NILMEV: Electric Vehicle disaggregation for residential customer energy efficiency incentives

Camilo Mariño *
Electrical Engineering Institute
Universidad de la República
Montevideo, Uruguay
Email: cmarino@fing.edu.uy

Guillermo Cossio *
Electrical Engineering Institute
Universidad de la República
Montevideo, Uruguay
Email: gcossio@fing.edu.uy

Pablo Massaferro
Electrical Engineering Institute
Universidad de la República
Montevideo, Uruguay
Email: pmassaferro@fing.edu.uy

Matias di Martino
Electrical Engineering Institute
Duke University
NC, United State
Email: matias.di.martino@duke.edu

Alvaro Gómez
Electrical Engineering Institute
Universidad de la República
Montevideo, Uruguay
Email: agomez@fing.edu.uy

Alicia Fernandez
Electrical Engineering Institute
Universidad de la República
Montevideo, Uruguay
Email: alicia@fing.edu.uy

***Abstract*—Due to its impact on household energy use and the adoption of renewable energies, the intelligent management of the power consumption of electric vehicles (EVs) is of great relevance. In the context of widespread clean energy adoption and growing environmental concerns, generating incentives through discounted rates for intelligent residential EV power consumption requires algorithms capable of measuring loads in a disaggregated manner. The deployment of smart meter networks offers the possibility of applying machine learning techniques to estimate EV residential consumption. This work presents an efficient algorithm for the Non Intrusive Load Monitoring (NILM) of EV consumption, which is an adaptation of a method previously proposed for high-powered water heaters. Its performance is compared with methods based on deep neural networks. Results from an actual power demand dataset are discussed, and a comparative analysis is carried out against billing rules based on time slots and historical power consumption data.**

***Index Terms*—NILM, Electric Vehicles, load disaggregation**

I. Introduction

The estimation of household electricity consumption, broken down by use, is essential information for both generation and distribution companies and users. In a context of change in the energy matrix and renewable energy adoption, it is critical to generate incentives for the intelligent use of resources, with the accompanying economic and environmental impact.

Analyzing household energy consumption patterns allows for identifying activities that can be deferred over time to adjust energy demand to generation. One of the most common incentives for customers is differentiated rate plans, typically by time slots. The hourly consumption information can be obtained directly from the meter installed at each customer's network access point. However, other energy rate plans could be envisioned, taking into account the use given to energy.

Electric transport is a growing area worldwide. EV charging will significantly impact all the electrical grid's processes, with demand management taking on particular relevance. The power demand of any system capable of accumulating energy is potentially deferrable. This property is one of the reasons why energy companies are interested in analyzing the consumption patterns of electric vehicles (EVs).

Data from intrusive meters and aggregated consumption meters for the whole household makes it possible to address the problem with supervised disaggregation strategies based on deep learning [1], [2], [3], something that is difficult when there are no intrusive measurements available.

The availability and characteristics of public or private datasets, with or without labels, impose restrictions on model selection and training. In a recent paper by [4], the problem of disaggregation of high-consumption water heaters is addressed with a suitable strategy for the case where intrusive data is unavailable to train disaggregation algorithms with ground truth. Given the level of power used by those heaters and the shape of their activations, a similar approach could be used to estimate EV consumption when intrusive measures are unavailable.

II. Proposed Approach

The consumption curves of measured EVs and the water heaters follow similar on-off patterns. In Uruguay, Fig. 12, [5], water heaters have consumption curves, similar to Costa Rica, Fig. 2 [4]. In both cases, the water heaters have ON-OFF characteristics similar to those of electric vehicles, see Fig. 1. All of them meet the working hypotheses of the algorithm. This similarity

* These authors contributed equally to this work.

makes it reasonable to think that disaggregation algorithms with good performance for high-power water heaters can be adapted for the disaggregation of EV consumption. This approach can be appropriated when there is an incipient EV fleet and little historical information on consumption patterns. In this context, we set out to adapt the algorithm proposed by [4] to disaggregate EV, taking into account its simplicity and the fact that it does not require large volumes of intrusive data for its parameterization. The code developed while designing and evaluating the algorithm will be made available.

The algorithm is comprised of three distinct stages. In the first stage, the active power consumed by the EV during charging periods is estimated. In the second one, the algorithm detects large jumps in the building's power consumption, which will be used to demarcate charging periods. Finally, the disaggregated average power of the EV during the charging periods is defined.

The following hypotheses are assumed when disaggregating loads with this algorithm: a) The rated charging power of the EV is greater than the rated power of all other appliances in a residential building. b) The load curve of the EV is quasi-rectangular. Its power consumption is approximately constant while it charges. c) The time series includes EV consumption (i.e., an EV charging period appears at some point in the series).

The different stages of the proposed algorithm are described below.

1) Stage 1 - EV power estimation: The estimation of EV charging power is based on the histogram of power samples from the aggregate household as in [4]. This is justified, for water heaters, because the disaggregated loads usually have the same power consumption and are turned on many times throughout the day, while most domestic loads have a more variable power consumption. Therefore, a peak with more samples than the rest should be visible in the power sample histogram, corresponding to the usual power consumption of the EV. This is particularly true when high-frequency samples are available, but it is still noticeable with a sample period of 15 minutes for EV charging patterns that usually last for many hours.

The following criterion is used to determine if bin i (the set of samples that form the i-th column of the histogram) is an outlier to the typical consumption of the rest of the appliances in the building:

$d_i \geq d_{Q3} + 1.5\,(d_{Q3} - d_{Q1})$ where d_i is the density of bin i, d_{Q3} is the third density quartile and d_{Q1} is the first one.

The histogram is computed with all samples above a certain power $P_{threshold}$, to filter out the samples from periods where the EV was not being charged. The *estimation_method* parameter is used to choose whether the samples or the jumps between samples are used. After computing the histogram, the algorithm checks if the bin with the most samples is an outlier. If so, the average power of the samples within it is taken as the estimate of P_{EV}.

As mentioned at the beginning of this section, it is assumed that there is at least one EV charging period in the time series. Therefore, if no outlier column is detected, the algorithm takes the median of all samples above $P_{threshold}$ as the estimate.

2) Stage 2 - Charging period detection: At this stage, the algorithm searches for jumps in the household's power consumption, defined as follows:

I. There is a positive jump at sample n if: $x[n] - x[n - n_0] \geq k_{jump} \cdot P_{EV}$

II. There is a negative jump at sample n if: $x[n] - x[n - n_0] \leq -k_{jump} \cdot P_{EV}$

with P_{EV} being the power of the EV, either as estimated in Stage II-1 or known beforehand, $n_0 \geq 1$ a number of samples and $k_{jump} \in (0, 1)$ a real number akin to a sensibility.

It is important to note that if positive jumps are shifted $\frac{n_0}{2}$ samples forward, and negative jumps are shifted back by the same amount, the detected jumps, and the ones effectively seen in the household's consumption line up better.

After detecting these jumps, we define a charging period as the set of samples between a positive jump and the next negative jump, having previously filtered the detected jumps.

The filtering process consists of several steps: a) If there are many successive positive jumps, the algorithm discards all but the last one. As for successive negative jumps, only the first one is kept. b) Considering the available data and the opinion of domain experts, at present, and with the chargers and vehicles available in Uruguay, the duration of a charging cycle does not exceed 24 hours in any case, a time we call T_{max}. Therefore, if a charging period exceeds this time, it is assumed that it has resulted from two false alarms (one positive and one negative) and is entirely discarded. c) If any sample in a charging period is less than $k_P \cdot P_{EV}$ with k_P a real number $\in (0, 1)$, it stops being considered as part of the charging period.

3) Stage 3 - Total consumption estimation: In this last stage, power consumption during the charging periods must be defined. As is the case in [4], the following rule is used:

$$\hat{P}_{EV}[n] = min(\{P_{EV}, P_{house}[n]\})$$

The algorithm predicts a rectangular consumption curve for the EV while avoiding nonsensical situations where the EV would have a power consumption higher than the aggregate household power.

4) Parameters: In addition to the previously mentioned parameters, the algorithm's behavior can be altered in other ways with the following parameters: a) *std_values*: Boolean. Adjust the power estimate from stage 1 to the closest value of the standard power ratings EVs can have in Uruguay. These standard values are determined by the circuit breaker tripping current multiplied by the grid's voltage, which is 230V. The resulting values are 7360W, 6900W, 3680W, and 2300W. b) $P_{threshold}$: Minimum power that a sample must have to be considered when computing the histogram from Stage 1.

III. Electric Vehicle DataSets

Three data sets were generated for use in the design and evaluation of EV charge disaggregation algorithms, all with a sample period of 15 minutes:

1) **AggregateEV**: A dataset consisting of remote measurements of consumption of 70 residential clients who possess an EV, with the measured period being around a year but varying between clients. Among these, 6 correspond to meters that measure EV chargers exclusively.
2) **AggregateNoEV**: A dataset consisting of remote measurements of consumption of 377 residential clients who did not have an EV.
3) **SyntheticEV**: A two-part synthetic dataset created through the following process:
    a) Identify the meters from **AggregateEV** whose power consumption comes entirely from the EV. In total, 6 meters were identified, and all of their loads have an instantaneous power of around 7kW.
    b) Randomly divides the EV meters into two subsets of three, A and B. Also, randomly assign half of the non-EV households from **AggregateNoEV** to subset A and the rest to subset B.
    c) For each non-EV household, randomly select one of the 3 EV meters from the same subset and apply a time shift to it so that both time series overlap. Then obtain a synthetic aggregate household with an EV by adding both series.

    Through this method, 377 synthetic household time series were generated, with subset A having 188 and subset B having 189. The reason for splitting the dataset in this way is that one can be used as the training set for the machine learning algorithm, while the other works as a test set. This is better than randomly dividing the whole dataset as it ensures that the same EV time series does not appear both in training and testing, which could lead to various models "memorizing" all possible EV consumption curves and having good test performance but with poor generalization to non-seen data.

IV. Experiments

A. *NILMEV performance on synthetic data.*

The proposed algorithm was compared against several state-of-the-art machine-learning-based disaggregators. The disaggregators used were: InceptionTime[6], InceptionTimePlus[6], BiLSTM [2], Denoising Autoencoder(DAE)[2], ResNet [7], Seq2Seq [8].

Although the proposed algorithm has a structure based on decision rules that do not follow the design of a traditional ML model, it requires that the choice of parameters follow the same principles as for any statistical learning model. Therefore, care must be taken not to overfit the training data, so performance must be evaluated in a test data set not seen in training for all the compared models.

With this in mind, the following testing procedure was devised. All disaggregation algorithms were trained on subset A of the synthetic dataset. A random search was conducted to find the proposed algorithm's parameters that minimize the Mean Absolute Error (MAE) on subset A. The synthetic households on subset B were disaggregated with the models mentioned in the previous two steps. The sample-wise power MAE was calculated for each model. The results of both disaggregation instances are shown in Table I and in Fig. 1. On the synthetic dataset, NILMEV appears to have outperformed all neural-network-based approaches. There are some caveats.

We observed that the trained NNs struggle to predict that no power is being used during non-charging periods, yielding small wattages even when no EV appears to be charging. This contrasts with NILMEV, which always predicts zero power consumption outside detected charging periods. Conversely, because of its relative simplicity, NILMEV provides only a rough estimate of power use while the EV is charging. In contrast, the NN-based approaches tend to perform better during these periods.

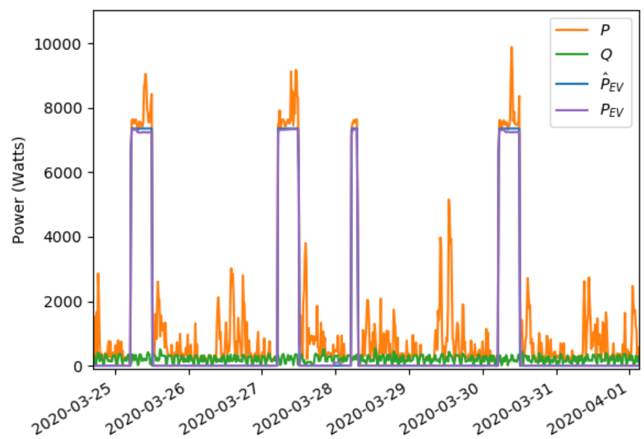

Fig. 1: Disaggregation of a synthetic household with NILMEV. The algorithm shows an impressive performance at detecting charging periods while still providing a decent estimate of the power demand during them.

This means that the MAE depends on the proportion of the time series in which the EV charges. Therefore, NILMEV would be more performant in cases where the EV is charged sparingly (e.g., a charger located in a residential household). At the same time, NNs would work best with heavy EV use. The energy use by the synthetic households was calculated from the disaggregation yielded by the algorithm. The disaggregation errors are shown in Fig. 2. It is important to remember that, in reality, EV charging powers can vary widely, so it should not be assumed that the

	NILMEV	InceptionTime	InceptionTimePlus	BiLSTM	DAE	ResNet	Seq2Seq
Trained on Subset A	5.54	30.9	30.6	21.9	30.2	16.4	31.0
Trained on Subset B	24.5	33.1	30.0	28.2	92.3	31.4	61.1
Average	15.0	32.0	30.2	25.0	61.3	23.9	46.0

TABLE I: Model performance evaluations. In the first row, the MAE over Subset B in models trained on Subset A is reported. In the second row, the sets are interchanged. The last row shows the mean performance over both sets.

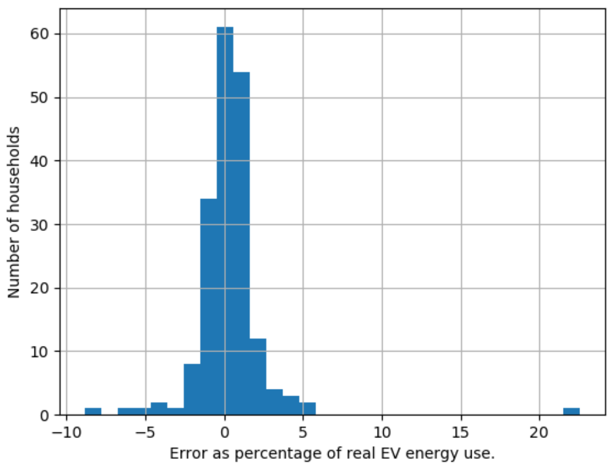

Fig. 2: Disaggregation error for households from subset A, using the best parameters found for subset B. A positive error means that EV consumption is being overestimated.

parameters found for these particular vehicles are the optimal choice for every EV.

B. NILMEV performance on a real dataset.

Although the penetration of EVs in Uruguay is still in its early stages, there are policies put in place by the national energy company (UTE) to stimulate the adoption and use of EVs. These policies include service agreements that offer discounted rates for charging during off-peak hours. As there is no sub-metering available, which would allow EV energy use to be measured separately from the rest of the household, the energy consumption of the EV must be estimated based on historical consumption data and the rate plan chosen by the customer.

The availability of the **AggregateEV** dataset allows for a comparative statistical analysis of the proposed disaggregation algorithm's results against the currently used criteria based on time slots and changes in historical consumption. It also allows us to analyze whether there is bias or differences between the results depending on the charging plan chosen by the client.

The disaggregation algorithm was run on **real data** of service points (SPs) that have EVs (**AggregateEV** dataset from section III). Disaggregation was also calculated using criteria that consider different scenarios based on time slots and historical measurements, and the different consumption values obtained were compared. Only scenarios with two or three-slot regimes were compared, as EV consumption is not discounted if the customer has a flat rate.

The criteria are as follows:

- **Two slot regime:** 40% of off-peak consumption is considered EV consumption.
- **Three slot regime:** 90% of consumption during the lowest-demand period is considered EV consumption.

These criteria are the ones currently used by the electrical utility company in Uruguay to offer discount rates for EV charging. [1] . The user chooses "Peak hours" from any of the three periods of 4 consecutive hours between 17:00 and 23:00. Due to this, the disaggregation for double slot regimes is separated into three groups. The utility company fixes the lowest-demand hours as the period from 0:00 to 7:00. Fig. 3 shows the results of applying the NILMEV disaggregation algorithm with the best parameters for Subset A. The algorithm shows an excellent ability to locate EV activity and precisely detect the absence of EV charging. In Fig. 4, the disaggregation of NILMEV is compared

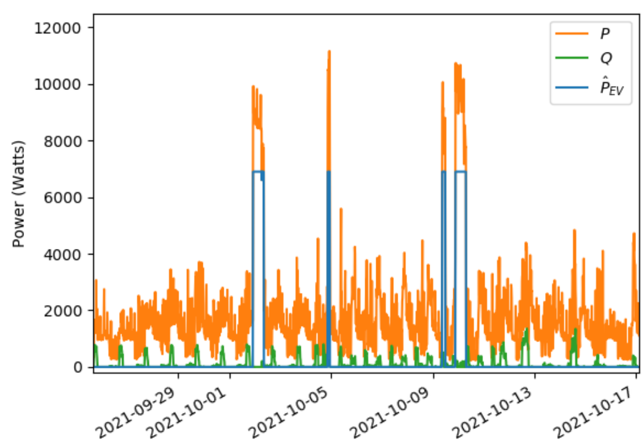

Fig. 3: Time series of a real Service Point (SP) with its disaggregation by NILMEV, using Subset A's best parameters. Note how the algorithm can identify extended periods where the EV does not appear to be loaded.

with the energy estimated under the different time slot regimes. It can be seen that the energy use allocated by current criteria is much greater than the NILMEV disaggregation algorithm. This is most likely because current criteria do not detect the absence of an EV and therefore function in practice as an extra

[1]https://portal.ute.com.uy/clientes/soluciones-para-el-hogar/planes-hogar/plan-inteligente

discount for *all* off-peak consumption instead of as an incentive for using a particular appliance. As an example, Fig. 4 shows that NILMEV does not disaggregate any energy use, which is consistent with the user's consumption pattern in that month. However, the estimates made with the time slot criteria do not reflect that. Fig. 5 shows the disaggregation for another user who

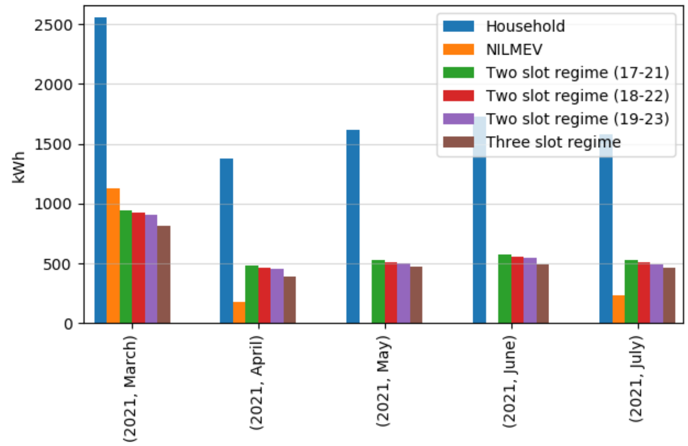

Fig. 4: Disaggregated energy per month for SP in Fig. 3, with NILMEV and current UTE criteria. We highlight the algorithm's ability to detect periods in which the EV does not appear to be charged.

has a three-slot regime in place. The user charges their vehicle in the lowest-demand hours to take advantage of lower prices. However, the jump in power consumption, along with the shape of the nightly consumption curve, suggests that they may be using other high-powered appliances in the same period. As Fig. 5 shows, the time-slot-based disaggregation estimates a higher energy use, leading us to believe that it may be counting other appliances used during the lowest-demand period as part of the EV.

V. Conclusions

The article shows that it is possible to accurately disaggregate EV power consumption with a simple algorithm suitable for situations with little to no sub-metered data. We show that the proposed algorithm can achieve performance levels similar to state-of-the-art disaggregation algorithms based on deep neural networks when trained and evaluated on a synthetic database generated by the research team. The ability to detect charging activity is highlighted, particularly the accuracy with which the absence of activity is detected. The evaluation of the results in the real dataset shows a similar behavior regarding detecting charging periods. The comparative analysis with currently used criteria based on historical data and discounted rate plans confirms the relevance of providing methods that estimate EV consumption in a disaggregated manner.

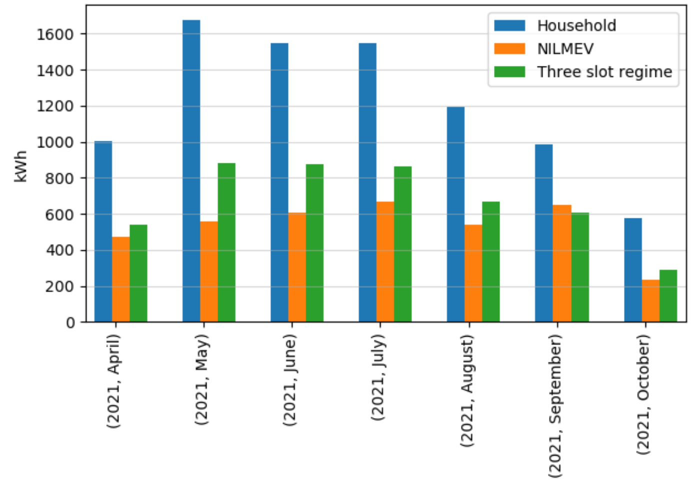

Fig. 5: Disaggregated energy per month for a user with a three-slot regime, using NILMEV and current UTE criteria. Three-slot disaggregation consistently overestimates EV energy use.

Acknowledgment

The authors thank UTE for their research support and for sharing their expertise and access to the datasets. In particular, authors greatly acknowledge engineers: L. Corna, G. Hernández, F. Santomauro, G. Alvez, S. Alpuy, S. Garabedian, P. Cabalo y M. Rey.

References

[1] C. Mariño, E. Masquil, F. Marchesoni, A. Fernandez, and P. Massaferro, "Nilm: Multivariate dnn performance analysis with high frequency features," in *2021 IEEE PES Innovative Smart Grid Technologies Conference - Latin America (ISGT Latin America)*, 2021, pp. 1–5.

[2] J. Kelly and W. Knottenbelt, "Neural nilm: Deep neural networks applied to energy disaggregation," in *Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments*. ACM, 2015, pp. 55–64.

[3] X. Wang, G. Tang, Y. Wang, S. Keshav, and Y. Zhang, "Evsense: a robust and scalable approach to non-intrusive ev charging detection," in *Proceedings of the Thirteenth ACM International Conference on Future Energy Systems*, 2022, pp. 307–319.

[4] T. Zufferey, G. Valverde, and G. Hug, "Unsupervised disaggregation of water heater load from smart meter data processing," in *The 12th Mediterranean Conference on Power Generation, Transmission, Distribution and Energy Conversion (MEDPOWER 2020)*, vol. 2020. IET, 2020, pp. 283–288.

[5] J. Chavat, S. Nesmachnow, J. Graneri, and G. Alvez, "Ecd-uy, detailed household electricity consumption dataset of uruguay," *Scientific Data*, vol. 9, no. 1, pp. 1–16, 2022.

[6] H. I. Fawaz, B. Lucas, G. Forestier, C. Pelletier, D. F. Schmidt, J. Weber, G. I. Webb, L. Idoumghar, P.-A. Muller, and F. Petitjean, "InceptionTime: Finding AlexNet for time series classification," *Data Mining and Knowledge Discovery*, vol. 34, no. 6, pp. 1936–1962, sep 2020. [Online]. Available: https://doi.org/10.1007/s10618-020-00710-y

[7] Z. Wang, W. Yan, and T. Oates, "Time series classification from scratch with deep neural networks: A strong baseline," 2016. [Online]. Available: https://arxiv.org/abs/1611.06455

[8] A. Reinhardt and M. Bouchur, *On the Impact of the Sequence Length on Sequence-to-Sequence and Sequence-to-Point Learning for NILM*. New York, NY, USA: Association for Computing Machinery, 2020, p. 75–78. [Online]. Available: https://doi.org/10.1145/3427771.3427857