 * \Delta s + B_3 * |20 - T| * aux * \Delta t + B_4 * \Delta H_{pos} + B_5 * \Delta H_{neg} + \varepsilon \quad (2)$$

Using available EV data, the energy consumption can be predicted based on this linear regression developed. along with the associated linear regression, relative error and a confidence interval can be established.

In *Driving Range Prediction of Electric Vehicles: A Machine Learning Approach* by Hasib *et al*, the prediction of EV driving range takes a machine learning algorithm approach to the linear regression application [5]. By creating a neural network of multiple linear regressions (MLR), this method aims to relate the single dependent variable of vehicle range to the multitude of independent variables. The MLR model can be expressed by Eq 3.

$$Z = a_0 + a_1X_1 + a_2X_2 + a_3X_3 + \cdots + a_kX_k + \mu \quad (3)$$

Where the dependent variable Z is controlled by each of the predictor terms a_iX_i and the error μ. The neural network takes into consideration vehicle features and processed data to output a predicted vehicle range based on the EV characteristics and the trained network (Figure 1). This model is designed in such a way that each predictor term a_iX_i can be tuned to a particular variable deemed as "data features," such as wind conditions or inclement weather.

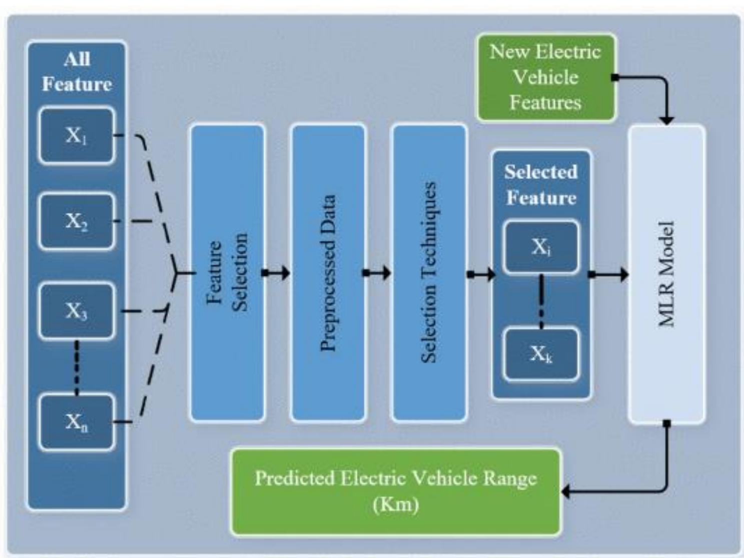

Figure 1: Neural network [5]

Results:

In De Cauwer *et al*, the linear regression coefficients were determined using past electric vehicle data. Acceleration coefficients accounted for almost 20% of the energy usage for the regression model and the rolling resistance contributed 55-78% of the energy usage (Table I). Additionally, positive incline and wind resistance accounted for up to 30% of the energy usage in some data points. Note for this statistical model the regression correlation coefficient (R^2) was 95.6%

Table I: Linear Regression Coefficients of EV Driving Range

Model 1	B_i	Standard Deviation	β_i	Contribution(%)
Constant	1.1E-1	2.65E-2	--	4
Rolling Resistance	1.32E-1	1.84E-3	0.83	78
Aerodynamics	5.00E-6	4.45E-7	0.11	6
Auxiliaries	1.83E-1	7.35E-3	0.14	7
Elevation (+)	3.08E-3	3.72E-4	0.23	26
Elevation (-)	-2.54E-3	3.68E-4	-0.19	-21%

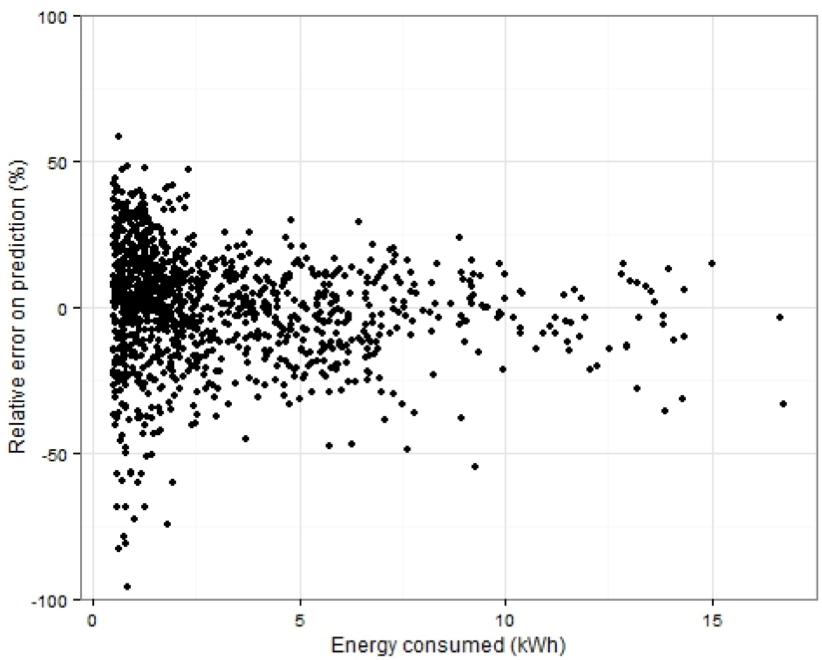

Figure 2: Relative error prediction for 99% CI. [2]

In Hasib *et al,* the MLR model was applied by training the neural network with historical EV data across all automotive manufacturers, as well as vehicle attributes such as brand, price, and number of seats. If the P value of the predictor variable was less than 5%, the predictor was used in the model, otherwise discarded. Using principal component analysis (PCA), the number of independent variables were reduced, in order to reduce the complexity of the network, while maintaining the information of the larger data set (Figure 3).

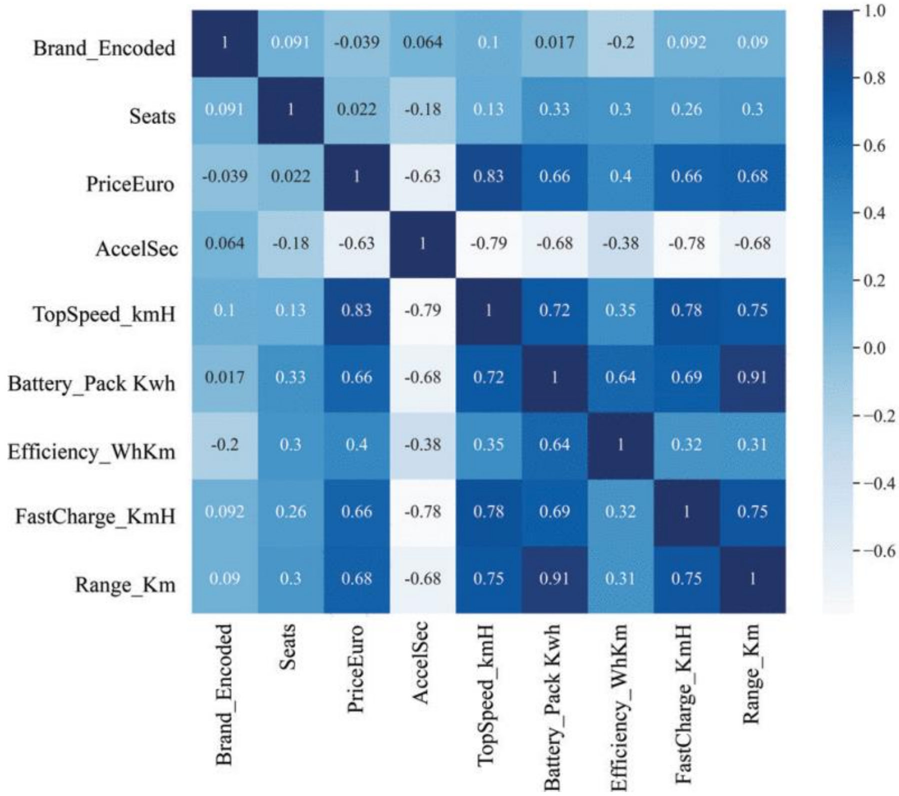

Figure 3: PCA correlation analysis [5]

As a result, the neural network was reduced to 8 arbitrary independent variables trained with the historical data set (Table II). Using the MLR model, a polynomial regression was obtained, with an R^2 value of 97.3% (Figure 4). Additionally, a random forest regression was performed, with R^2 value of 89%, a simple linear regression performed with R^2 value of 82.5%, and a support vector regression yielding an R^2 value of 79%.

Table II: MLR Model Results

	Coefficient	Standard Error	t	P>|t|	[0.025	0.975]
Con-stant	338.6275	2.133	158.725	0.000	334.391	342.864
x1	-4.7656	2.187	-2.179	0.032	-9.108	-0.423
x2	-13.5350	2.934	-4.613	0.000	-19.361	-7.709
x3	-6.9614	2.333	-2.984	0.004	-11.594	-2.329
x4	5.1840	2.502	2.072	0.041	0.215	10.153
x5	23.9997	3.922	6.120	0.000	16.212	31.788
x6	119.7420	4.678	25.595	0.000	110.452	129.032
x7	-50.2471	2.972	-16.908	0.000	-56.148	-44.346
x8	14.9095	3.621	4.118	0.000	7.719	22.099

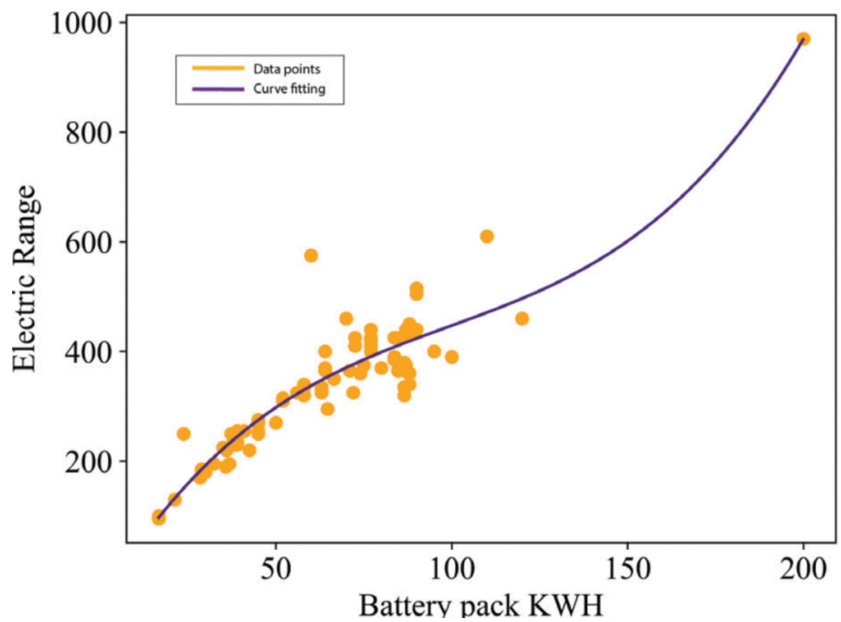

Figure 4: Polynomial regression of EV range [5]

Discussion:

While both methods of producing prediction models for EV range successfully give a mileage estimate, the factors used in each model differ greatly, and are weighted according to each author's own specifications. In *Energy Consumption prediction for Electric Vehicles Based on Real-World Data,* the regression model is successful with a correlation coefficient of 95.6%. It is important to note this linear regression analysis placed a greater emphasis on environmental driving factors such as wind resistance, uphill and downhill roads, and inclement weather. Conversely, *Driving Range Prediction of Electric Vehicles: A Machine Learning Approach* emphasized the data features of individual brands and vehicle types, even conducting a PCA analysis to determine the weighing of the most relevant predictor factors to drive the prediction model.

Despite the two statistical models differing in their methods, they both use linear regressions to predict the range. While a machine learning algorithm can be applied to build a more complex EV range predictor model, using a regression analysis provides the best interpolation of the given data. The greatest difference between models lies in deriving the number regression coefficients and the value of their terms. In other EV range predictor models presented in literature, while the methods may differ, a regression model of some degree is used in predicting and interpolating the EV driving range [1][3].

While these statistical models are effective in accurately predicting EV driving range, these methods are limited in predicting vehicle range due to difficult to model factors such as variable regenerative breaking and constant vehicle motion (i.e. interstate driving), where both can greatly reduce the energy used by the vehicle. Despite these limitations, both models include error terms to account for general noise that is difficult to account for. As a result, while the predicted range may vary slightly from driver to driver depending on their driving habits and conditions, the regression models provide a valid method of EV range prediction given real world data.

Conclusion:

The use of statistics aids in the analysis and sorting of real-world data. Given the real-world data collected by EVs, regression models can be developed to predict the effective driving range. Whether by a machine learning algorithm, or a complex linear regression, this data can be processed to develop these models. These models illustrate the value of even the most basic of statistical regression analyses in modern day data analysis methods, and how even a simple linear regression model can be used in the intricate design of EVs.

Bibliography:

[1] A. Bolovinou, I. Bakas, A. Amditis, F. Mastrandrea, and W. Vinciotti, “Online prediction of an electric vehicle remaining range based on regression analysis,” *2014 IEEE International Electric Vehicle Conference (IEVC)*, 2014.

[2] C. De Cauwer, J. Van Mierlo, and T. Coosemans, “Energy consumption prediction for electric vehicles based on real-world data,” *Energies*, vol. 8, no. 8, pp. 8573–8593, 2015.

[3] J. Zhang, Z. Wang, P. Liu, and Z. Zhang, “Energy consumption analysis and prediction of electric vehicles based on real-world driving data,” *Applied Energy*, vol. 275, p. 115408, 2020.

[4] “Range Tips,” *Tesla Support*. [Online]. Available: https://www.tesla.com/en_eu/support/range#:~:text=Displayed%20range%20in%20your%20Tesla,the%20onboard%20computer%20calculates%20range. [Accessed: 24-Mar-2023].

[5] S. A. Hasib, D. K. Saha, S. Islam, M. Tanvir, and M. S. Alam, “Driving range prediction of electric vehicles: A machine learning approach,” *2021 5th International Conference on Electrical Engineering and Information Communication Technology (ICEEICT)*, 2021.