Article

Estimation Model of Total Energy Consumptions of Electrical Vehicles under Different Driving Conditions †

Seyed Mahdi Miraftabzadeh *, **Michela Longo** * and **Federica Foiadelli**

Department of Energy, Politecnico di Milano, Via La Masa, 34, 20156 Milan, Italy; federica.foiadelli@polimi.it

* Correspondence: seyedmahdi.miraftabzadeh@polimi.it (S.M.M.); michela.longo@polimi.it (M.L.);
  Tel.: +39-02-2399-3759 (M.L.)

† This paper is an extension of the conference paper presented at IEEE EEEIC 2020, Madrid, Spain,
  9–12 June 2020.

Abstract: The ubiquitous influence of E-mobility, especially electrical vehicles (EVs), in recent years has been considered in the electrical power system in which CO_2 reduction is the primary concern. Having an accurate and timely estimation of the total energy demand of EVs defines the interaction between customers and the electrical power grid, considering the traffic flow, power demand, and available charging infrastructures around a city. The existing EV energy prediction methods mainly focus on a single electric vehicle energy demand; to the best of our knowledge, none of them address the total energy that all EVs consume in a city. This situation motivated us to develop a novel estimation model in the big data regime to calculate EVs' total energy consumption for any desired time interval. The main contribution of this article is to learn the generic demand patterns in order to adjust the schedules of power generation and prevent any electrical disturbances. The proposed model successfully handled 100 million records of real-world taxi routes and weather condition datasets, demonstrating that energy consumptions are highly correlated to the weekdays' traffic flow. Moreover, the pattern identifies Thursdays and Fridays as the days of peak energy usage, while weekend days and holidays present the lowest range.

Keywords: electric vehicles (EVs); big data; smart grid; energy demands; estimation model; energy consumption

1. Introduction

The green city concept is one of the realities of the future. The interactions between cars, houses, trains, or any other component of the smart city, are not the only contributing factor in the formulation of this concept; from an energetic point of view, it is necessary that the components use green energy not to pollute the environment [1,2]. The emission reduction targets motivate manufacturers to invest more in electric mobility and lead to rapid growth in electric vehicles (EVs). Therefore, electric mobility is one of the most critical elements that should be considered and investigated more deeply. Even in the era of electric vehicles, the problem of range anxiety has pushed back people from adopting this technology. People have developed a low level of trust in the EVs' battery capacity, causing uncertainty for reaching their destination [3].

Nevertheless, to invest in EV consumption behavior and prediction algorithms, it is necessary to gain people's trust to popularize EVs usage. The main problem of range prediction is obtaining values from a colossal amount and variety of data generated at a high frequency. Some of the values are produced by the EV itself, while the others are from different sources. This problem of volume, variety, and velocity leads us to use big data techniques to better predict energy consumption [4]. In the near future, most of the conventional cars will be substituted with electric versions, and based on the Bloomberg executive report, more than 54% of new car sales across the globe are EVs or hybrid electric vehicles (HEV) [5,6]. Therefore, it is essential to know both EVs' energy consumption and

charging patterns in a city to tune the electrical grid productions in order to continuously provide energy and prevent any disturbances. Estimating the total EV energy demand is a complicated issue since it involves many different factors, such as human behavior, city infrastructure, electricity price, chosen route, charging time, working schedules, etc. Moreover, there are currently very few EVs in the cities to study their total demand behavior, and their datasets are usually protected by company policies or customers' privacy. Thus, there is a need to make a new strategy to use other public data to study the cars' movement patterns, so as to link them to a massive penetration of EVs in the future and develop a new estimation model in order to tune electricity generation to supply the EVs and prevent disturbances such as generator and branch tripping, bus faults, sudden large load changes, or blackouts [7–10].

This paper provides a new strategy and algorithm to model and estimate the electricity demand based on the large-scale real travel trajectory data of public fleets and weather condition data. The proposed algorithm and method are applied to the New York City taxi fleet and weather conditions datasets. Although results are achieved based on the taxi fleet examination in New York City, the method and algorithm can be implemented for private vehicles in any city with similar travel trajectory data. The remainder of the article is structured as follows: Section 2, state of the art with the study of energy consumption prediction models; Section 3, the methodology which is used, big data and its concepts, are presented. In contrast, Sections 4 and 5 are allocated to results, discussion, and conclusion.

2. State of the Art

This section reviews related state-of-the-art works in recent years that sustain the necessity of developing a new method for estimating EVs' total energy consumption, such as the one presented in this article.

Seungwoo Jeon and Bonghee Hong have proposed a new statistical modeling method that finds the best historical dataset to accurately predict the traffic flow by day of the week [11]. Using the Intelligent Transportation System (ITS), it is possible to know the location of currently congested roads and find the shortest routes by using real-time traffic data. However, the more difficult job of traffic flow prediction is to predict the traffic flow of the next day of the week; it depends upon an irregular pattern of traffic flow which may be occurred repeatedly. For a more accurate forecast, the authors argued that historical traffic data and suitable time series forecasting methods are necessary. They proposed a three-step filtering algorithm to find the optimal historical data range for the same time slot using optimal parameters to select data. The optima time-series forecasting method was chosen based on the minimum mean absolute Ppercentage error (MAPE) metric applied to a combination of historical and simulation data. Because both methods involve big data processing, they constructed a big data processing framework to handle the overall prediction process and calculate large traffic data. The prediction outcomes reveal that the proposed model can obtain a high forecasting precision for each road in their increasing dataset. Having more accurate traffic data is entirely related to and helpful for predicting the energy consumption of electric vehicles throughout the week, and for having a better insight to know where to put the charging stations within the city, and for better prediction of electrical consumption patterns.

In [12], Mariz B. Arias and Sungwoo Bae (2016) used historical traffic data and weather data of South Korea to formulate the forecasting model. The considered variables in this study were the charging starting time determined by the real-world traffic patterns and the initial state-of-charge of a battery. The different charging load profiles of electric vehicles in the residential and commercial sites were shown by studying electric vehicles' charging demand during weekdays and weekends in summer and winter. Furthermore, in [13], electric cars and electric buses were considered in forecasting the EV charging demand. This paper, which examined both slow and fast charging classifications, proposed an EV charging demand forecasting method based on big data, including real-world traffic distribution data and weather data collected in South Korea every hour. Aditionally, this

study developed an EV charging demand forecasting program of which the technical design architecture is based on MATLAB built-in functions. The technical architecture of this program has four layers, including data sources (collections of tabular text files which contain historical traffic volume data and weather data), data storage (it can provide data in a chunk-wise manner for quick and efficient access to the target data), data management (manages the data stored in the previous layer), and data processing. This last layer, big data handling processes, includes a cluster analysis to classify traffic patterns in each division, a relational analysis to identify influential factors affecting the traffic patterns, and a decision tree to establish classification criteria. The presented forecasting model may allow power system engineers to anticipate electric vehicle charging demand based on historical traffic data and weather data. Therefore, the proposed electric vehicle charging demand model can be the foundation for the research regarding the impact of charging electric cars on the current power system.

In [14], a stochastic model is proposed to enable the optimal charging of EVs in the presence of renewable energy sources. The EVs are not considered individually but are clustered into fleets. Only the energy demand of the fleet is considered. A model to estimate a fleet of plug-in electric vehicles energy demand is proposed in [15]. This model is used to evaluate the impacts of charging in distribution networks. This work uses the Monte Carlo algorithm to generate vehicles' trips according to users' behavior, types of vehicles, periods of charging, and the temperature effect. In [16], a regression model is presented to estimate the travel time.

In [17], a Fuzzy Logic (FL) control strategy is proposed for an EV energy management system with dual-source power (battery and supercapacitor). This architecture is proposed to satisfy EVs energy requirements, improving both the EVs efficiency and the overall performance of the system. A regulatory framework and business models for network operators and EVs aggregators are described in [18] regarding opportunities for vehicle-to-grid (V2G) applications, such as peak power and frequency regulation. Models regarding EVs' bidding and optimal charging strategies are proposed in [19]. Although the research work presented in [20] considered EV model characteristics, mobility patterns, charging processes, and social and economic variables, it did not consider the real-world data of weather and traffic volumes, which may change the EV charging demand. A recent study [21] also presented different forecasting methods based on historical charging data; these datasets include charging records from customer profiles prone to privacy invasion issues and station measurement records containing a large volume of data requiring a long processing time. On the other hand, this paper presents an EV charging demand forecasting method that considers real-world traffic volume data with weather conditions that can resolve tprivacy invasion issues and data processing speed concerns, as seen in the previous research works, also taking into account other variables. In [22], vehicle speed, acceleration, and roadway grades were used to model EV energy consumption. However, most studies considered driving patterns for predicting EV charging demand.

In [23], the authors estimated that each EV unit consumes on average 7.29 kWh since, in South Korea, the average daily mileage per vehicle is 73.9 km, and EVs' fuel efficiency is 6.02 km/kWh. Then, they developed different scenarios based on two days of survey data of vehicle usage of 418 adults and their preferred time of day to charge their car. Finally, the market expansion scenarios concluded that the average daily electricity required was 194–447 MWh.

Gomez-Quiles et al. [24] used connection attempt real data—data of an EV connected to a charging station regardless of its state of charge (SoC)—to measure EVs' power consumption based on geographical zones each hour. This information was then used to perform hourly EV power consumption predictaion via ensemble learning with ARIMA, GARCH, and PSF algorithms for different zones.

Morlock et al. [25] developed an algorithm to predict a speed profile for a given route based on a real-time traffic dataset in order to forecast an EV's energy consumption using an EV powertrain loss model for two driving styles (desired cruise speed, and acceleration and

deceleration behaviors). Such a model considers the formula of the required kinetic power, stationary power loss, and auxiliary power demand based on the ambient temperature.

In [26], information of a gas station and a month of a taxi fleets' travel patterns were used to promote the development of charging infrastructure. A large number of studies focus on energy consumption and charging behaviors of a single EV; nevertheless, a few have been explored how these behaviors would change for larger penetration of EVs, or how the other source data could be paired and used with the EVs data to better understand energy consumption patterns [27].

In the literature, several works studied individual electrical vehicle energy consumption and modelled prediction models with reasonable accuracy. However, none of them considered the scenario concerning the amount of energy needed when the majority of cars in the same area are electrical. Therefore, this study provides a new idea and strategy to estimate the total energy consumption of EVs. In other words, the main differences between previous researchers and this work are as follows:

1. An estimation of how much energy is needed at the desired intervals of the day, month, and year, to adjust the rate of power patterns and prevent any disturbance in the power system network;
2. Useful to find the critical points of a city to locate charging points based on needs by following taxi routes and their consumptions;
3. Identifying required charging types (regular, fast, superfast, wireless) according to energy usage patterns;
4. Since this model uses the fast-response big data platforms and algorithms, it is possible to feed more data and obtain a more accurate estimation.

3. Materials and Methods

Figure 1 outlines the overall model framework employed in this study. The model uses three different datasets: taxi trajectory, weather, and EV datasets. In the first phase, the average speed was calculated based on each trip's start and end time and distance covered; the corresponding average temperature was determined from the weather dataset. After obtaining the EV consumption curve from the EV database, the remaining range was selected to calculate total energy consumptions in a chosen window. In the final stage, based on the desired window width and time interval, for example for each minute for each month of a year, the total energy consumption was computed using the proposed algorithm. The dependency of EV power consumption on the driving speed, temperature, auxiliary services, traffic, and covered distance, which are the most influential factors, are taken into account in this framework [28]. In the rest of the article, the detailed descriptions for each step are discussed.

This study manages datasets coming from the routes of the 13,587 existing taxis in New York City as a case study, containing a massive amount of information organized in 103 million rows and 18 columns for 2018 (volume property of big data). Analyzing this amount of data with traditional methods such as spreadsheets is almost impossible and needs a robust computational system; for example, it is difficult to load the whole amount of this size of data in Microsoft Excel, or it would take a lot of time to perform each change analysis step in modeling. Thus, two of the free data science softwares that are usually used to handle big data problems are used for this study: Python and KNIME. Moreover, it is much more comfortable with regard to big data platforms to access and use diverse managing and modeling techniques, such as machine learning algorithms to analyze, model, and gain useful insights efficiently out of the massive amount of information received. The KNIME provides a user-friendly interface based on drag and drop tools; additionally, after each stage, it nicely demonstrates the result [29]. However, Python [30] adds more flexibility for users to speed up heavy computational tasks and improve their algorithm's performance by applying different packages such as Panda, TensorFlow, etc. Therefore, the core parts of this paper's proposed model are written in Python, and some parts of data manipulation and visualization are conducted in KNIME.
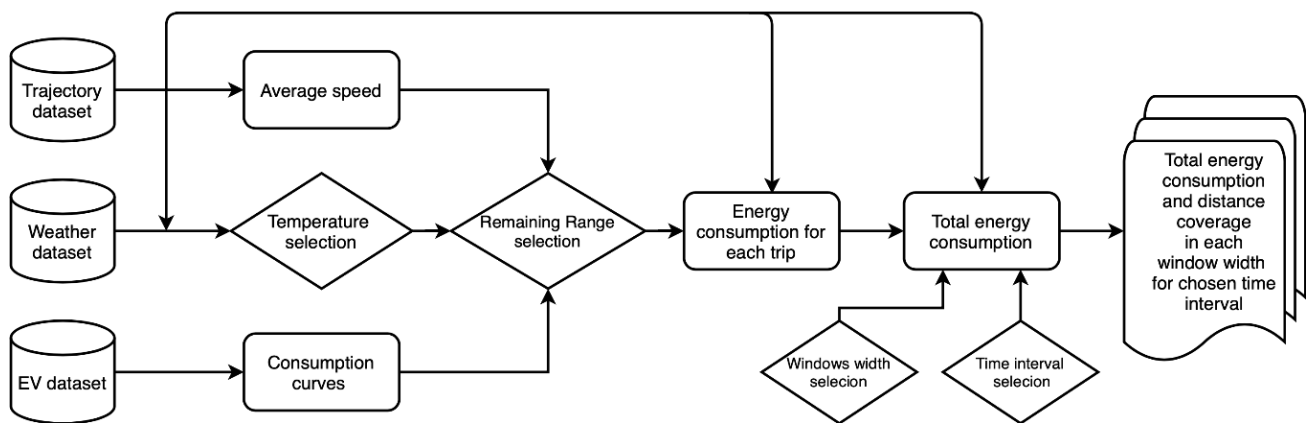
Figure 1. The model framework.

3.1. Datasets

The proposed method and algorithm was tested by using internal combustion engine (ICE) taxis as a probe to model human fleets on different days of a year, to provide us with the following information: the distance between the origin and destination point with corresponding coordinates, number of passengers, pick up (start) and drop off time (end), districts, and so on, for every time someone used the taxi service. The required time to cover the routes depends on real-time traffic and other limitations. Considering the weather condition correlation to energy consumption, especially for EVs' battery efficiency, the independent weather dataset for New York City was acquired. Next, two datasets were merged based on the corresponding time of each taxi's route to feed the next stage of the process. This dataset was used to calculate if taxis were EV ones and how much energy they would consume to reach their destinations. Finally, the proposed algorithm introduced in Section 3.3 is used to estimate energy consumption in the desired time interval and window width.

3.2. Big Data Methodology

The term big data is used to refer to any dataset that is difficult to manage using traditional database systems. Several years ago, the Gartner group noticed this undue attention focused on size and proposed the now-famous "3Vs" of big data [31]. Big data is a high-volume, high-velocity, and high-variety information asset that demands cost-effective, innovative forms of information processing for enhanced insights and decision making [32]. A similar definition is provided in [33]. However, IBM tried to add a 4th and 5th V to big data: veracity and variability [34]. Currently, in the data science community, big data is considered to have a high volume, velocity, variety, and veracity (or at least the first three) of data—both structured and unstructured—that overwhelms a business on a day-to-day basis. Figure 2 shows a general view of the big data or data science process steps. It can be dissected for insights that lead to better resolutions and strategic business progress.

In particular, the different steps are:

1. Acquiring data: Improving analysis capabilities is obtained by available dataset identification, data retrieval, and data query. The result of receiving data from different sources (individual, structured, unstructured data, etc.) enhances the procedure of finding the correlation between variables, enhancing model efficiency, and providing more valuable insights. In this work, the weather dataset comes from the National Oceanic and Atmospheric Administration [35], and the taxi fleet information is acquired from the official website of the City of New York [36].

2. Preparing data: This phase demands a considerable amount of time to proceed correctly. This is a very critical step to secure a purposeful analysis. The result of preparing the data is to eliminate any percentage of error and to adjust the data for the next steps. Knowing the nature of the data and applying statistical analysis, it is

possible to obtain insight on how to deal with missing values, invalid records, outliers, or duplicate values. After exploring and pre-processing data in this step, its quality increases significantly, and it will result in a suitable structure for a better model.

3. Analyzing data: At the end of this step, an accurate model is built in order to enhance business success. Having only a strong background in statistics is not enough to handle high volume sets of data. State-of-the-art machine learning techniques, such as Neural Network or Regression Tree, address the meaningful classification, clustering, regression, association analysis, and graph analytics on the scale of big data problems. After describing the correlation between the data, the developed business outcomes are obtained, followed by choosing and establishing the appropriate analytical technique and creating the model that better fits to the data and problem. Since each step of this methodology is scalable, if, based on accuracy metrics, the results of the mathematical model are meaningless or irrelevant, data scientists repeat the analysis and put more thorough attention to details.

4. Reporting insights: There are different kinds of visualization tools that can be used to simplify the presentation of information to the public, individuals, or companies. The display must make sense and be easily understandable for communication. It is very challenging to present results conventionally. Thus, in this work, the authors used various graphs instead of tables in order to make it more straightforward to be followed.

5. Insight into action: The value to be derived from this methodology lies in a procedure capability to turn customer insights into actionable decisions that boost business opportunities. When the question is reasonably understood, evaluation of the result gives the idea of needing to return to some previous steps or that real-time action should be addressed. After this step, the energy consumption patterns are reasonably identified and justified with human behaviors during working-days and holidays.

This methodology is considered an iterative procedure, since after each step, it is possible to face some outcomes that were not expected or do not align with accuracy metrics. For instance, after cleaning and analyzing the data of the proposed model, 1 million rows of information were automatically omitted by the algorithm applied because it contained false values or was missing some critical information due to incorrect data acquisition.
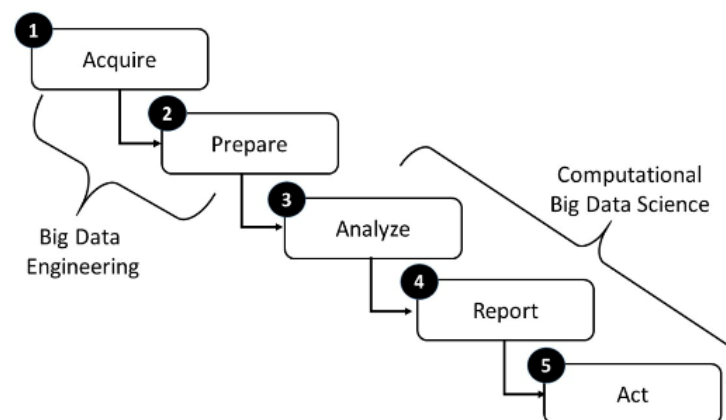

Figure 2. Big data procedure.

3.3. Methods and Model

In this model, each taxi cab is substituted with the specific electric vehicle characteristics in order to calculate energy usage in each route. Finally, by aggregating all the energy consumption, the algorithm provides the total energy consumption at any selected time interval.

In this study, as it is presented in Figure 3, three different datasets were used:

1. Taxis dataset (Yellow New York taxi).

2.　　Weather dataset (New York city weather).
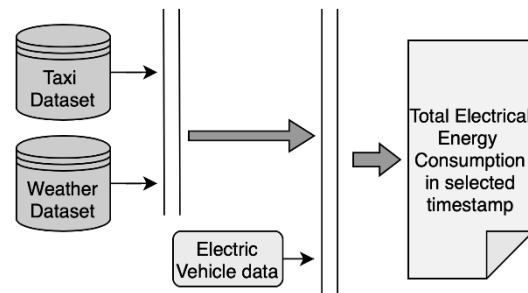3.　　Electric vehicle dataset (Renault Zoe).

Figure 3. The datasets' schema.

After cleaning the datasets, they were merged with the same date and time to be ready for the next step. Finally, each trip's energy consumption based on electric vehicle specifications (such as the remaining range based on battery capacity, speed, and temperature) is calculated. The total energy consumption of all trips based on selected time intervals and the window width is obtained using the proposed algorithm. One of the main advantages of this algorithm is that it does not need to wait until the end of the route to determine energy consumption. In other words, it follows all changes in real-time and provides us with the ability to tune timestamps at any selected time interval.

The range (R) [km] based on average speed (V) [km/h] for selected temperatures [°C] for the specific EV that was used in this study is presented in Figure 4. The study assumed that the battery was new and full (SoC is equal to 100%), and did not lose its capacity after charging.
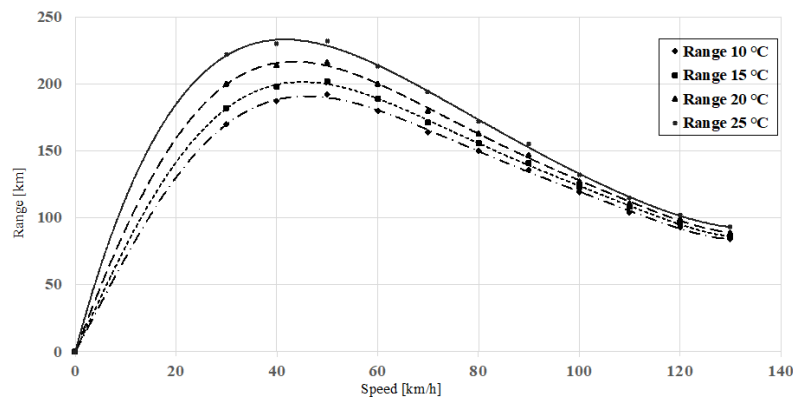

Figure 4. Electric vehicle characteristics: range vs. speed vs. temperature.

The lines for each temperature are presented by polynomial linear regression with a degree of seven, which has the maximum R-squared value, with an average of 0.998, suggesting that this curve is the most similar to the real one. Equation (1) gives the R for 10 °C with respect to average speed (V), while (2)–(4) relate to the temperatures of 15 °C, 20 °C, and 25 °C respectively:

$$R = 8.47 \times 10^{-4} + 6.97 \cdot V + 5.16 \times 10^{-2} \cdot V^2 + (-5.09 \times 10^{-3}) \cdot V^3 + 7.71 \times 10^{-5} \cdot V^4 + (-4.85 \times 10^{-7}) \cdot V^5 + 1.14 \times 10^{-9} \cdot V^6 + (-1.11 \times 10^{-11}) \cdot V^7 \quad (1)$$

$$R = -0.15 + 6.29 \cdot V + 1.81 \times 10^{-1} \cdot V^2 + (-1.06 \times 10^{-2}) \cdot V^3 + 1.84 \times 10^{-4} \cdot V^4 + (-1.57 \times 10^{-6}) \, V^5 \cdot + 6.7 \times 10^{-9} \cdot V^6 + (-1.14 \times 10^{-11}) \cdot V^7 \quad (2)$$

$$R = 0.06 + 7.62 \cdot V + 1.65 \times 10^{-1} \cdot V^2 + (-1.14 \times 10^{-2}) \cdot V^3 + 2.10 \times 10^{-4} \cdot V^4 + (-1.88 \times 10^{-6}) \cdot V^5 + 8.41 \times 10^{-9} \cdot V^6 + (-1.51 \times 10^{-11}) \, V^7 \quad (3)$$

$$R = -0.12 + 6.30 \cdot V + 1.89 \times 10^{-1} \cdot V^2 + (-1.10 \times 10^{-2}) \cdot V^3 + 1.93 \times 10^{-4} \cdot V^4 + (-1.68 \times 10^{-6}) \cdot V^5 + 7.36 \times 10^{-9} \cdot V^6 + (-1.29 \times 10^{-11}) \cdot V^7 \quad (4)$$

These equations help us calculate each desired speed range quickly because they are less than 55 mph, or 90 km/h, the maximum speed limit in New York City [37]. On the other hand, in a city such as New York with high traffic congestion, taxis' average pace is much less than 90 km/h. However, the proposed algorithm is not affected by EVs power consumption models, thus the proposed model can work with any EVs, or work with different versions of EVs at the same time so as to calculate the total energy consumptions for each window width.

Knowing the range, speed, distance, and battery capacity, it is possible to calculate the energy consumption for each route individually using Equation (5), which is identified by $A_i\{energy\}$ in the dataset:

$$dE_i = B_j(capacity) \cdot (dS_i/R_j) \tag{5}$$

where dE_i and dS_i are the energy consumption and distance covered in ith trip, B_j and R_j are the battery capacity and the range of jth electric vehicle model with full battery capacity for the jth electric vehicle model.

If the window width were equal or bigger than the interested time interval, total energy consumption would be simply achieved by adding all dE_i together; however, this is not always the case, and in order to do so, Equation (6) should be used:

$$E_{total} = \sum dE_i \cdot dt_i \tag{6}$$

In this equation, the dt_i is the ratio of the trip duration that is inside the window to window width. In order to calculate the corresponding dt_i, four scenarios and algorithm are defined in this study.

Before moving forward to link the trajectory of each taxicab to the total energy consumptions, there are a few terms that need to be defined to determine overall energy usage at each time interval. In this study, the time interval refers to the requested period, which might be in a day, month, or a year. Whereas, a window width is the length of the moving window over the time interval; it is possible to consider it as the desired precision of the calculation for the time interval to illustrate the total energy consumption for the chosen interval (seconds, minutes, hours, or days). The maximum accuracy or minimum window width (W) can be in the order of the seconds, due to available dataset configurations, and it is recognized by starting time (T_S) and finishing time (T_F). $A_i = \{t_p, t_d, \ldots\}$ describes each row of the dataset or route, which contains pick-up (t_p) and drop-off (td) timestamps—or the start and end of the trip for a private EV—distance, etc.

The four different scenarios contribute to each window for calculating the energy consumption, while the window is swept in the time interval, as shown in Figure 5. In this figure, the starting or ending round arrows means that the pick-up or drop-off time for the corresponding route (or row of data) is inside the window, and when it is otherwise, it is represented by dashed lines. For the sake of simplicity, suppose energy consumption of a route in every condition is proportionally linear to time. Condition one occurs when both pick-up and drop-off of a passenger(s) is inside the window's width, and condition two relates to a pick-up timestamp inside a window's width, while a drop-off timestamp is after the upper boundary of a window. The situation where the pick-up of a passenger(s) is before the lower limit of a window and the drop-off is inside the window refers to condition three. The last condition (4) includes a scenario that a taxi overpasses a window entirely, where the pick-up time is before T_s and drop-off after T_F.

The Equations (7)–(10) formulate the above conditions in the mathematical form in order to be applied to the model, in which x is a taxi event or row of the dataset matrix A:

$$Cond\ 1 = \{x \in interval \mid T_S < x(t_p, t_d) < T_F\} \tag{7}$$

$$Cond\ 2 = \{x \in interval \mid T_S < x(t_p) < T_F \wedge T_F < x(t_d)\} \tag{8}$$

$$Cond\ 3 = \{x \in interval \mid x(t_p) < T_S \wedge T_S < x(t_d) < T_F\} \tag{9}$$

$$Cond\ 4 = \{x \in interval \mid x(t_p) < T_S \wedge T_F < x(t_d)\} \tag{10}$$
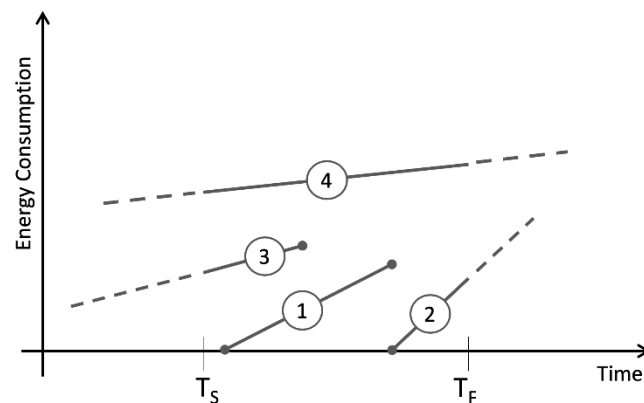
Figure 5. Different conditions in the window width.

Henceforth, we need a fast response algorithm to take into account all the specific scenarios for each window until it covers the time interval completely. Additionally, in order to follow the big data methodology in Figure 2, this study developed a model in Python to manipulate data (cleaning and preparing data) and to perform the analysis. The proposed algorithm, as it is shown in Algorithm 1, considers all four scenarios to calculate total energy consumption in any desired window's width and time interval. The vectorization property of the Python package, Pandas, dramatically speeds up the computational procedure time, which gives leverage to analyze the big data scaled datasets in this study [38].

Algorithm 1: Calculating total energy consumption in tuned window width in the coveted time interval

---

1: procedure
2:     System initialization:
3:     $W \leftarrow$ *Window's width*
4:     $T_F \leftarrow$ *Lower interval boundary*
5:     **while** $T_F \leq$ *Upper interval boundary* **do**   • stopping condition
6:         $T_S \leftarrow T_F$ • move the window
7:         $T_F \leftarrow T_S + W$
8:         **if** $t_d < T_F$ **then**
9:             **if** $t_p \geq T_S$ **then**
10:                 $Energy_{total}\; +=\; \sum\limits_{A_i \in\; Cond\; 1} A_i\{Energy\}$
11:             **else if** $t_d > T_S$ **then**
12:                 $Energy_{total}\; +=\; \sum\limits_{A_i \in\; Cond\; 3} (A_i\{t_d\} - T_S) * A_i\{Power\}$
13:         **else**
14:             **if** $t_p < T_F$ **then**
15:                 **if** $t_p \geq T_S$ **then**
16:                     $Energy_{total}\; +=\; \sum\limits_{A_i \in\; Cond\; 2} \left(T_F -\; A_i\{t_p\}\right) * A_i\{Power\}$
17:                 **else**
18:                     $Energy_{total}\; +=\; \sum\limits_{A_i \in\; Cond\; 4} (W) * A_i\{Power\}$
19:     **return:** $Energy_{total}$ • total energy consumption for each window

---

Algorithm 1 starts to determine A_i relating to each condition after tuned of its initialization variables (interval time and width of the window), in which $A_i\{Power\}$ is the ith row of the dataset matrix A corresponding to the calculated power column, and $A_i\{t_d\}$ and $A_i\{t_p\}$ are the ith row of the dataset matrix A corresponding to the pick-up and drop-off, respectively. Then it starts to calculate the corresponding energy for each Ai in every different scenario and makes the summation of them before reporting the total. The main

advantage of this algorithm is that it follows every taxi's routes and correctly calculates their energy consumption. As a matter of fact, EV energy consumption adds uncertainty to the total calculation, as well as data acquisition mistakes. However, since the later uncertainty has random behaviour, normal distribution does not change the identified trends of changing data during the time.

4. Results and Discussion

The results of the presented model show new insights about the future energy consumption of EVs in vast usage. Figure 6 shows the energy consumptions for each month of 2018. The order of energy usage is between 6 to 8 thousand megawatts hour. It is shown in the histogram that the highest energy usage is for April, 7655.18 MWh, and the lowest for July, 6148.67 MWh, with an average of 6946 MWh of energy usage per month, represented by the red dashed line, and 83,352 MWh for the total year. Based on the data analysis, the difference usage in months demonstrates that fewer or more people used taxi services due to different reasons, such as weather conditions, holiday pause, summer break, etc.
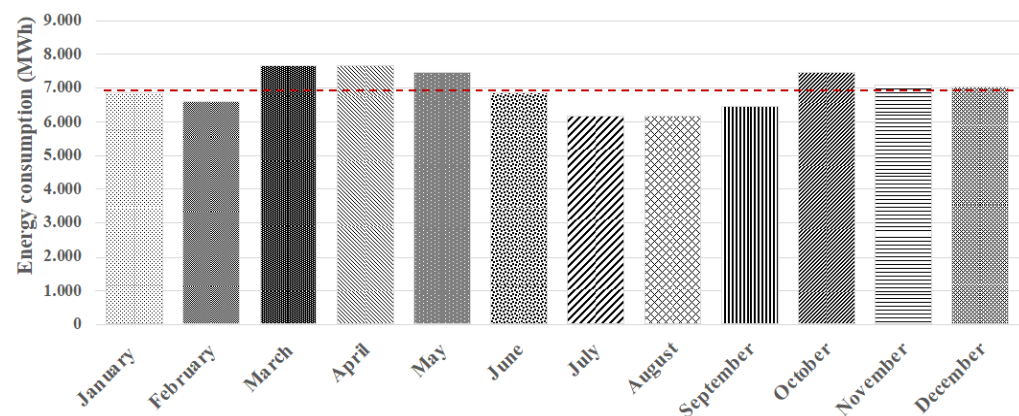

Figure 6. Energy consumption for each month of 2018, red line demonstrates the average.

It is supposed that an EV roughly consumes between 0.15 to 0.2 kWh per kilometer igure [39,40]. For example, in November, all taxis had covered 38,147,929 km, and if we consider them as an electric model, based on this rough estimation they had consumed 5,722,189 MWh to 7,629,585 MWh. The proposed model calculated 710,708 MWh—in the acceptance range—which validates the model's accuracy. Therefore, the proposed model not only estimates more accurately the total energy consumption of EVs, but also retrieves the usage pattern, which is very useful for the electrical grid side.

The maximum and minimum energy consumption value in the first week of December 2018 is demonstrated in Figure 7. In this radar, the inner-circle has the minimum amount, and the outer circle relates to the maximum range, 20 MWh. In this graph, the time intervals are four days, and the window's width is tuned to an hour. The first window is from midnight to 01:00, the second one from 01:00 to 02:00, and so on.

On the one hand, Figure 7 shows how the energy consumption pattern changed in each hour during the day, where the minimums occurred in the window width of 04:00–05:00, while the values of the maximum were achieved in the period from 18:00 to 23:00. Moreover, energy consumption slowly decreased from midnight until 06:00, whereas on weekdays, it rapidly increased in a shorter period, from 06:00 to 08:00. Later on, it reached the peak average around 09:00 and remained in the outer circles until 23:00.
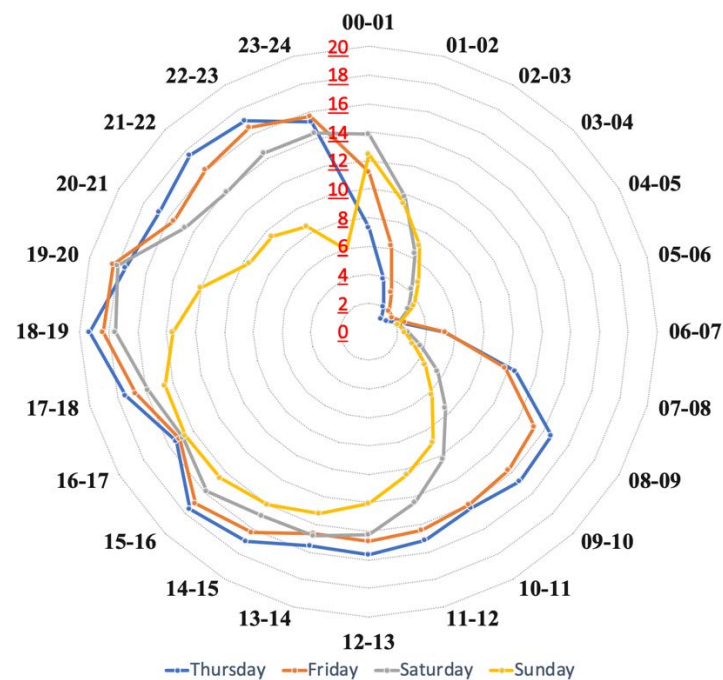
Figure 7. Hourly energy consumption for Thursday (6 December), Friday (7 December), Saturday (8 December), Sunday (9 December) in the radar representation graph with the scale of a megawatt-hour (MWh).

Additionally, it shows the various energy consumption rates on different days of the week in terms of patterns and quantities. As it is explained later, generally, the maximum usage during a week occured on Thursdays and Fridays; consequently, energy consumption on these days remained in outer circles, from 16 to 20 MWh. Additionally, weekend days, particularly Sundays, were identified as the minimum rated days of a month in which from 06:00 to 12:00 the energy consumption slowly increased to its average peak.

These patterns could help to identify the best charging solutions (regular, fast, super-fast, wireless) considering the energy usage during a day. For example, during the hours that showed less energy consumption, from midnight to 06:00 or 07:00 am, it is supposed that people would be charging their cars. For this period, the regular charging types are sufficiently enough. However, if any electrical taxi needs to re-charge before the peak hours (starting from 18:00), the supercharge solutions and/or charging wireless models should be integrated around the city and along main routes with traffic congestion. On the grid side, this information not only gives a general view of total electrical energy demand (for charging EVs) that must be provided constantly during a day, but also estimates hours in which EVs would need to re-charge based on their energy consumption patterns. Therefore, the electrical grid schedules generators and tunes demand trends accurately to avoid any disturbances in the power system or blackouts in the city.

Figure 8 shows the energy usage in the winter season, the first three months of the year, for each day. As known, the length of the months varies from 28 to 31 days. Moreover, similar analyses were performed for the rest of the months, as shown in Figures 9–11 for spring, summer, and autumn seasons, respectively. These graphs show the energy consumption patterns based on weekdays, as seen in Figure 8, where the minimum range in the month of June occurred in the 3rd, 10th, 17th, 24th, and 30th day, which are Sundays. On the other hand, the maximums were achieved on Thursdays and Fridays, demonstrating the presence of more taxi fleets on service. The further valuable insight coming from this analysis shows energy consumptions during holidays with a lower range compared to any other days, even in the case of Thursdays or Fridays. For instance, on the 12th and 22nd of November (respectively Veterans Day and Thanksgiving, American holidays), as presented in Figure 10, the energy consumption was comparably the lowest rate. Therefore, this

study concludes that generally, the maximum energy consumption occurs on Thursdays and Fridays, while minimum consumption occurs on weekend days and other holidays.
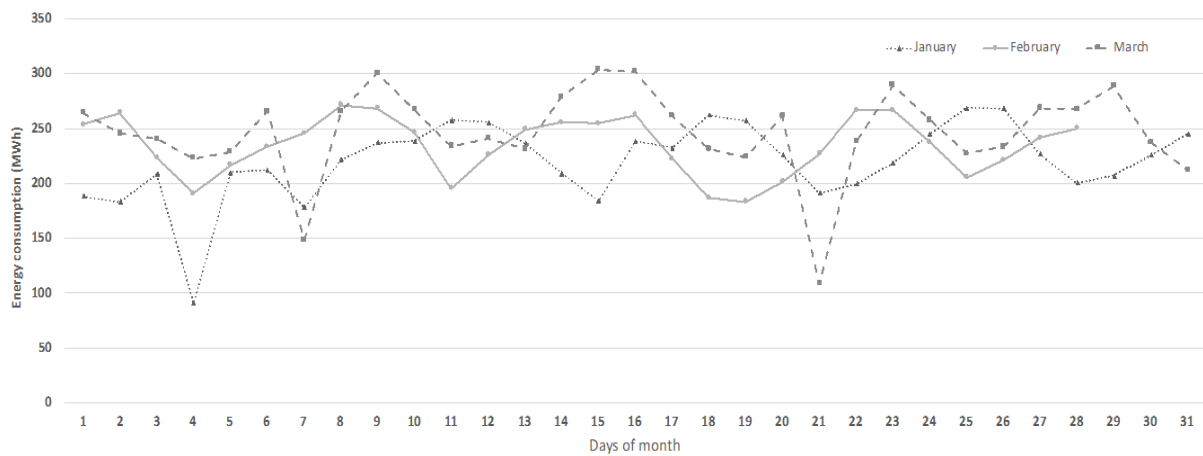

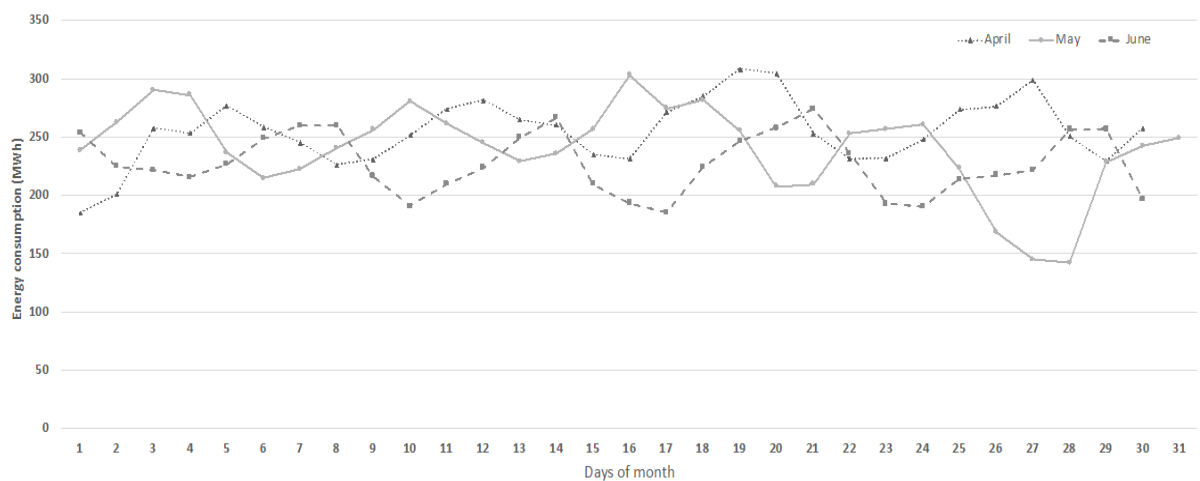**Figure 8.** Energy consumption in winter, 2018.

Figure 9. Energy consumption in spring, 2018.

The presented algorithm in this study is capable of calculating energy consumption at any window width and timestamp intervals. The energy usage per each minute term shows that rush hour is between 18:00 to 19:00, the end of most working days, as demonstrated in Figure 12. Figure 11 displays the outcomes of the model in a window's width tuned per minute, along the randomly selected day: Thursday, 6 March 2018. On the contrary, the period between 02:00 to 07:00 shows the lowest outcomes; however, on some special days such as New Year's Eve, these minimum values could move forward because of celebration events.
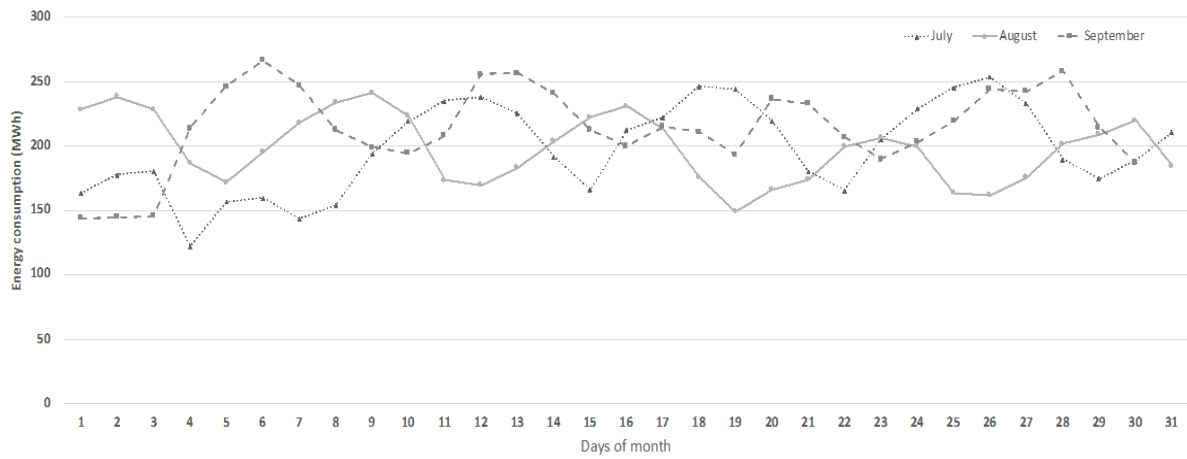
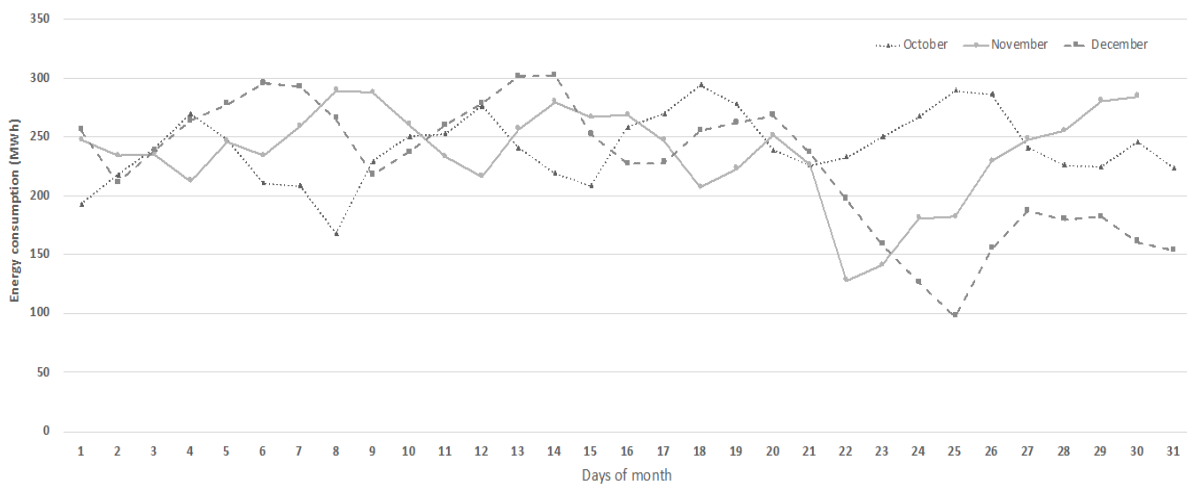**Figure 10.** Energy consumption in summer, 2018.

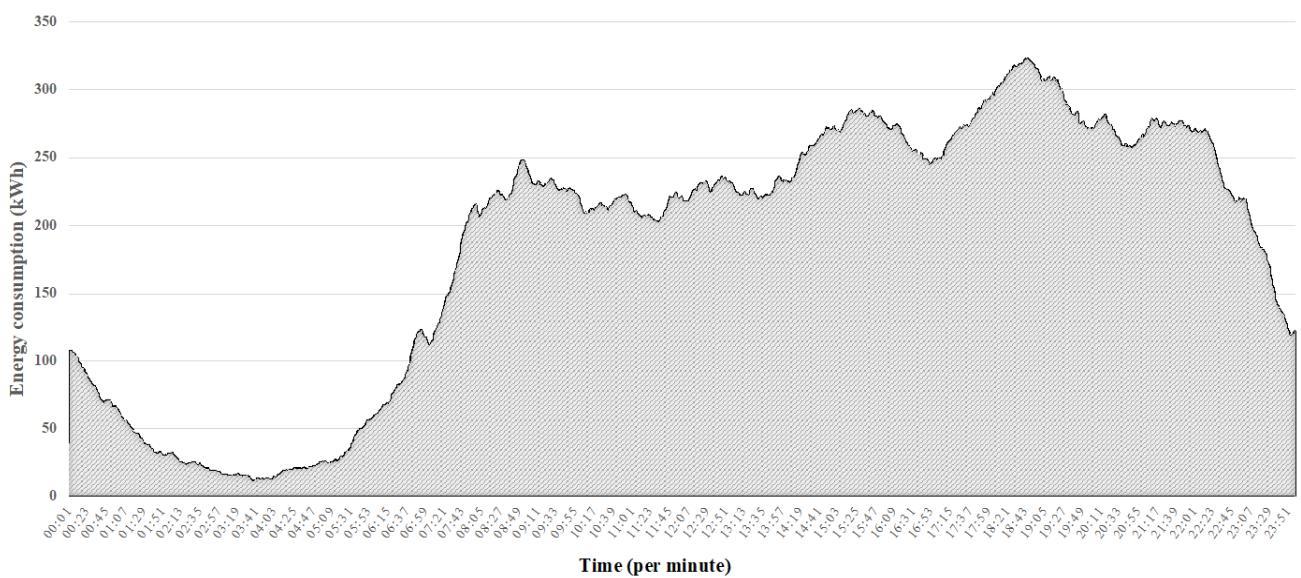**Figure 11.** Energy consumption in autumn, 2018.

Figure 12. Energy consumption of a random day (March 2018) for each minute.

5. Conclusions

In this study, a novel estimation model is proposed and applied to the real-world data of the New York City taxi fleets in order to obtain the total energy consumption of EVs for different time intervals, in the scenario that all the yellow taxis are replaced with a specific type of electric vehicle. While other studies have focused on energy consumption prediction of an individual EV, the outputs and results of this study provide a promising approach to estimating the total energy demands of EVs. One of the main outcomes of this article is the identification of demand patterns at any time intervals (day, month, or year) with any desired level of accuracy, to adjust the power generation's schedules and prevent any disturbances due to the consequences of charging a massive amount of EVs simultaneously. The output shows that energy consumptions are highly correlated to the weekdays; for instance, Thursdays and Fridays are usually identified as the high peak, while weekend days and holidays present the lowest energy usage. The next step of this work is to investigate more deeply the importance of other influential factors.

References

1. International Energy Agency and International Union of Railways. *Railway Handbook 2015—Energy Consumption and CO2 Emissions*; International Energy Agency and International Union of Railways: Paris, France, 2015.
2. Kanchev, H.; Colas, F.; Lazarov, V.; Francois, B. Emission reduction and economical optimization of an urban microgrid operation including dispatched PV-based active generators. *IEEE Trans. Sustain. Energy* **2014**, *5*, 1397–1405. [CrossRef]
3. Mehta, D.; Sapun, P.; Hamke, A.K. *In-Depth: Emobility 2018. Statista Mobility Market Outlook*; Trend Rep: Luxembourg, 2018.
4. Su, W.; Eichi, H.; Zeng, W.; Chow, M.-Y. A Survey on the Electrification of Transportation in a Smart Grid Environment. *IEEE Trans. Ind. Inform.* **2011**, *8*, 1–10. [CrossRef]
5. Bloomberg, L.P. *The Relentless March toward an EV Future is Good News for Miners*; Bloomberg: New York, NY, USA, 2020.
6. IEA. *Global EV Outlook 2019*; Technical Report; IEA: Paris, France, 2019. Available online: https://www.iea.org/reports/global-ev-outlook-2019 (accessed on 24 August 2020).
7. Fernandez, L.P.; Roman, T.G.S.; Cossent, R.; Domingo, C.M.; Frias, P. Assessment of the impact of plug-in electric vehicles on distribution networks. *IEEE Trans. Power Syst.* **2011**, *26*, 206–213. [CrossRef]
8. Mu, Y.; Wu, J.; Jenkins, N.; Jia, H.; Wang, C. A Spatial–Temporal model for grid impact analysis of plug-in electric vehicles. *Appl. Energy* **2014**, *114*, 456–465. [CrossRef]
9. Qian, K.; Zhou, C.; Allan, M.; Yuan, Y. Modeling of Load Demand Due to EV Battery Charging in Distribution Systems. *IEEE Trans. Power Syst.* **2010**, *26*, 802–810. [CrossRef]
10. Neaimeh, M.; Wardle, R.; Jenkins, A.M.; Yi, J.; Hill, G.; Lyons, P.F.; Hübner, Y.; Blythe, P.T.; Taylor, P.C. A probabilistic approach to combining smart meter and electric vehicle charging data to investigate distribution network impacts. *Appl. Energy* **2015**, *157*, 688–698. [CrossRef]
11. Jeon, S.; Hong, B. Monte Carlo simulation-based traffic speed forecasting using historical big data. *Future Gener. Comput. Syst.* **2016**, *65*, 182–195. [CrossRef]
12. Arias, M.B.; Bae, S. Electric vehicle charging demand forecasting model based on big data technologies. *Appl. Energy* **2016**, *183*, 327–339. [CrossRef]
13. Zhang, P.; Wu, X.; Wang, X.; Bi, S. Short-term load forecasting based on big data technologies. *CSEE J. Power Energy Syst.* **2015**, *1*, 59–67. [CrossRef]
14. Pantoš, M. Stochastic optimal charging of electric-drive vehicles with renewable energy. *Energy* **2011**, *36*, 6567–6576. [CrossRef]
15. Shaaban, M.F.; Atwa, Y.M.; El-Saadany, E.F. PEVs modeling and impacts mitigation in distribution networks. *IEEE Trans. Power Syst.* **2013**, *28*, 1122–1131. [CrossRef]
16. Mu, T.; Jiang, J.; Wang, Y. Heterogeneous delay embedding for travel time and energy cost prediction via regression analysis. *IEEE Trans. Intell. Transp. Syst.* **2013**, *14*, 214–224. [CrossRef]

17. Silva, M.A.; de Melo, H.N.; Trovao, J.P.; Pereirinha, P.G.; Jorge, H.M. An integrated fuzzy logic energy management for a dual-source electric vehicle. In Proceedings of the IECON 2013—39th Annual Conference of the IEEE Industrial Electronics Society, Vienna, Austria, 10–13 November 2013; pp. 4564–4569.

18. Galus, M.D.; Zima, M.; Andersson, G. On integration of plug-in hybrid electric vehicles into existing power system structures. *Energy Policy* **2010**, *38*, 6736–6745. [CrossRef]

19. Sortomme, E.; El-Sharkawi, M.A. Optimal charging strategies for unidirectional vehicle-to-grid. *IEEE Trans. Smart Grid* **2010**, *2*, 131–138. [CrossRef]

20. Olivella-Rosell, P.; Villafafila-Robles, R.; Sumper, A.; Bergas-Jané, J. Probabilistic agent-based model of electric vehicle charging demand to analyse the impact on distribution Networks. *Energies* **2015**, *8*, 4160–4187. [CrossRef]

21. Majidpour, M.; Qiu, C.; Chu, P.; Pota, H.R.; Gadh, R. Forecasting the EV charging load based on customer profile or station measurement? *Appl. Energy* **2016**, *163*, 134–141. [CrossRef]

22. Fiori, C.; Ahn, K.; Rakha, H.A. Power-based electric vehicle energy consumption model: Model development and validation. *Appl. Energy* **2016**, *168*, 257–268. [CrossRef]

23. Moon, H.; Park, S.Y.; Jeong, C.; Lee, J. Forecasting electricity demand of electric vehicles by analyzing consumers' charging patterns. *Transp. Res. Part D Transp. Environ.* **2018**, *62*, 64–79. [CrossRef]

24. Gomez-Quiles, C.; Asencio-Cortes, G.; Gastalver-Rubio, A.; Martinez-Alvarez, F.; Troncoso, A.; Manresa, J.; Riquelme, J.C.; Riquelme-Santos, J.M. A Novel Ensemble Method for Electric Vehicle Power Consumption Forecasting: Application to the Spanish System. *IEEE Access* **2019**, *7*, 120840–120856. [CrossRef]

25. Morlock, F.; Rolle, B.; Bauer, M.; Sawodny, O. Forecasts of Electric Vehicle Energy Consumption Based on Characteristic Speed Profiles and Real-Time Traffic Data. *IEEE Trans. Veh. Technol.* **2020**, *69*, 1404–1418. [CrossRef]

26. Cai, H.; Jia, X.; Chiu, A.S.; Hu, X.; Xu, M. Siting public electric vehicle charging stations in Beijing using big-data informed travel patterns of the taxi fleet. *Transp. Res. Part D Transp. Environ.* **2014**, *33*, 39–46. [CrossRef]

27. Li, B.; Kisacikoglu, M.C.; Liu, C.; Singh, N.; Erol-Kantarci, M. Big Data Analytics for Electric Vehicle Integration in Green Smart Cities. *IEEE Commun. Mag.* **2017**, *55*, 19–25. [CrossRef]

28. Uimonen, S.; Lehtonen, M. Simulation of Electric Vehicle Charging Stations Load Profiles in Office Buildings Based on Occupancy Data. *Energies* **2020**, *13*, 5700. [CrossRef]

29. Berthold, M.; Cebron, N.; Dill, F.; Gabriel, T.; Kötter, T.; Meinl, T.; Ohl, P.; Thiel, K.; Wiswedel, B. KNIME–The Konstanz information miner: Version 2.0 and Beyond. *SIGKDD Explor.* **2009**, *11*, 26–31. [CrossRef]

30. Van Rossum, G.; Drake, F.L. *Python 3 Reference Manual*; CreateSpace: Scotts Valley, CA, USA, 2009.

31. Pattern-Based Strategy: Getting Value from Big Data, Gartner Group Press Re-Lease. Available online: http://www.gartner.com/it/page.jsp?id=1731916 (accessed on 14 April 2020).

32. Gartner IT Glossary. Available online: http://www.gartner.com/it-glossary/big-data/ (accessed on 14 April 2020).

33. TechAmerica Foundation's Federal Big Data Commission. Demystifying Big Data: A Practical Guide to Transforming the Business of Government. 2012. Available online: http://www.techamerica.org/Docs/fileManager.cfm?f=techamerica-bigdatareport-final.pdf (accessed on 14 April 2020).

34. The 4 V's of Big Data. Available online: http://www.ibmbigdatahub.com/tag/587 (accessed on 14 March 2020).

35. National Oceanic and Atmospheric Administration. Available online: https://www.noaa.gov (accessed on 9 March 2020).

36. The Official Website of the City of New York. Available online: https://www1.nyc.gov (accessed on 9 March 2020).

37. Laws of New York. Available online: http://Public.leginfo.state.ny.us (accessed on 15 March 2020).

38. McKinney, W. AQR Capital Management, Pandas: A Python Data Analysis Library. Available online: http://pandas.sourceforge.net (accessed on 15 March 2020).

39. Wu, X.; Freese, D.; Cabrera, A.; Kitch, W.A. Electric vehicles' energy consumption measurement and estimation. *Transp. Res. Part D Transp. Environ.* **2015**, *34*, 52–67. [CrossRef]

40. EV Charging 101—How Much Electricity Does an Electric Car Use? Available online: https://www.virta.global/blog/ev-charging-101-how-much-electricity-does-an-electric-car-use (accessed on 13 December 2020).