

Article

Data Driven Approaches for Sustainable Development of E-Mobility in Urban Areas

Marialisa Nigro ^{1,*}, Marina Ferrara ¹, Rosita De Vincentis ¹, Carlo Liberto ² and Gaetano Valenti ²

¹ Department of Engineering, “Roma Tre” University, 00146 Rome, Italy; marina.ferrara@uniroma3.it (M.F.); ros.devincenit@stud.uniroma3.it (R.D.V.)

² ENEA, Italian National Agency for New Technologies, Energy and Sustainable Economic Development, Laboratory of Systems and Technologies for Sustainable Mobility, 00196 Rome, Italy; carlo.liberto@enea.it (C.L.); gaetano.valenti@enea.it (G.V.)

* Correspondence: marialisa.nigro@uniroma3.it

Abstract: This study focuses on a modeling framework to support mobility planners and energy providers in the sustainable development of electric mobility in urban areas. Specifically, models are provided to simulate measures for the optimal management of energy demand and thoughtful planning of charging infrastructures in order to avoid congestion on the power grid. The measures, and consequently the models, are classified according to short-term initiatives based on multimodality between electric vehicles and public transport (Park and Ride), as well as medium to long-term initiatives based on the development of an energy-oriented land use of the city. All the models are data-driven, and different sets of floating car data available for the city of Rome (Italy) have been exploited for this aim. The models are currently being implemented in an agent-based simulator for electric urban mobility adopted by the National Agency for Energy and Environment in Italy (ENEA).



Citation: Nigro, M.; Ferrara, M.; De Vincentis, R.; Liberto, C.; Valenti, G. Data Driven Approaches for Sustainable Development of E-Mobility in Urban Areas. *Energies* **2021**, *14*, 3949. <https://doi.org/10.3390/en14133949>

Academic Editors: Kazimierz Lejda, Artur Jaworski and Maksymilian Mądziel

Received: 12 May 2021
Accepted: 28 June 2021
Published: 1 July 2021

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Reducing road vehicle emissions in urban areas represents a big challenge for Europe [1]. Electric mobility (e-mobility) applied to both private and public transport is one of the most promising solutions to reduce air pollution in densely populated areas as well as to support energy diversification and cut greenhouse gas emissions [2].

Electric vehicles (EVs) have expanded significantly over the last decades, largely driven by the mass expansion of this transport mode in China accounting for nearly half of all vehicle sales [3]. Despite its rapid global growth, the European EV market has shown moderate progress, often connected to support schemes. Most EVs are concentrated in a few northern and western EU Member States, although southern and eastern ones have recently recorded the highest sales growth. However, according to forecasts from Deloitte for 2030 [4], EV sales will achieve a domestic market share of around 42 percent in Europe. Moreover, the IEA’s Stated Policies Scenario [3] forecasts that the share of demand due to electric vehicles in total electricity consumption in Europe by 2030 will grow to as high as 4%.

The wide-scale adoption of EVs in urban areas poses new challenges with respect to charging infrastructure, covering a wide variety of transport patterns and needs. Due to the impact on the electricity grid, infrastructure improvements need to be made to optimize EV charging operations and to ensure a smooth integration with renewables-based energy generation systems. The investments and infrastructure requirements can vary significantly from one urban context to another according to specific market characteristics, such as local infrastructure, land use, energy system, mobility culture, and patterns.

This paper proposes a modeling framework that can be useful to evaluate short-term and medium to long-term measures to avoid energy congestion in urban areas; specifically, it aims to simulate multimodality between EVs and public transport (Park and Ride), as well as to estimate energy demand requested by EVs as a function of the land use of the city these are moving to.

The correct simulation of a short-term measure, such as promoting multimodality between EVs and public transport in Park and Ride sites located in suburban areas of the city, envisages a twofold goal from the energy viewpoint. Firstly, energy demand can be shifted differently in space and time, with respect to the request derived by the usual commuter trips adopting only private e-vehicles; moreover, it would promote an efficient charging management through a vehicle-to-grid approach in which the EVs assume the role of “demand peak stabilizers” in the parking sites.

Indeed, a correct estimation of energy demand as a function of the land use can help in the medium- to long-term in moving urban planning toward a more energy-saving approach.

Thus, the entire modeling framework, once implemented into a simulator for EVs, can be adopted by local governments, in cooperation with mobility planners and energy providers, to promote a sustainable development of e-mobility.

In both cases, specific models adopting-data-driven approaches have been realized and tested. Different sets of floating car data (FCD) derived by probe vehicles moving on the road network of Rome (Italy) have been exploited for this aim; hence, the city of Rome is the reference case for the study.

The paper is structured as follows: Section 2 presents the literature supporting the modeling framework and the research contribution. The methodology behind the developed models is described in Section 3: here, the main dataset, i.e., the FCD dataset, is also described to obtain a useful insight into Rome’s private vehicle travel patterns. Section 4 shows the resulting models and their calibration and validation for the Rome case study. Possible policy implications about multimodality through Park and Ride for EVs and energy-oriented land use are reported in Section 5. Conclusions and further developments follow.

2. Literature Outcomes and Research Contribution

The literature supporting the development of this study has been divided according to the following topics: (1) simulation tools for EVs; (2) planning measures for the sustainable development of EVs in urban areas; (3) multimodal models for estimating Park and Ride demand; and (4) how to realize urban development oriented to e-mobility.

In [5], a detailed review of simulation tools for EVs is reported. The main elements of EV simulators are the simulation of trips, the modeling and analysis of charging stations, the charging behavior, and the impact of EV user decisions on the electrical network distribution. EV simulation can be implemented by adding specific modules inside flexible traffic simulators such as SUMO [6] or MatSim [7] or, alternatively, by developing ad-hoc instruments. Most of the simulation performed by the aforementioned tools deals with EV consumption models, charging behavior, charging design/management, and vehicle to grid strategies [8–10].

Indeed, charging management and vehicle-to-grid are also adopted to reshape the charging load profile on the power grid, thus reducing line congestion or voltage deviations. Alternative strategies to reduce the energy load can be more “transport oriented”, since EVs and their supporting charging infrastructure act as a glue between the transportation and the electrical system [11].

The concentration of energy demand can be verified in terms of both time and space, and it is well known from the literature that a specific activity distribution or travel lifestyle can impact charging patterns and, as a consequence, energy congestion [12,13].

In this paper, the models developed aim to be implemented into an ad-hoc simulator for EVs. As with similar tools (i.e., MatSim), the simulator relies on an agent-based system,

which is characterized by two main levels: (1) a synthetic trajectory generator and (2) a recharge behavior model [14].

In contrast, regarding the models here proposed, their aim is to simulate actions for smoothing the charging loads and specifically:

1. Reducing traffic congestion and related concentration of charging behavior through the modal shift toward public transport by means of Park and Ride sites. It is assumed that by promoting the Park and Ride choice, EVs can also contribute to vehicle-to-grid strategies in the parking sites;
2. Supporting urban planners, providing instruments able to evaluate the energy demand as a function of the travel pattern, and, consequently, as a function of the land use of the city.

As for simulating actions to promote/push Park and Ride with EVs, this is usually done by calibrating modal shift models. Most Park and Ride models are usually based on random utility theory such as logit models [15]. Through logit models, the main variables affecting the Park and Ride demand can be investigated. For example, authors in [16] found that the lack of parking spaces and the transfer time are the main variables in defining the level of Park and Ride demand; in [17], the importance of car accessibility to Park and Ride sites and the waiting time of public transport is underlined. Only recently have other techniques based on machine learning been investigated also for modal shift models: Tai-Yu Ma [18] analyzed multimodal choice behaviors through a Bayesian network, comparing it with a multinomial logit model. Thanh et al. [19] adopted decision trees for the binary choice “shifting or not to other travel options”; a similar approach has been proposed by [20] but using a support vector machine. Regarding the adoption of machine learning in combination with open data, tracking data, or real data not derived by survey, few examples can be found in the literature, but these are mainly related to estimating the level of occupancy of a parking site rather than to understand the potential for a modal shift [21,22]. Thus, there is still work to exploit these combinations of methods and data for modeling a modal shift.

Interaction between urban development and EVs is a significant topic in the literature. Indeed, EV deployment would play an important role in the spatial organization and structure of cities, and the land use will influence the charging request. Most of the contributions about land use and EVs deal with the optimal location of charging infrastructures as a function of the different functional areas of the cities or evaluating which facilities or points of interest are more interesting for EV charging [23]. For example, Zhi Wu et al. [24] work on locating multi-functional charging stations (MFCS) in a medium-sized city with different functional areas. However, the difference between the land use is only adopted to classify the vehicle travel destinations (assuming, for example, commuting trips mainly between residential areas and industrial or commercial areas and social and recreational trips mainly between commercial and residential areas). Other contributions deal with the effects of EVs in terms of urban sprawl: indeed, it is assumed that EV owners would drive more than they used to, even if they are shifting from very efficient ICEVs [25].

Regarding the influence of the land use, mobility patterns, and related energy consumption, the only recently integrated land use–transport–energy model, able to incorporate the energy system explicitly, has started to be developed. Thus, there is evidence that new models to support the right association between these three systems are of strategic importance. Indeed, the spatial heterogeneity of urban areas calls for extreme care when evaluating the effectiveness of sustainable energy-oriented policies [26].

This study will try to take a step forward with respect to the analyzed literature. The models that will be developed, specifically a Park and Ride model and an energy-oriented land-use model able to describe the transport energy profiles, will be (i) data-driven and (ii) easily implementable in any agent-based simulator.

3. Methodology

Since the models have the common factor of being data-driven, the methodology section starts describing the dataset adopted for calibration and validation. Then, modeling details are given about both multimodality and energy-oriented land use.

3.1. Available Dataset

In this study, the main data sources are timestamped geospatial locations provided by probe vehicles, i.e., floating car data (FCD). Additional data on mobility services and transport infrastructure are also collected, namely:

- Park and Ride facilities for public transport inside the Rome metropolitan area;
- GTFS data published by the Mobility Agency of the city of Rome that contains schedules, fares, and geographic transit information;
- Census data about population and activities.

FCD are currently available worldwide; most of the vehicles (both cars and trucks, as well as fleets of vehicles such as a public transport fleet) are provided with an on-board unit, and tracking data are commercialized by several solution providers, thus ensuring accurate, reliable data for transport applications [27]. The available FCD samples refer to private cars moving in the metropolitan area of Rome during two different time horizons: about 150,000 cars (160 million records, penetration rate 5%) tracked during May 2013 and about 240,000 vehicles (317 million records, penetration rate 7%) tracked during November 2015. The locations of points during the FCD trip are recorded together with the start and stop points. The recording frequency of the adopted data is set as follows [2]: (1) every 30 s if the probe vehicle moves along the motorway network or some main roads in the metropolitan area; (2) every two kilometers elsewhere. The information available in the FCD is the vehicle ID, the geo-referenced position of the vehicle, the time and date, the instant speed, the state of the engine (on or off), the distance traveled from the previously recorded position, and GPS signal quality. Thus, it is possible to extract useful insight from FCD collections such as distances traveled, average speeds and related travel times, the final destination or the intermediate destinations of the trip, the parking time, etc.

The FCD have been used to extract vehicle home locations that can be taken as input to explain or predict multimodality and energy-oriented land use. In the first case, it is assumed that most of the trips connected to Park and Ride are home-based trips, since Park and Ride is usually adopted for commuting trips. In the second case, defining whether the starting/ending zone of a trip is the home location can help in understanding the motivation behind the trip, thus better defining the land use.

The residence location procedure (RLP) is conducted on a zoning system that represents a compromise between the census zones (13,656 zones) and urban zones (155 zones) of Rome, consisting of 1409 zones with an average area of 86 ha and an average population per zone of 1900 inhabitants.

The procedure works as follows:

- trip destination locations of each FCD vehicle are extracted and mapped;
- all the destinations are clusterized, adopting a DBScan clustering technique with parameter ϵ (i.e., the proximity radius between points in the same cluster) equal to 120 meters and the minimum number of points for each cluster equal to one;
- overnight stops of each FCD vehicle based on arrival time and departure time at/from destination are extracted.
- the residence is the cluster of destinations with higher values of overnight stops.

The procedure has been validated by comparing the resulting number of residences in each zone with the census population, thus showing a high correspondence (R^2 equal to 0.7465, Figure 1).

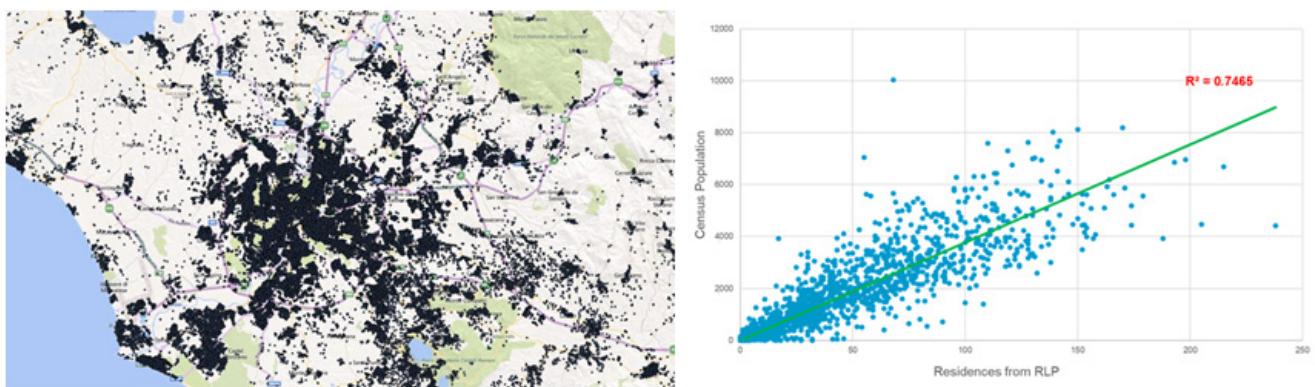


Figure 1. Map of the residences after the application of the residence location procedure and comparison with census data.

3.2. Multimodality

Starting from the multimodality measure, we focus on the development of a suite of models able to represent multimodal choice behaviors between EVs and public transport in the parking sites located close to the metro stations in the suburbs of the city, i.e., Park and Ride (P&R) models.

All the models are data-driven, but two different approaches have been followed (Figure 2): (1) an aggregated approach where both dependent variable and explanatory variables are aggregated variables at zonal level; (2) a disaggregated approach where the decision of each user is simulated as a function of several explanatory variables related to both individual features and zonal ones.

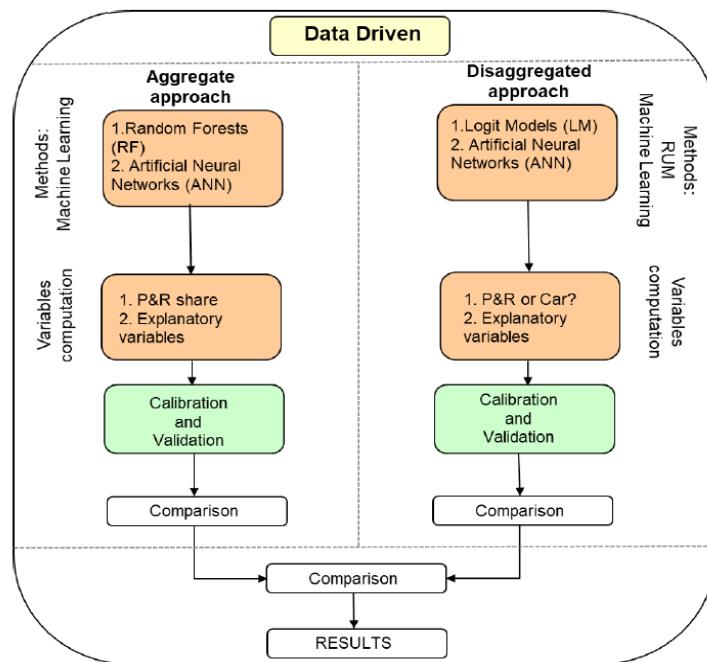


Figure 2. Methodological framework of multimodality models.

3.2.1. Formulation of the Park and Ride Models

Starting from the aggregated approach, the model has to be able to represent multimodal choice behaviors, thus assessing the share of the private EVs that can be moved onto public transport by using Park and Ride sites. The model assumes the following structure:

$$(G_O^{\text{Park}} / G_O) |_{D_t} = f(x_1, x_2, \dots, x_n) \quad (1)$$

where G_O^{Park} / G_O is the share of generated trips from the traffic zone O that would make P&R in the time interval Dt. The explanatory variables x_1, x_2, \dots, x_n are mainly related to: (1) the accessibility of the Park and Ride sites, (2) impedances due to the travel times on the public transport, (3) the location and occupancy of the P&R sites, and (4) characteristics related to the activity system in the starting point of the trip. The proposed explanatory variables were firstly proposed in [28] and then updated according to [14]. Specifically, the following ten explanatory variables have been adopted:

1. Variables depending on the starting point of the trip (origin/start traffic zone O) and the starting time interval (Dt) of the trip:
 - Impedance on private transport x_1 (2), i.e., the average access time from the traffic zone O to the P&R sites belonging to the catchment area of O, weighted for the O-P&R flows;
 - Impedances on public transport: x_2, x_3 : x_2 (3) is an average travel time from the P&R sites belonging to the catchment area of O to all the destinations, weighted for the O-P&R flows and for the attractiveness of each destination; In x_3 (4) the travel times on public transport are weighted for the origin-destination (OD) flow;
 - Travel time differences between P&R and car-only mode x_4 (5); x_4 represents the travel time advantage the user may have following the adoption of the P&R mode. The advantage is positive if the total travel time, the sum of the parking access plus the travel time by public transport, is less than the time required to go from O to D by car;
 - The average occupancy of the P&R sites belonging to the catchment area of the origin/start traffic zone O, x_5 (6).
2. Variables depending only on the characteristics of the starting point of the trip origin/start traffic zone O) and not by the time interval, specifically:
 - The variable x_6 that represents the geographical extent of the origin/start traffic zone O [km^2];
 - The variable x_7 that represents the population density of O (between 18 and 70 years old as possible users of the P&R service);
 - The variable x_8 , i.e., a binary variable equal to 1 if O is within the Main Ring Road of Rome (where the metro lines are located), 0 otherwise;
 - The variable x_9 that represents the number of urban railway stations and metro stations in the starting zone O.
3. Variables depending only on the time interval, specifically:
 - The dummy variable x_{10} that is equal to 1 if the starting time interval of the trip is inside the morning peak (between 6:00 a.m. and 11:00 a.m. for homework trips by public transport), 0 otherwise.

$$x_1 = \left(\sum_{P \in O} t_{OP} f_{OP} \right) / \left(\sum_{P \in O} f_{OP} \right) = \left(\sum_{P \in O} t_{OP} f_{OP} \right) / G_O^{\text{Park}} \quad (2)$$

$$x_2 = \left(\sum_{P \in O} f_{OP} \left(\frac{\sum_D t_{PD}^{TP} A^D}{\sum_D A^D} \right) \right) / \left(\sum_{P \in O} f_{OP} \right) = \left(\sum_{P \in O} f_{OP} \left(\frac{\sum_D t_{PD}^{TP} A^D}{\sum_D A^D} \right) \right) / G_O^{\text{Park}} \quad (3)$$

$$x_3 = \left(\sum_D f_{OD} A^D \left((\sum_{P \in O} t_{PD}) / n_{P \in O} \right) \right) / \left((\sum_D f_{OD}) (\sum_D A^D) \right) \quad (4)$$

$$x_4 = \sum_D A^D (t_{OD} - \left(\sum_{P \in O} (t_{OP} + t_{PD}) / n_{P \in O} \right)) / \sum_D A^D \quad (5)$$

$$x_5 = \left(\sum_{P \in O} r_P \right) / n_{P \in O} \quad (6)$$

where:

t_{OP} = travel time by car from the origin/start traffic zone O to the P&R site P;

f_{OP} = vehicle flow from the origin/start traffic zone O to the P&R site P;

G_O^{Park} = generated trips from the origin/starting traffic zone O making P&R;

$t_{P \rightarrow *}^{TP}$ = average travel time on public transport to reach all the destinations of the city from the P&R site P;

f_{OD} = vehicle flow from the origin/starting traffic zone O to the final destination D;

t_{PD} = travel time by public transport from the P&R site P to the final destination D;

$n_{P \in O}$ = number of P&R sites belonging to the catchment area of the origin/starting traffic zone O;

G_O = generated trips from the origin/starting traffic zone O

t_{OD} = travel time by private transport from the origin/starting traffic zone O to the final destination D;

A_D = attractiveness of the destination zone D;

r_P = occupancy of the P&R site P.

In the case of the disaggregated approach, the output of the model is the choice for each user between Park and Ride or car, while the explanatory variables contain both aggregated and disaggregated information. Regarding the aggregated ones, we refer to the zonal variables described for the aggregated approach since these are a function of the starting zone of the trip. The only variable not adopted from the aggregated set is the dummy variable x_{10} , since the disaggregated model has been calibrated considering only those trips related to the morning peak hours. Regarding the disaggregated ones, six new variables have been added to describe the behavior of each individual user, specifically:

- the traveled distance by Park and Ride x_{11} ;
- the total P&R travel time x_{12} ;
- the number of transfers by public transport x_{13} ;
- the average waiting time by public transport x_{14} ;
- the traveled distance by car x_{15} ;
- the travel time by car x_{16} .

Traveled distances and travel times have been considered in the models as the difference between Park and Ride and car options.

Since each variable needs to be calculated regarding whether the choice of each user is car or Park and Ride, some of them are not directly available by FCD (e.g., for each FCD trip stopping in a Park and Ride site, the final destination is unknown); thus, some assumptions are required as follows:

- Only the main destinations accounting for the 70% of the demand for each starting zone are considered if the destination is unknown;
- The features related to destinations are weighted for the attractiveness AD ;
- For users using a car, the potential Park and Ride site is considered the one closest in time with respect to the starting point of the trip.

3.2.2. Calibration and Validation Methods

For the aggregated approach, two machine learning (ML) techniques have been adopted for calibrating f in (1): random forest, RF, and artificial neural networks, ANN [29]. The motivation behind the adoption of ML lies in the lack of an explicit functional form for f .

RF is a supervised learning technique adopted for both classification and regression and was firstly proposed by Breiman [30]. RF works by creating an ensemble of decision trees and randomly selecting a subset of features to grow each tree. Then, it provides the class mode (if classification is required) or the average prediction (if regression is required) of individual trees, finally generating the resulting model through a combination of several parallel decision trees. For the training and the test phase of each tree, the data are usually divided into two subsets. In transport engineering, advanced ML techniques such as RF and ANN are emerging in recent years due to the increasing availability of traffic data.

Both RF and ANN are mainly used as data-driven approaches for travel time prediction [31–33]. Contributions in literature dealing with the adoption of RF for demand forecasting are harder to find. For example, in Ghasri et al. [34], decision tree and RF

are adopted to reproduce individual mobility patterns. Ashqar et al. [35] used RF to predict the number of available bikes in bike-sharing stations. Liu et al. [36] compared RF with well-known regression methods such as autoregressive integrated moving average (ARIMA) and support vector regression (SVR) for the demand estimation of four different on-demand shared transport services.

Regarding ANN, one of the first contributions dealing with transport demand modeling was by Hensher et al. [37]: they compared ANN with logit models underlining a high potential of ANN to describe individual mode choice. The same has been reported in Cantarella et al. [38]. Nijkamp et al. [39] in 2004 adopted ANN for modal choice for freight. Gazder et al. [40] proposed a logit-ANN ensemble for mode choice modeling based on survey data. In contrast, recently, authors in [41] calibrated a convolutional neural network (CNN) for the mode choice reconstructing trajectories and related additional features through GPS, obtaining good results in terms of model accuracy.

In our study, the RF was built by varying the main model parameters: the number of trees and the maximum depth. Regarding ANN, a multi-layer feedforward neural network has been trained by different algorithms (Levenberg–Marquardt, Bayesian regularization), modifying the activation function between layers (rectified linear unit—ReLU—or logistic regression—Sigmoid), the number of hidden layers, and the number of neurons in each hidden layer. In both cases, the validation adopts out-of-sample data, unseen by the model during the training phase.

For the disaggregated approach, ANNs have been also considered. However, since the output of the approach is mainly a binary output (car or Park and Ride), a classification with respect to the most complicated regression is required in (1); both a Perceptron network as well as a multi-layer feedforward neural network have been trained. The Perceptron is basically simpler than the multi-layer because it does not adopt hidden layers. As done for the aggregated approach, the ANNs have been trained with different algorithms, activation functions, and numbers of neurons in the hidden layer (the last one only for the multi-layer). Due to the high number of instances in the disaggregated approach (one instance for each FCD trip) the solution algorithms are different with respect to the aggregated approach, i.e., stochastic gradient descent (SGD) and Adam.

Apart from ANNs, a classical behavioral approach based on random utility theory has also been applied for the disaggregated approach. A simple binomial logit has been considered due to the need to represent the binary choice between car or Park and Ride. The maximum likelihood method is adopted for calibration, while validation is conducted through statistical tests on the estimated coefficients [42].

3.3. Energy-Oriented Land-Use Model

A methodology has been developed able to link a specific land use to its potential energy demand.

Firstly, the land use for each zone of the city is assessed through the adoption of mobility patterns from FCD. Similar to Caceres and Benitez [43], we exploited a clustering approach (K-means clustering, [44]). Specifically, FCD trips have been divided between weekend and weekdays; then, since the data contain information about start and stop points of the vehicle's engine together with the corresponding timestamp [45], these are converted into generated and attracted trips by each zone, and profiles during the daytime have been derived. The K-means algorithm works in order to clusterize profiles, finding the optimal number of clusters k^* able to minimize the distance between observed profiles in each cluster. Each $k \in k^*$ corresponds to a land use and, after the clustering phase, it is possible to find different values of k^* for attracted/generated trips during weekends/weekdays. To finally assign a land use to each zone, the ones having the same land use with respect to both generated and attracted trips are assigned univocally, while the other ones are considered mixed land use with a prevailing function given by the higher impact of a land use with respect to the other ones.

The resulting land use of each zone has been validated, comparing it with aggregated available statistics such as population density or residential buildings with respect to the total buildings. Relating the land use to mobility patterns allows the improvement of traffic prediction models and redefining zoning regulations, something that simply cannot be done relating the land use to the usual existing statistics.

Finally, the energy demand can be computed. Specifically, for each zone z with land use k (z_k), for both home-based and not home-based trips, i and time interval h , the energy demand, are computed as follows:

$$E_h^{z_k} [kWh] = f \times \sum_{i=HB, NHB} Trips_{i,h}^{z_k} * ParkingTime_{i,h}^{z_k} * Power\ col_i |_{HB, NHB} \forall z_k, \forall h \quad (7)$$

where:

f = recharge frequency computed as a function of distance traveled and battery life before charging;

$Trips_{i,h}$ = number of trips of type i arriving in k during time h ;

$ParkingTime_{i,h}$ = parking time of trips of type i arriving in k during time h , where distribution of parking times can be exploited by FCD;

$Power\ col_i$ = Recharge power. It can be assumed equal to 3.7 kW for home-based trips and 7.4 kW for not-home-based trips (thus assuring, respectively, an average recharge time of about 8 and 4 h).

The methodology has been applied to several zones of the city of Rome characterized by different land use, showing the value of the energy demand and its distribution in time and giving fundamental information for energy providers and urban planners. Indeed, the methodology requires calibration with respect to the current data of mobility patterns, but it allows the simulation of how the demand for energy can change as a function of the land use (something that can be realized only in the medium to long term by urban planners).

4. Results

4.1. Multimodality: Results

Starting from the aggregated approach, the database counts about 1000 records and 10 independent variables. It has been divided according to a cross-validation method into two subsets: 70% used to train the model (training data set) and the remaining 30% (validation data set) to evaluate its predictive capability.

According to the computed performance indices, the comparison between the two different machine-learning techniques suggests that in the P&R demand model RF gives better results than ANN. The RF model succeeds in generalizing the relationships in the data with higher prediction accuracy (Table 1): specifically, RF shows a mean squared error between the model and the data of 0.560 with a correlation index (R^2) of about 0.73; in contrast, ANN shows a mean squared error of 1.085 with a correlation index (R^2) of about 0.66.

Table 1. Best resulting models for the different techniques in the aggregated approach.

ML Technique	Best Architecture	R2 Train	R2 Test	MSE (All)
Random Forest	1000 number of trees 20 max depth	0.891	0.734	0.560
Neural Network	2 hidden layers 8 neurons	0.704	0.657	1.085

Regarding the analyses conducted on the parameters of the machine learning techniques:

- for RF, the analyses showed that the variability of the results decreases as the number of trees increases (Figure 3a);
- a clear difference exists between the RF with 10 depth levels and the others (Figure 3b)

- the optimal parameter values for the RF have been set equal to 1000 as number of trees and 20 as the maximum depth (Figure 3);
- about the different algorithms adopted to train the ANN, the Levenberg–Marquardt algorithm resulted to be the most performing one. Indeed, despite the mean squared error seems to decrease more with the Bayesian regularization (Figure 4a,b), the error with respect to the only test set demonstrates the higher accuracy of the Levenberg–Marquardt (Figure 4c);
- the optimal parameter values for the ANN have been set as 2 hidden layers and 8 neurons for each hidden layer (Figure 4c).

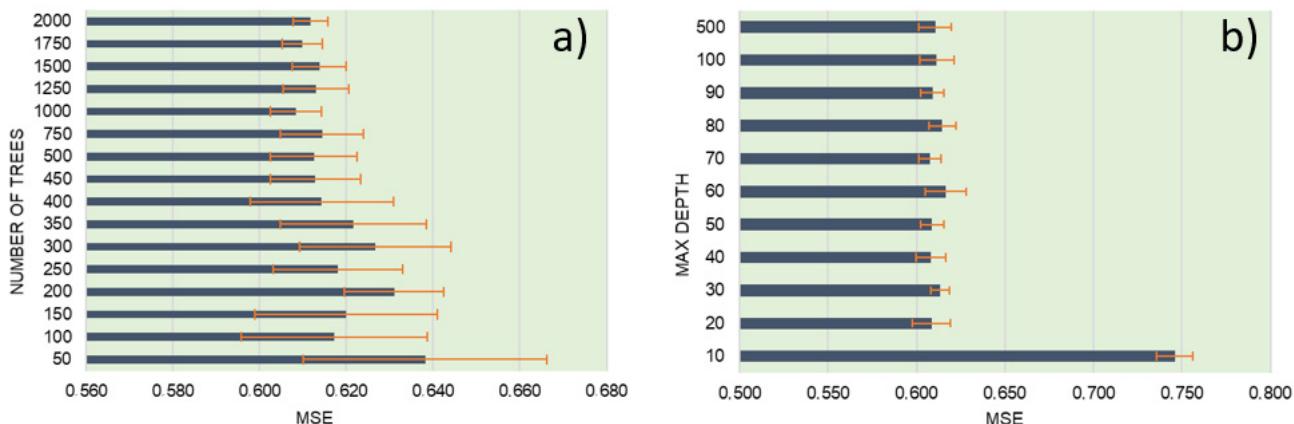


Figure 3. Mean square error (MSE) as a function of number of trees (a) and maximum depth (b) [28].

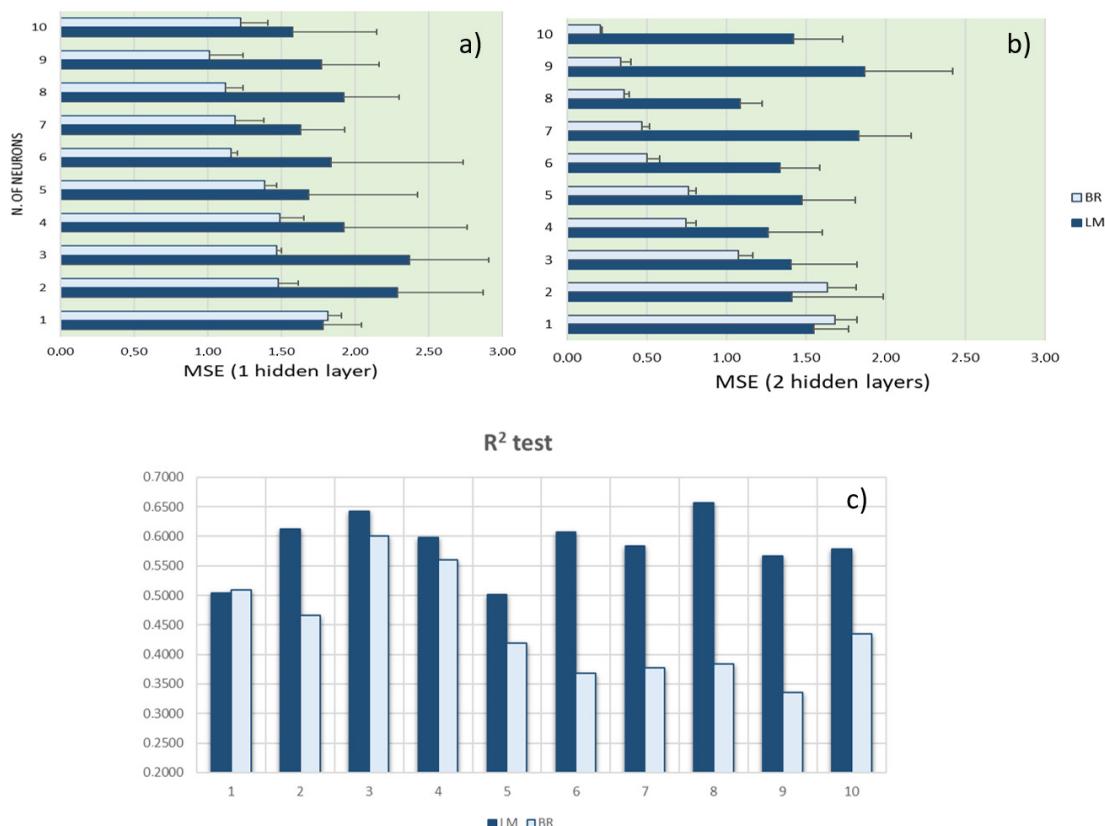


Figure 4. Mean square error (MSE) as a function of number of neurons for one hidden layer (a) and two hidden layers (b). R² on only the test set comparing Levenberg–Marquardt and Bayesian regularization with different numbers of neurons for two hidden layers (c).

The final RF model has been tested to controlled changes of the independent variables. Indeed, the calibration of the model generated the feature's importance of each independent variable: this allows establishing which variable has the higher predictive power in the model. The most important variable is the parking occupancy with a nonlinear trend in defining the Park and Ride demand: a given P&R share can correspond to both a low occupancy state (high capacity available, filling phase of the P&R site) and a high occupancy state (no more capacity, no more demand).

Regarding the features' importance in the case of ANN, this can be computed by adopting the formula in Kayri [46]. It is possible to observe (Figure 5) a flattening of the weights, especially when adopting the best architecture with two hidden layers. It seems to demonstrate that ANN is not able in the aggregated approach to capture correctly the different hierarchy between the explanatory variables of the model.

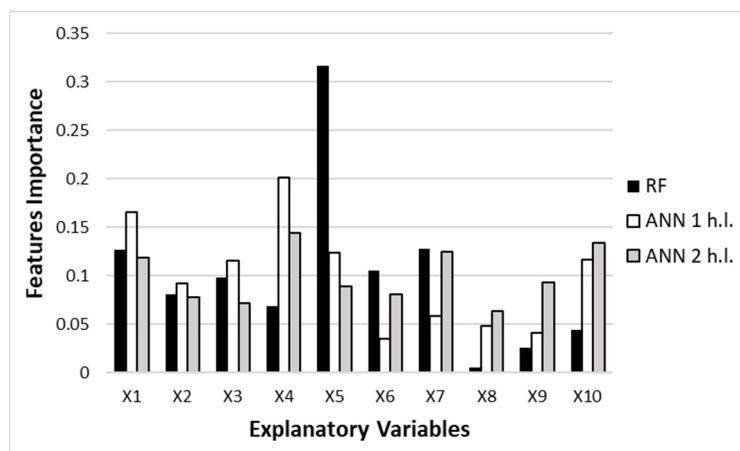


Figure 5. Comparison between features' importance of RF and ANN (with one or two hidden layers) in the aggregated approach.

Moving to the disaggregated approach, the database counts about 1.6 million records and 13 independent variables. As in the aggregated case, it has been divided according to a cross-validation method (70% training data set; 30% validation data set). The resulting binomial logit (Table 2), calibrated through the adoption of BIOGEME [42], assumes the following expression for the systematic utilities:

$$\begin{aligned} V_{P\&R} &= -2.57 - 2.18 x_5 + 0.0155 x_9 \\ V_{car} &= -0.01 x_4 + 0.269 x_8 \end{aligned} \quad (8)$$

Table 2. Calibrated values of binomial logit parameters and statistical significance of the model.

Parameters	Value	Std Error	t-Test	p-Value
ASC_P&R	-2.570	0.0225	-114.43	0.00
β_4	-0.010	0.00055	-18.26	0.00
β_5	-2.180	0.0261	-83.41	0.00
β_8	0.269	0.0163	16.47	0.00
β_9	0.0155	0.00486	3.18	0.00

Thus, the choice of Park and Ride is a function of a subset of the explanatory variables, the parking occupancy (x_5), and the number of stations (x_9); instead, the choice of the car is mainly due to the difference in travel time between car and Park and Ride (x_4) and to the location of the starting zone of the trip (inside or outside the Main Ring Road). However, although this is the most reliable random utility model obtained, the resulting accuracy is very low. The model is able to reproduce only the car choices, while it is not able to

reproduce any Park and Ride choices. This is probably due to the low number of trips by Park and Ride (about 16,000 with respect to the total sample of 1.6 million records). Additional behavioral models can be calibrated starting with the same data structure, for example, mixed logit able to consider for panel data (several trips related to the same vehicle as in the FCD sample) or a logit model with a penalized likelihood such as the weighted exogenous sample maximum likelihood [47]. However, problems in the model applicability or in the reproduction of the sample can, respectively, occur.

In contrast, by adopting the ANNs, it is possible to observe a high accuracy in reproducing not only car choices but also Park and Ride. It happens specifically in the case of the adoption of the Adam algorithm on a multi-layer feedforward neural network. This is shown in Figure 6 where the structure of the network can be formalized as $x-i-j-y$, with x the number of explanatory variables, y the output, i and j , respectively, the number of neurons in the first and in the second layer of ANN.

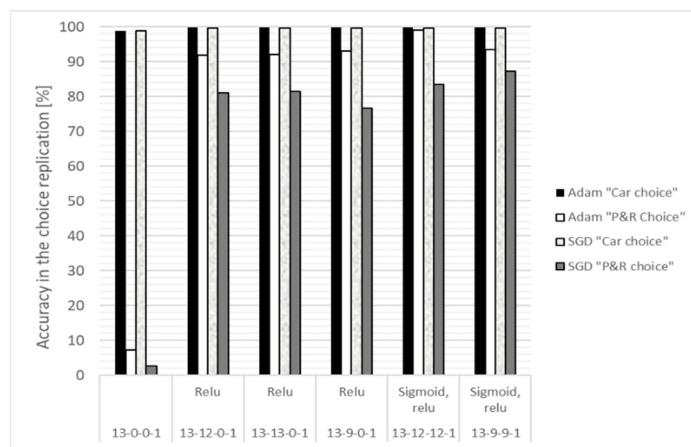


Figure 6. Results of different calibrated ANNs (disaggregated approach) for different activation function (Relu and Sigmoid, relu), numbers of hidden layers, and number of neurons in the hidden layers.

If $i-j$ are equal to zero, i.e., no hidden layers, a Perceptron network is obtained (architecture 13-0-0-1 in Figure 6), and the model is not able to reproduce Park and Ride but only car choices as obtained by the binary logit. The best architecture is the one with two hidden layers and 12 neurons for each layer (Architecture 13-12-12-1, Figure 6).

Regarding the adopted activation function, while a Sigmoid is always adopted as an output of the architecture, we tested a Sigmoid between the input layer and the first hidden layer and a ReLu between the first hidden layer and the second hidden layer (Figure 6).

Considering the objective of the multimodality, i.e., to build a suite of models that can be adopted to simulate multimodality for electric vehicles and that are mainly data-driven, it is possible to summarize the obtained results as follows:

1. In the case of the aggregated approach:
 - the resulting variable is the share of Park and Ride demand generated by each traffic zone of the city;
 - only machine learning techniques have been adopted since we are dealing with a regression model; the best form of machine learning was the random forest. However, the results of RF are also not sufficiently able to reproduce the observed outcomes: this is probably due to the low number of examples adopted as an input.
2. In the case of the disaggregated approach:
 - the resulting variable is the choice of Park and Ride or car to perform the trip; thus, it is a binary variable, and it is related to each single trip;

- both random utility models and machine learning (specifically neural networks) have been tested; however, random utility models were not satisfactory due to the oversampling of cars with respect to the Park and Ride choices. In contrast, specific architectures of neural networks combined with the right optimization algorithm can reproduce perfectly the sample despite the oversampling and thanks to the high number of examples adopted as an input.

4.2. Energy-Oriented Land-Use Model: Results

After training the model, we found four clusters based on trips on the weekdays (mainly residential, business/industrial, commercial, and mixed, Figure 7) and only three clusters on the weekend (mainly residential, mixed, and nightlife/leisure, Figure 8).

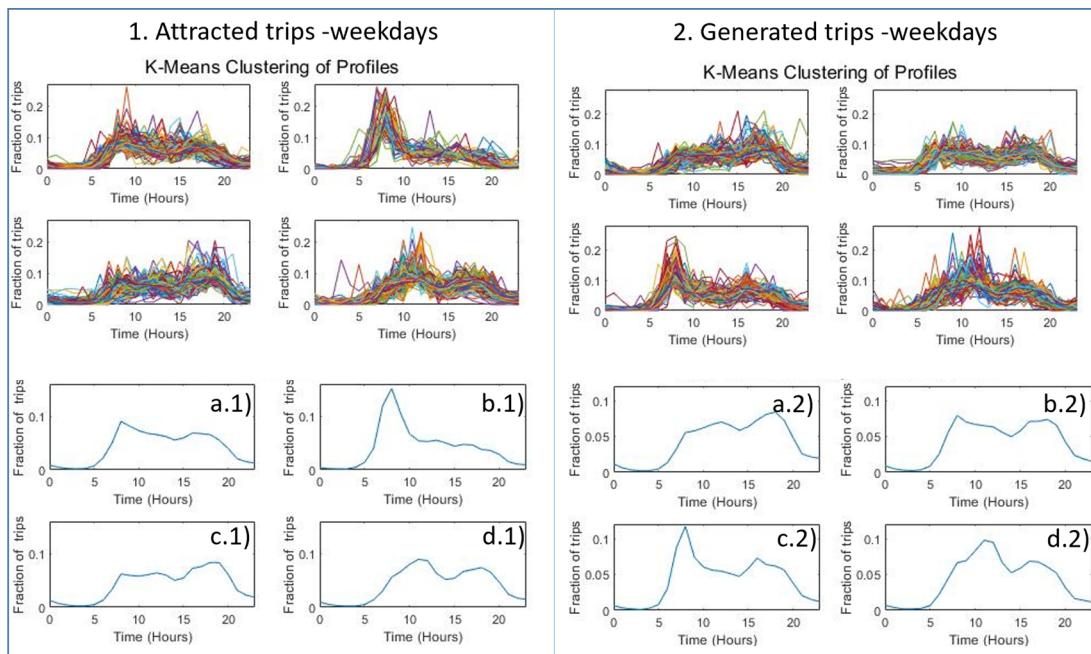


Figure 7. K-means clustering applied to (1) attracted trips and (2) generated trips on the weekdays. Resulting clusters: mixed-use (**a1,b2**), business/industrial (**b1,a2**), residential (**c1,c2**), and commercial (**d1,d2**).

Joining the clusters for both weekdays and weekend, as well as for generated and attracted trips, the city of Rome was 70% mixed land use, while only 8.4% of zones were classified as strictly residential, 8.6% strictly nightlife and leisure, 8.4% strictly commercial, and 2.1% strictly business/industrial. Inside the mixed land use, there is a prevailing residential function in 41% of the cases.

Results were validated comparing the resulting land use of each zone with census data:

- strictly residential zones correspond to intermediate/high values of population density (about 3500–9000 inhabitants per km^2) and residential buildings are more than 80% of the total buildings;
- strictly business/industrial zones have a low population density (mostly below 450 inhabitants per km^2), and residential buildings are less than 70% of the total buildings;
- strictly commercial zones showed average population density while quite high values of residential buildings: this is probably due to the typical configuration of shopping activities in Rome, where most of the shops are located on the ground floor of residential buildings facing the street.

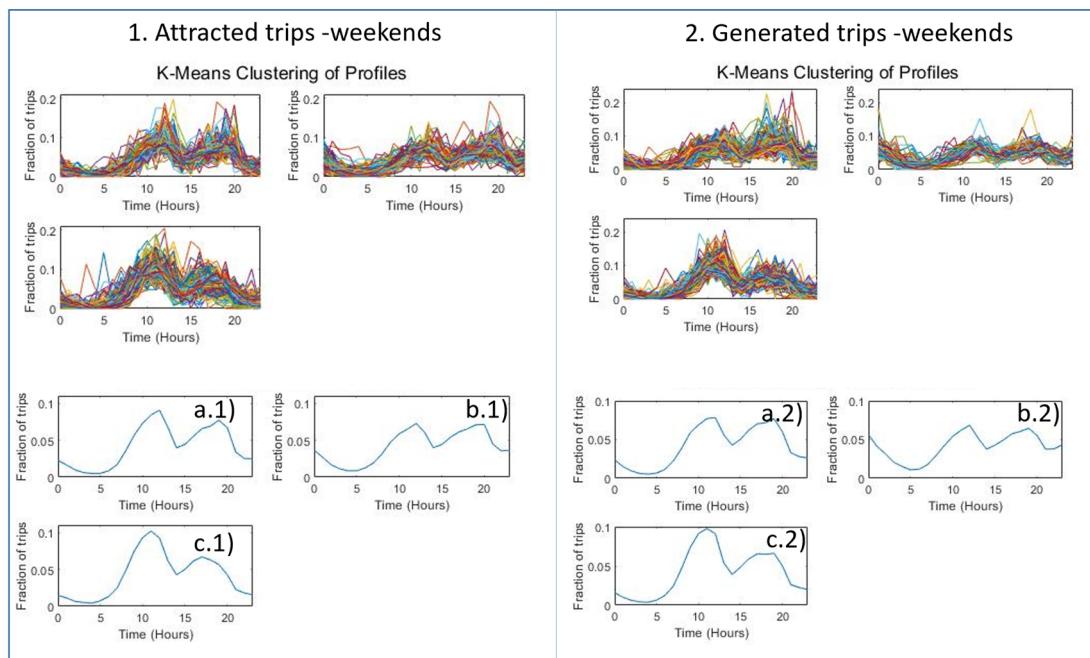


Figure 8. K-means clustering applied to (1) attracted trips and (2) generated trips on the weekends. Resulting clusters: residential (a1,a2), mixed-use (c1,a2), and nightlife and leisure (b1,b2).

Once the land use was defined, we moved to the computation of the energy demand for each zone of the city as described in Section 3.3. Most of the input required in (7) can be derived directly from FCD; for example, the recharge frequency is a function of the distance traveled (on average, each FCD in Rome travels for about 26 km in a day) and the kilometers to be traveled within the battery life (which are assumed equal to 170 km in this study). Parking times have also been exploited by FCD; these are a function of the land use of the arrival zone and can be different depending on whether the trip is home-based or not, as represented in Figure 9.

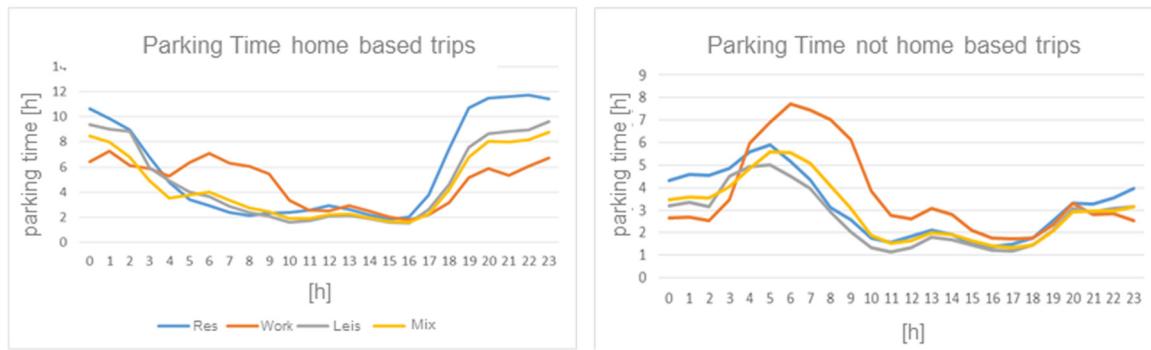


Figure 9. Parking time distribution during the daytime as a function of the land use of the arrival zone for home-based and not-home-based trips.

Below, we reported the application of (7) for two specific zones of the city characterized by a different land use: for the considered residential zone, a request of about 60 kWh in the evening was estimated, while in a typical working zone, a peak value of about 120 kWh in the morning time was estimated (Figure 10).

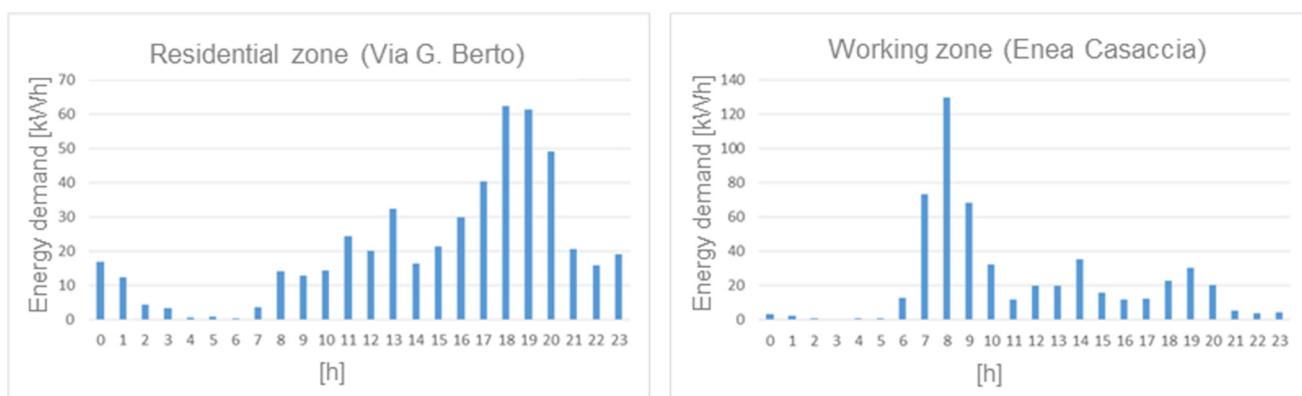


Figure 10. Examples of energy demand computation for two different land use zones in the city of Rome.

It is quite easy to convert these energy values into the entire energy demand if the penetration rate of FCD is known. However, we remind that the penetration rate of FCD in each zone could be different with respect to the average penetration rate, thus to obtain a refined estimation of the energy demand, it is suggested to compute a penetration rate for each zone. This can be done by comparing FCD flows with respect to traffic counts for different count sections located in the city.

5. Discussion and Policy Implications

The main research question explored in this work concerns how the future spread of EVs in urban areas can be tackled both from a mobility planning and an energy management perspective. Specifically, the proposed modeling framework is aimed to assess: (1) short-term measures driven by the characterization of the multimodal interchange between public transport and electric private mobility; (2) medium- to long-term measures based on the evaluation of the EVs' energy demand as a function of the land use.

The models developed here are currently under testing inside an agent-based simulator for electric urban mobility adopted by the National Agency for Energy and Environment in Italy (ENEA).

Regarding multimodality, the Park and Ride module generates the changes in the Park and Ride demand in the different zones of the city as a function of possible scenarios. Two scenarios are now implemented in the simulator: (1) a scenario promoting the efficacy and effectiveness of the public transport network (thus a pull policy to increase the potential shift from EVs to public transport); (2) a scenario introducing congestion charging schemes to limit the circulation of EVs in central areas of the city (thus a push policy to increase the adoption of public transport through the Park and Ride sites). The change in Park and Ride demand is followed by a redistribution of the final destinations and a definition of the new attraction points (EVs stopping in the Park and Ride sites). The agent model receives the new distribution of the attraction points of EVs and simulates the resulting mobility patterns and energy requests. Thus, integrating our models into a standard agent-based simulator can potentially help local governments and energy providers to form a correct assessment of measures for promoting multimodality and consequently for redistributing the charging load by EVs.

At the current stage, first results demonstrated the dependency of the potential modal shift to the Park and Ride sites by the residual capacity of the public transport system, together with a higher effectiveness of a pull policy (scenario 1) compared to a push policy (scenario 2).

Additional work is required to introduce in the multimodal model a dependency between the Park and Ride demand and possible incentives for promoting vehicle-to-grid schemes in the parking sites (e.g., through a discount on public transport subscription or on the electricity bill). However, the model structure can be easily extended to include this policy [28].

Regarding energy-oriented land use, it works by firstly relating the land use to mobility patterns, allowing improving traffic prediction and redefining zoning regulations, something that simply cannot be done by relating the land use to the usual existing statistics. Moreover, the main result of the proposed approach is in its capability to link the energy demand to the land use, thus providing information on how the energy demand is distributed in space and time for the different land use zones.

From a policy perspective, this information can be adopted by energy providers for energy demand forecasts or for the optimal location of charging stations. Moreover, it can be adopted by urban planners in order to understand the impact of the land use on energy demand, thus working on a more sustainable and energy-oriented plan for the city. Moving urban planning toward a more energy-saving approach is closely related to a smart-city viewpoint and to the objectives of the European Green Deal [48].

6. Conclusions and Further Developments

Electric mobility is the “closest” future of transport systems. Especially in urban areas, a strong increase of EVs is expected for private mobility and public transport, as well as for the other mobility solutions (e.g., mobility as a service solution). However, local governments and urban/transport planners require tools able to assess the sustainable development of electric mobility, especially avoiding energy congestion phenomena.

Promoting multimodal electric mobility and sustainable energy-oriented land use are two of the key challenges in energy congestion reduction towards a greater sustainability of urban context.

In this study, we developed a multimodal model that considers the change between EVs and public transport in Park and Ride sites. Moreover, a methodology to assess the land use of the city and to correlate the land use to the energy demand for mobility is proposed. Both have been modeled using data-driven approaches based on extended collections of floating car data (FCD) in the city of Rome (Italy). The final aim is to create a modeling framework able to assess short- and medium- to long-term measures for a sustainable development of e-mobility in urban areas.

Specifically, we focused on the development of Park and Ride models based on machine learning techniques and random utility theory, working with two different approaches. In the aggregated approach, where the output is the P&R share for each traffic zone of the city, results suggest that random forest outperforms ANN in reproducing the P&R share. In the disaggregated approach, ANN shows high accuracy in modeling the behavior of each vehicle, and it is able to perfectly reproduce the choice between Park and Ride and a car. In contrast, the behavioral model based on random utility theory is not satisfactory due to the oversampling of cars with respect to the Park and Ride choices. There are still several research perspectives for the problem, for example, expanding data sources for calibration and validation purposes (e.g., mobile phone data) or testing additional sophisticated machine learning techniques.

Regarding the procedure to compute the energy demand as a function of the land use, firstly the land use for each zone of the city has been defined, adopting FCD trip profiles both for home-based and not home-based trips. The results reveal four clusters on the weekdays (mainly residential, business/industrial, commercial, and mixed), and three clusters on the weekend (mainly residential, mixed, and nightlife/leisure). Specifically, joining the clusters for both weekdays and the weekend, as well as for generated and attracted trips, the city of Rome had 70% mixed land use, with 8.4% of zones strictly residential, 8.6% strictly nightlife and leisure, 8.4% strictly commercial, and 2.1% strictly business/industrial. Finally, we provided first estimates of energy demand for each zone of the city. The energy demand estimates could be improved in further developments of the study using the history of each trip in the process. This can potentially provide more refined values for the energy computation.

All models and methods here proposed are currently being implemented in an agent-based simulator for e-mobility to be adopted to guide sustainable urban planning considering both the demand for mobility and the demand for energy.

Author Contributions: conceptualization, M.N.; data curation, M.F. and R.D.V.; funding acquisition, G.V.; methodology, M.N.; resources, C.L.; supervision, C.L. and G.V.; validation, M.F. and R.D.V.; writing—original and final draft, M.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by ENEA through the National Research Project on Electric Systems 2015–2017, AdP Mise-Enea 2015–2017.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: GTFS data for the city of Rome are open data available in <https://dati.comune.roma.it/>; Census data for the city of Rome are open data available in <https://www.istat.it>; Floating Car Data adopted in this study are the property of ENEA.

Acknowledgments: The authors acknowledge funding from the Italian Ministry of Education, University and Research (MIUR), in the frame of the Departments of Excellence Initiative 2018–2022, attributed to the Department of Engineering of Roma Tre University.

Conflicts of Interest: The authors declare no conflict of interest.

References

- European Commission, Directorate-General Mobility and Transport. Transport in the European Union—Current trends and issues. Available online: <https://ec.europa.eu/transport/sites/transport/files/2019-transport-in-the-eu-current-trends-and-issues.pdf> (accessed on 17 January 2020).
- Liberto, C.; Valenti, G.; Lelli, M.; Ferrara, M.; Nigro, M. Evaluation of the impact of e-mobility scenarios in large urban areas. In Proceedings of the 2017 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS), Naples, Italy, 26–28 June 2017; pp. 379–384.
- IEA—International Energy Agency, Global EV Outlook. Available online: <https://www.iea.org/reports/global-ev-outlook-2020> (accessed on 20 April 2021).
- Deloitte “Electric Vehicles—Setting a Course for 2030”. 2020. Available online: https://www2.deloitte.com/content/dam/Deloitte/rs/Documents/about-deloitte/DI_Electric-Vehicles.pdf (accessed on 12 February 2021).
- Canizes, B.; Soares, J.; Costa, A.; Pinto, T.; Lezama, F.; Novais, P.; Vale, Z. Electric Vehicles’ User Charging Behaviour Simulator for a Smart City. *Energies* **2019**, *12*, 1470. [[CrossRef](#)]
- Krajzewicz, D.; Erdmann, J.; Behrisch, M.; Bieker-Walz, L. Recent Development and Applications of SUMO—Simulation of Urban MObility. *Int. J. Adv. Syst. Meas.* **2012**, *5*, 128–138.
- Axhausen, K.W. *The Multi-Agent Transport Simulation MATSim*; Ubiquity Press: London, UK, 2016.
- Kurczveil, T.; López, P.A.; Schnieder, E. Implementation of an Energy Model and a Charging Infrastructure in SUMO. In Proceedings of the Transactions on Petri Nets and Other Models of Concurrency XV, Berlin, Germany, 7 November 2014; pp. 33–43.
- Luin, S.P.; Al-Mansour, F. *Microsimulation of Electric Vehicle Energy Consumption*; Energy Scenario Toward 2030; Elsevier: New York, NY, USA, 2019; Volume 174, pp. 24–32.
- Fiori, C.; Marzano, V.; Punzo, V.; Montanino, M. Energy Consumption Modeling in Presence of Uncertainty. *IEEE Trans. Intell. Transp. Syst.* **2020**, *1*–12. [[CrossRef](#)]
- Farid, A.M. A Hybrid Dynamic System Model for Multimodal Transportation Electrification. *IEEE Trans. Control. Syst. Technol.* **2016**, *25*, 940–951. [[CrossRef](#)]
- Kassakian, J.G. *The Future of the Electric Grid: An Interdisciplinary MIT Study*; MIT Press: Cambridge, MA, USA, 2011. Available online: http://web.mit.edu/mitei/research/studies/documents/electric-grid-2011/Electric_Grid_Full_Report.pdf (accessed on 10 May 2021).
- Al Junaibi, R.; Farid, A.M. A method for the technical feasibility assessment of electrical vehicle penetration. In Proceedings of the 2013 IEEE International Systems Conference (SysCon), Orlando, FL, USA, 15–18 April 2013; pp. 606–611.
- Ferrara, M.; Monechi, B.; Valenti, G.; Liberto, C.; Nigro, M.; Biazzo, I. A simulation tool for energy management of e-mobility in urban areas. In Proceedings of the 2019 6th International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS), Cracow, Poland, 5–7 June 2019; pp. 1–7.
- RAND Europe. Modelling Park-and-Ride. Understanding Multi-Modal Travel. 2007. Available online: www.rand.org/randeurope (accessed on 26 February 2021).
- Zhao, X.; Li, Y.; Xia, H. Behavior decision model for park-and-ride facilities utilization. *Adv. Mech. Eng.* **2017**, *9*, 1–9. [[CrossRef](#)]

17. Sharma, B.; Hickman, M.; Nassir, N. Park-and-ride lot choice model using random utility maximization and random regret minimization. *Transportation* **2017**, *46*, 217–232. [CrossRef]
18. Ma, T.-Y. Bayesian networks for multimodal mode choice behaviour modelling: A case study for the cross border workers of Luxembourg. In Proceedings of the 18th Euro Working Group on Transportation, EWGT 2015, Delft, The Netherlands, 14–16 July 2015.
19. Thanh, T.T.M.; Ly, H.-B.; Pham, B.T. A Possibility of AI Application on Mode-choice Prediction of Transport Users in Hanoi. In *Lecture Notes in Civil Engineering*; Springer Science and Business Media LLC: Cham, Switzerland, 2019; Volume 54, pp. 1179–1184.
20. Guan, Y.; Wang, Y.; Yan, X.; Guo, H.; Zhou, Y. A Big-Data-Driven Framework for Parking Demand Estimation in Urban Central Districts. *J. Adv. Transp.* **2020**, *2020*, 1–13. [CrossRef]
21. Xiao, J.; Lou, Y.; Frisby, J. How likely am I to find parking?—A practical model-based framework for predicting parking availability. *Transp. Res. Part B Methodol.* **2018**, *112*, 19–39. [CrossRef]
22. Schuster, T.; Volz, R. Predicting Parking Demand with Open Data. In Proceedings of the Transactions on Petri Nets and Other Models of Concurrency XV; Springer Science and Business Media LLC: Cham, Switzerland, 2019; pp. 218–229.
23. Karolemeas, C.; Tsigdinos, S.; Tzouras, P.; Nikitas, A.; Bakogiannis, E. Determining Electric Vehicle Charging Station Location Suitability: A Qualitative Study of Greek Stakeholders Employing Thematic Analysis and Analytical Hierarchy Process. *Sustainability* **2021**, *13*, 2298. [CrossRef]
24. Wu, Z.; Zhuang, Y.; Zhou, S.; Xu, S.; Yu, P.; Du, J.; Luo, X.; Abbas, G. Bi-Level Planning of Multi-Functional Vehicle Charging Stations Considering Land Use Types. *Energies* **2020**, *13*, 1283. [CrossRef]
25. Orsi, F. On the sustainability of electric vehicles: What about their impacts on land use? *Sustain. Cities Soc.* **2021**, *66*, 102680. [CrossRef]
26. Zhang, R.; Zhang, J.; Long, Y.; Wu, W.; Liu, J.; Jiang, Y. Long-term implications of electric vehicle penetration in urban decarbonization scenarios: An integrated land use–transport–energy model. *Sustain. Cities Soc.* **2021**, *68*, 102800. [CrossRef]
27. Tizghadam, A.; Khazaei, H.; Moghaddam, M.H.Y.; Hassan, Y. Machine Learning in Transportation. *J. Adv. Transp.* **2019**, *2019*, 1–3. [CrossRef]
28. Ferrara, M.; Liberto, C.; Nigro, M.; Trojani, M.; Valenti, G. Multimodal choice model for e-mobility scenarios. *Transp. Res. Procedia* **2019**, *37*, 409–416. [CrossRef]
29. Hastie, T.; Tibshirani, R.; Friedman, J. The Elements of Statistical Learning: Data Mining, Inference and Prediction. *Math. Intell.* **2005**, *27*, 83–85.
30. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
31. Zhang, Y.; Haghani, A. A gradient boosting method to improve travel time prediction. *Transp. Res. Part C Emerg. Technol.* **2015**, *58*, 308–324. [CrossRef]
32. Kazic, B.; Mladenović, D.; Košmerlj, A. Traffic Flow Prediction from Loop Counter Sensor Data using Machine Learning Methods. In Proceedings of the International Conference on Vehicle Technology and Intelligent Transport Systems, VEHITS, Lisbon, Portugal, 20–22 May 2015; SCITEPRESS—Science and Technology Publications: Setúbal, Portugal, 2015; pp. 119–127.
33. Zarei, N.; Ghayour, M.A.; Hashemi, S. Road Traffic Prediction Using Context-Aware Random Forest Based on Volatility Nature of Traffic Flows. In *Transactions on Petri Nets and Other Models of Concurrency XV*; Springer Science and Business Media LLC: Cham, Switzerland, 2013; Volume PART 1, pp. 196–205.
34. Ghasri, M.; Rashidi, T.H.; Waller, T. Developing a disaggregate travel demand system of models using data mining techniques. *Transp. Res. Part A Policy Pr.* **2017**, *105*, 138–153. [CrossRef]
35. Ashqar, H.I.; Elhenawy, M.; Almanna, M.; Ghanem, A.; Rakha, H.A.; House, L. Modeling bike availability in a bike-sharing system using machine learning. In Proceedings of the 2017 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS), Naples, Italy, 26–28 June 2017; pp. 374–378. [CrossRef]
36. Liu, X.; Liu, K.; Li, M.; Lu, F. A ST-CRF Map-Matching Method for Low-Frequency Floating Car Data. *IEEE Trans. Intell. Transp. Syst.* **2016**, *18*, 1241–1254. [CrossRef]
37. Hensher, D.A.; Ton, T.T. A comparison of the predictive potential of artificial neural networks and nested logit models for commuter mode choice. In *Transportation Research Part E: Logistics and Transportation Review*; Elsevier: New York, NY, USA, 1999; Volume 36, pp. 155–172.
38. Cantarella, G.E.; de Luca, S. Multilayer feedforward networks for transportation mode choice analysis: An analysis and a comparison with random utility models. *Transp. Res. Part C* **2005**, *13*, 121–155. [CrossRef]
39. Nijkamp, P.; Reggiani, A.; Tsang, W.F. Comparative modelling of interregional transport flows: Applications to multimodal European freight transport. *Eur. J. Oper. Res.* **2004**, *155*, 584–602. [CrossRef]
40. Gazder, U.; Ratnou, N.T. A new logit-artificial neural network ensemble for mode choice modeling: A case study for border transport. *J. Adv. Transp. J. Adv. Transp.* **2015**, *49*, 855–866. [CrossRef]
41. Dabiri, S.; Heaslip, K. Inferring transportation modes from GPS trajectories using convolutional neural network. *Transp. Res. Part C Emerg. Technol.* **2018**, *86*, 360–371. [CrossRef]
42. Bierlaire, M. A short introduction to PandasBiogeme. *Tech. Rep. Transp. Mobil. Lab. ENAC EPFL* 2020. Available online: <https://transp-or.epfl.ch/documents/technicalReports/Bier20.pdf> (accessed on 13 January 2021).
43. Caceres, N.; Benítez, F.G. Supervised Land Use Inference from Mobility Patterns. *J. Adv. Transp.* **2018**, *2018*, 1–12. [CrossRef]

44. MacQueen, J. Some methods for classification and analysis of multivariate observations. In Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, CA, USA, 21 June 1965; Volume 1, pp. 281–297.
45. Nigro, M.; Cipriani, E.; Del Giudice, A. Exploiting floating car data for time-dependent Origin–Destination matrices estimation. *J. Intell. Transp. Syst.* **2017**, *22*, 159–174. [[CrossRef](#)]
46. Kayri, M. Predictive Abilities of Bayesian Regularization and Levenberg–Marquardt Algorithms in Artificial Neural Networks: A Comparative Empirical Study on Social Data. *Math. Comput. Appl.* **2016**, *21*, 20. [[CrossRef](#)]
47. Train, K. *Discrete Choice Methods with Simulation*; Cambridge University Press. Available online: <https://eml.berkeley.edu/books/train1201.pdf> (accessed on 13 January 2021).
48. European Commission (EC). *The European Green Deal: Striving to Be the First Climate Neutral Continent*; European Commission: Brussels, Belgium, 2019. Available online: <https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1596443911913&uri=CELEX:52019DC0640#document2> (accessed on 26 August 2020).