engineering proceedings

Proceeding Paper

Comparative Analysis of Residential Load Forecasting with Different Levels of Aggregation †

Ana Apolo Peñaloza *, Roberto Chouhy Leborgne and Alexandre Balbinot

Grupo de Modelagem e Análises de Sistemas de Potência and Laboratório de Instrumentação Eletrônica e Inteligência Artificial, Universidade Federal do Rio Grande do Sul UFRGS, Porto Alegre 90020-000, RS, Brazil; roberto.leborgne@ufrgs.br (R.C.L.); alexandre.balbinot@ufrgs.br (A.B.)

* Correspondence: ana.apolo@ufrgs.br; Tel.: +55-(51)-3308-3129

† Presented at the 8th International Conference on Time Series and Forecasting, Gran Canaria, Spain, 27–30 June 2022.

Abstract: Microgrids need a robust residential load forecasting. As a consequence, this highlights the problem of predicting electricity consumption in small amounts of households. The individual demand curve is volatile, and more difficult to forecast than the aggregated demand curve. For this reason, Mean Absolute Percentage Error (MAPE) varies in a large range (of 1% to 45%), depending on the number of consumers analyzed. Different levels of aggregation of household consumers that can be used in microgrids are analyzed; the load forecasting of the single consumer and aggregated consumers are compared. The forecasting methodology used is the most consolidated of Recurrent Neural Networks, i.e., LSTM. The dataset used contains 920 residential consumers belonging to the Commission for Energy Regulation (CER), a control group that is in the Irish Social Science Data Archive (ISSDA) repository. The result shows that the forecasting of groups of more than 20 aggregated consumers has a lower MAPE that individual forecasting. On the other hand, individual forecasting is better for groups with fewer than 10 consumers.

Keywords: load forecasting; LSTM; residential load forecasting; aggregation

check for updates

Citation: Peñaloza, A.A.; Leborgne, R.C.; Balbinot, A. Comparative Analysis of Residential Load Forecasting with Different Levels of Aggregation. *Eng. Proc.* **2022**, *18*, 29. https://doi.org/10.3390/engproc2022018029

Academic Editors: Ignacio Rojas, Hector Pomares, Olga Valenzuela, Fernando Rojas and Luis Javier Herrera

Published: 21 June 2022

1. Introduction

Load forecasting is essential to ensure a balance between demand and generation. Thus, utilities need highly accurate forecasts to maintain the security and stability of power supply [1]. At the same time, the complexity of the distribution network has continued to grow, which has created uncertainty in the grid, especially with the increase in microgeneration from renewable energy, and charging of electric vehicles [2].

Smart meters allow residents to monitor their consumption in real time. In addition to that, these meters provide large amounts of data from utilities [3]. These measurements allow for the enhanced measurement of consumption and energy control, allowing greater flexibility to the distribution network [4].

With this amount of data coming from smart meters, it becomes possible to perform a validation of demand forecasting at household or building levels. At these levels, consumption profiles are volatile [5,6]. Most load forecasting work is focused instead on large substations with tens of MW or transmission grids with tens of GW. Forecasting is assessed by the Mean Absolute Percentage Error (MAPE) metric which is generally below 2% for a substation at transmission level, while it can reach up to 30% for residential consumers [4]. Figure 1 shows aggregated demand curves for different amounts of consumers. This figure expresses that with 100 consumers, the demand curve is quite smoothed.

This work aims to analyze the demand forecasting in the context of microgrids. A dataset of 30 min demand of an individual residential consumer is used. This allows a comparison between demand forecasting of individual and aggregated consumers. This comparison is performed with different numbers of aggregated consumers (5, 10, 20, 30, 50 and 100). The selection of these consumers is random from a set of 256 residential consumers. The error metric used is the MAPE and the forecasting method is the Long Short-Term Memory (LSTM) recurrent neural network. LSTM is a consolidated methodology in load forecasting.

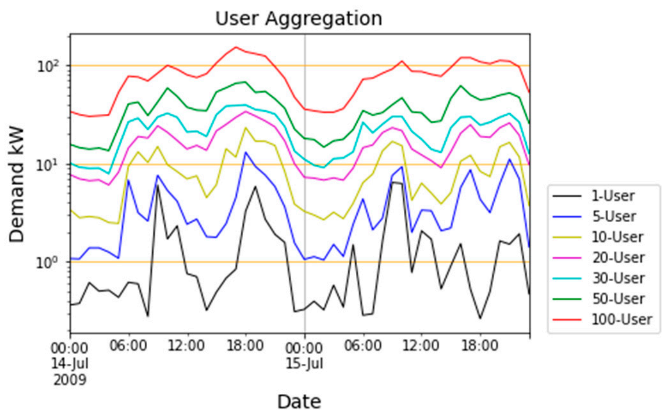

Figure 1. Different number of consumers.

2. Literature Review

The literature for demand forecasting is quite extensive, so it is possible to find statistical or machine learning methods for this purpose. In recent years, deep-learning methods are becoming popular for demand forecasting. Thus, recurrent neural networks, with their variant LSTM and GRU, are the most widely used methodologies for forecasting, as can be found in [2,7–13] for LSTM and [14,15] for GRU.

The forecasting literature is divided into forecasting of energy consumption and forecasting of load demand. Energy consumption forecasting has been used for few consumers, usually households or buildings, as found in [15–19]. On the other hand, load forecasting is used to forecast substations' demand, as found in [5,13,20–22]. These two approaches have completely different forecast errors using the same performance metrics. This is due to the difficulty of predicting power consumption.

Some authors that forecast energy consumption had aggregated individual energy consumption to obtain the building energy consumption [23].

There are few papers in the literature that compare load forecasting results and performance when using individual or aggregated consumers. Methodology performance for different levels of aggregation was evaluated in [7]. The authors found that with an aggregation of 5 consumers, the forecast errors are between 30 and 50%, while with 1000 aggregated consumers, the errors drop to less than 5%. However, the authors have not analyzed the forecasting performance for single users.

One of the first to analyze individual load forecasting is [13]. The individual's demand volatility makes the prediction difficult. Forecasting methodology is using an LSTM neural network. Demand forecasting is done for 69 consumers who belong to a set of 10,000 consumers from Australia for a period of three months. Other quantities of aggregated consumers are not analyzed. Results show that the non-aggregate forecast is better than the aggregate set forecast.

Comparisons of demand forecasting with different levels of aggregation are presented in [4], but the results are ambiguous and do not consider a non-aggregated methodology. Different aggregation levels using LSTM are shown in [24]. Their results suggest that demand forecasting of more than 200 aggregated residential consumers hardly decreases error, with the error curve becoming almost constant. However, the work of [24] only considers the increase in consumers and how this reflects on the aggregate predictor error. Non-aggregated consumers are not considered.

A methodology for demand forecasting of 200 consumers that are divided into groups of 50, 100, and 150 is presented in [25]. Forecasting is performed by LSTM and k-means clustering. The methodology without clustering proved to be better. They do not perform analyses with groups smaller than 50 consumers, which would be a typical microgrid context.

3. Demand Forecasting with RNN

Recurrent Neural Networks (RNNs) have become the most widely used methodology to perform residential demand forecasting [6,8,11,26–28]. The preference for RNNs is due to the fact that their models are sequence-based [6]. RNNs can process large data or text size series [29]. Therefore, RNNs are widely used in text translation and forecasting of time series data [30]. Energy demand is a time series, as observed in Figure 1. Figure 2 shows the typical architecture of RNNs and their unfolding in the earlier and later times.

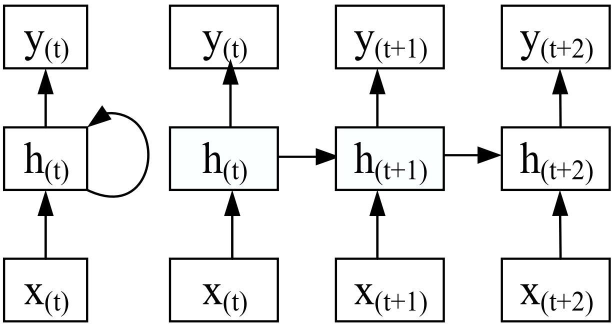

Figure 2. Recurrent neural network architecture. Adapted of [6].

There are two types of RNNs: Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU). LSTM and GRU are implemented to solve the problems of gradient bursting, when gradients approach infinity, or gradient fading, with gradients close to zero [30]. These problems are associated with successive multiplications of the weight matrices W [30].

LSTM is a solution to the long-time gradient fading and bursting problems through finite gradient control [30]. GRU, in turn, is a simplified variant of LSTM introduced by Cho [31], which performs the same gradient control using fewer gates.

Long Short-Term Memory

The LSTM is trained by the Backpropagation Through Time (BPTT) algorithm [32]. Figure 3 illustrates a time step with two LSTM cells, showing the internal connection of an LSTM cell. Figure 3 allows one to observe the updates of hidden state (h_t) and cell state (c_t) after a time step [30]. The key to the LSTM cell is cell state (c_t). The cell c_t moves from an earlier time step to the later time step and can be called a long-term memory term.

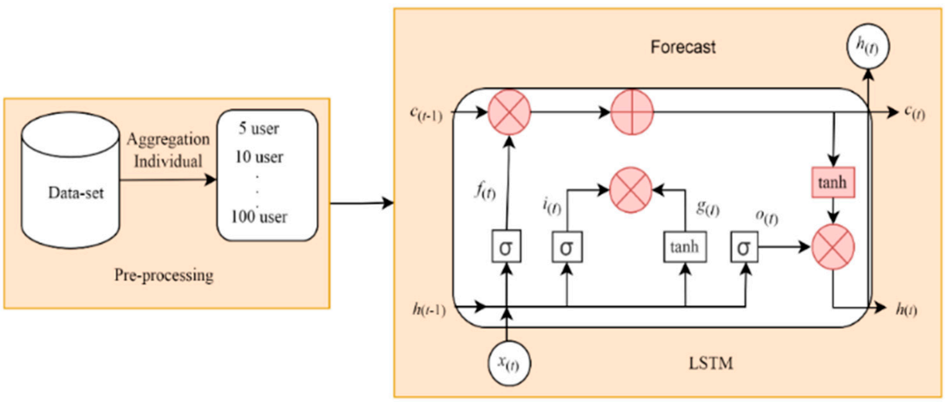

Figure 3. Individual and aggregated load forecasting.

The equation of c_t is shown in (1) and is divided into two parts. The first part is controlled by the forget gate (f_t). The f_t is in charge of defining the elements of the input

x_t to be kept or removed and the elements of the hidden state h_t. Its formula is shown in (2) [6,30]. The input gate (i_t) and input state (g_t) determine the input values that will be kept by the LSTM cell.

The input gate is shown in (3) and the input state (g_t) is shown in (4). So, g_t creates the values that can be added to the cell state while it decides which input values (input x) will be updated. The output gate (o_t) is also divided in two parts. In the first part, the values of the input (x) are placed, and the hidden state (h_t) is used in the output, and in the second part the same happens for the values of ct.

Finally, the sum of the two parts is transformed by the hyperbolic tangent function (tanh) to obtain values in the range from $[-1, 1]$ [6,29,30,33].

$$c_t = g_t \odot i_t + (c_{t-1} \odot f_t) \tag{1}$$

$$f_t = \sigma\left(W_{f \cdot x} x_t + W_{f \cdot h} h_{t-1} + b_f\right) \tag{2}$$

$$i_t = \sigma(W_{i \cdot x} x_t + W_{i \cdot h} h_{t-1} + b_i) \tag{3}$$

$$g_t = tanh\left(W_{g \cdot x} x_t + W_{g \cdot h} h_{t-1} + b_g\right) \tag{4}$$

$$o_t = \sigma(W_{o \cdot x} x_t + W_{o \cdot h} h_{t-1} + b_o) \tag{5}$$

$$h_t = tanh(c_t) \odot o_t \tag{6}$$

where: c_t is the cell state; h is the hidden state; f_t is the forget gate; i_t is the input gate; g_t is the input activation gate; o_t is the output gate. x is the input, W is the weight matrix, b is the bias in gate, σ is the sigmoid function, and *tanh* is the hyperbolic tangent.

4. Methodology for Comparison and Case Study

Figure 4 presents the methodology used to generate the forecasting for the comparison of the performance of the aggregated and individual forecast. The load forecasting is performed by the LSTM Recurrent Neural Network.

Figure 4 shows the methodology divided into two parts. The first is the pre-processing step, which contains normalization and missing data checking. The dataset used in this work belongs to the Smart Metering Electricity Consumer Behavior Trials project, carried out by the Commission for Energy Regulation (CER), Ireland's energy generation and distribution regulatory institution. The project that generated this database aimed to analyze the energy consumption per hour with different residential and industrial consumption tariffs by time of use (ToU). Its concern was to present the behavior of consumers in each of the different price ranges and their adaptation in relation to time of use. The project relied on the use of smart meters with a real-time digital panel, access to consumption via the internet, and detailed bimonthly consumption [34,35].

The dataset CER included 4225 residential consumers. However, this paper focused on the control group that included 657 consumers with measurements over a period of one and a half years, from 14 July 2009 to 31 December 2010. Data acquisition was carried out every 30 min. As the focus of this paper is influence of load aggregation on demand, forecasting groups with 5, 10, 20, 30, 50 and 100 consumers were randomized. The dataset CER does not contain missing data. The sampling period is one hour, and only one season is considered to perform the forecasting. Thus, 1600 samples per user are selected to perform the different forecasts.

In order to verify the performance of the forecast for aggregated and individual consumers, the following set of consumers {5, 10, 20, 30, 50 and 100} is performed. The selection of consumers for each set is randomized among the consumers. In order to verify the influence of consumers on each set, the random selection is repeated 10 times. Once the consumers in the set are selected, their demand is summed to create a single demand curve with 1600 time samples.

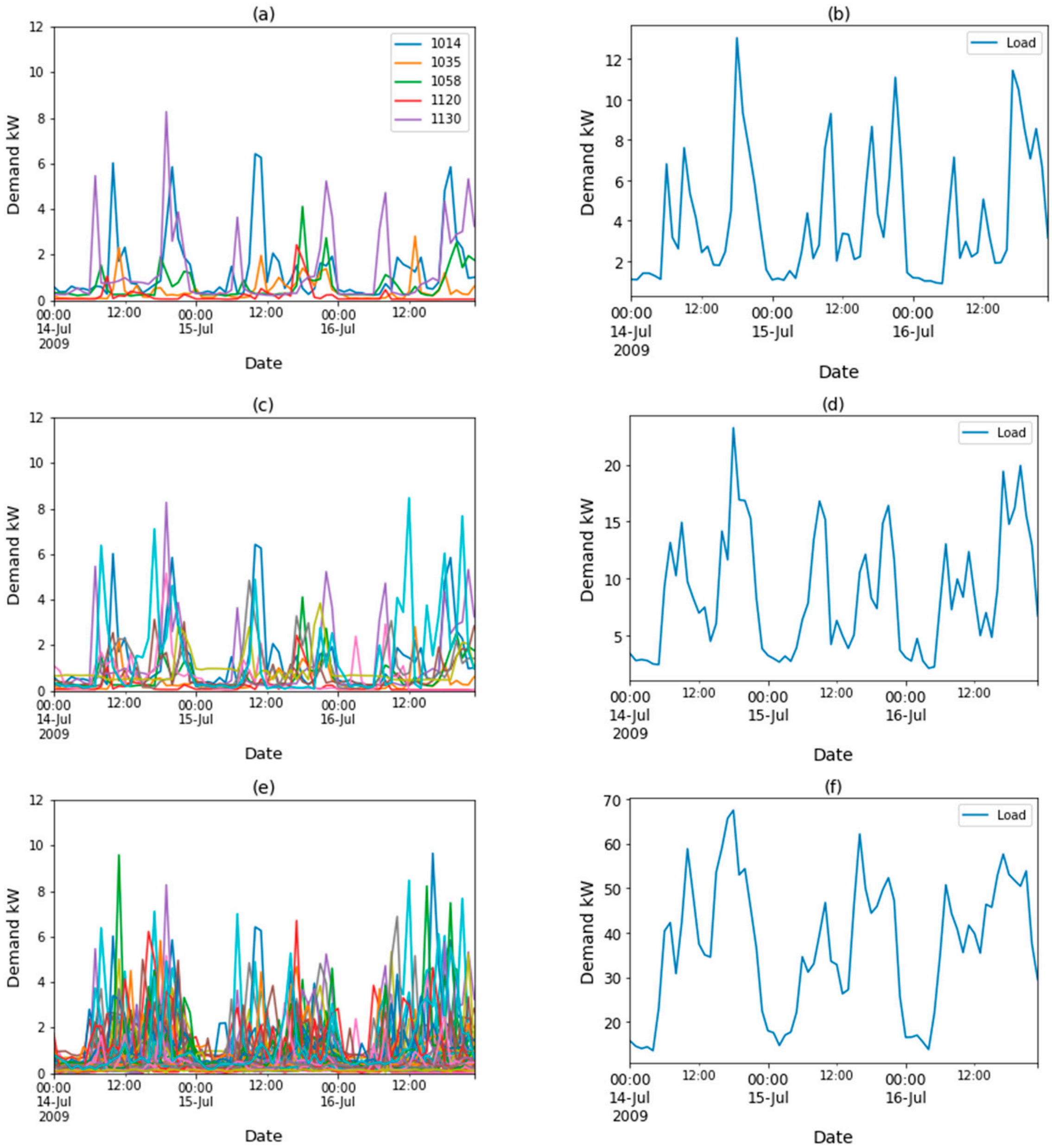

Figure 4. Individual and aggregated demand. (**a**) 5 individual consumers; (**b**) 5 aggregated consumers; (**c**) 10 individual consumers; (**d**) 10 aggregated consumers; (**e**) 50 individual consumers; (**f**) 50 aggregated consumers.

On the other hand, the curve called individual is composed of 1600 samples from each individual in the set (example: from a set with 20 consumers, the number of samples used is 32,000). Next, each individual consumer is predicted and then summed up; finally, the result is compared with the aggregated curve.

Figure 4a,c,e presents on the left the individual demand of 5, 10 and 50 consumers and on the right the aggregated demand of the consumers (see Figure 4b,d,f). Figure 4d,f shows that the demand waveform softens with the increase in the number of consumers, such that with 50 consumers the aggregated peaks (see Figure 4e) are smoother compared to the waveform of 5 consumers (see Figure 4b).

Table 1 shows the LSTM configuration for the aggregated and individual forecasting. The selection of these parameters is based on [8]. The forecast error metric is Mean Absolute Percentage Error (MAPE), which is the most widely used error measure according to the literature.

Table 1. LSTM hyper-parameters.

Hyper-Parameters	Aggregated	Individual
Number of Neurons	512	512
Dropout	0.2	0.2
Epochs	100	100
Optimizer	Adam	RMSProp
Activation Function	tanh	tanh

5. Results

To perform the forecasting, the CER database is divided into 70% for training and 30% for testing. The consumers were aggregated in 5, 10, 20, 30, 50 and 100 consumers to analyze the forecasting performance. The load demand curve of the previous 48 h of each group is used by the algorithm to predict the demand for next hour. In the case of aggregate forecasting, 48 time steps of aggregated demand are used, and the forecast is the aggregated demand of the following hour. In the case of individual forecasting, 48 time steps of individual demand are used for each consumer and the forecast is the individual demand of the next hour for each consumer.

Figure 5a–f show the forecast curve and the actual curve for aggregated and individual consumers, respectively. Figure 5a shows the forecast of individual consumers. Figure 5a–c highlight the difficulty of forecasting microgrids with few consumers (fewer than 10), where uncertainty is relatively more significant due to the volatility of individual consumers. Figure 5d–f show that after 20 aggregated consumers, the volatility decreases, smoothing the curves and facilitating the forecast. Finally, Figure 5f shows the best result because the curve is smoother and therefore less volatile than the curve in Figure 5a.

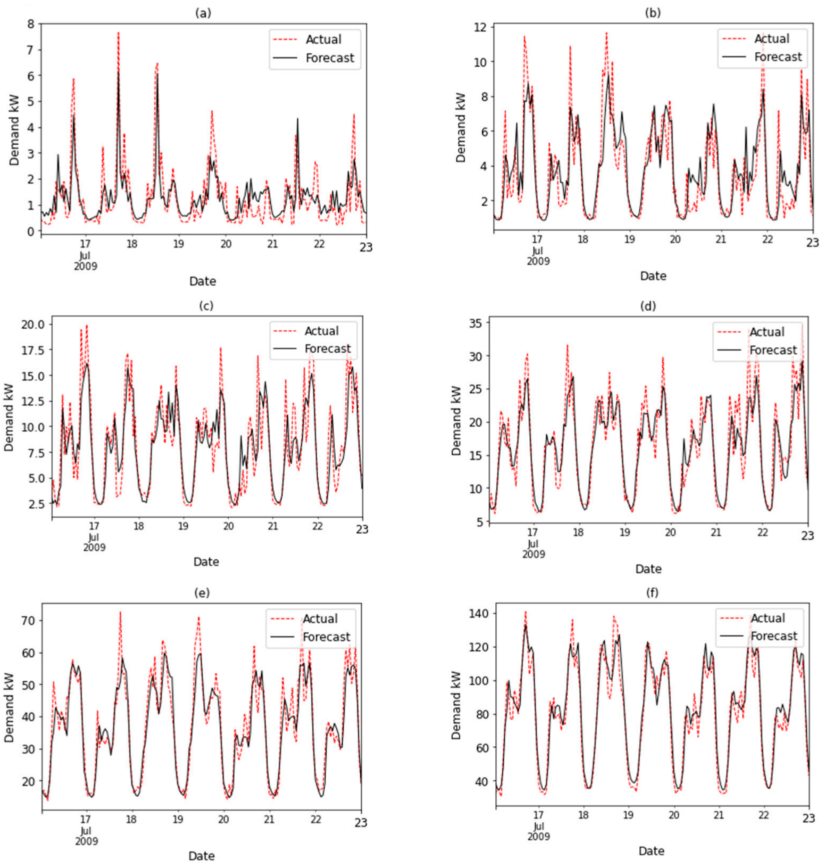

Figure 5. Actual vs. forecast demand curve. (**a**) Individual consumers; (**b**) 5 aggregated consumers; (**c**) 10 aggregated consumers; (**d**) 20 aggregated consumers; (**e**) 50 aggregated consumers; (**f**) 100 aggregated consumers.

Figures 6 and 7 show the average MAPE for the individual and aggregated consumers of the 10 forecast simulations for each group {5, 10, 20, 30, 50 and, 100} consumers. Figure 6 shows the MAPE for training whereas Figure 7 shows the MAPE for testing. Figures 6 and 7 show that MAPE decreases in training and testing when the number of aggregated consumers increases, because the aggregated demand curve is smoother with a larger aggregation of consumers.

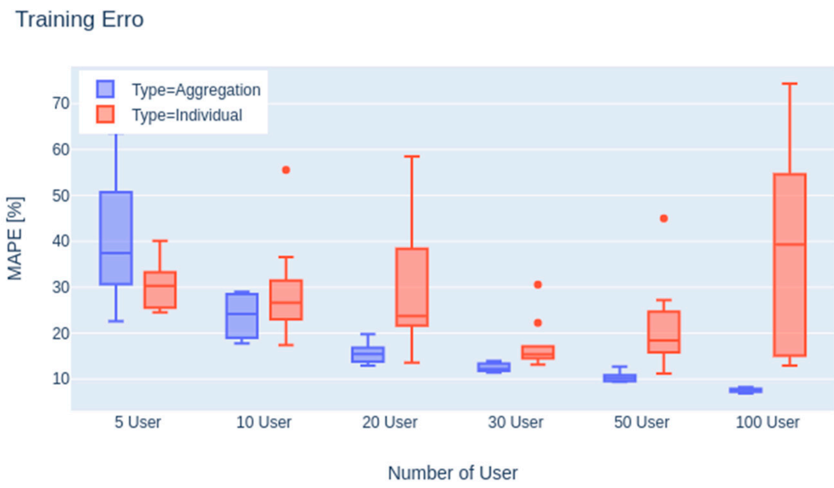

Figure 6. MAPE in the training of individual and aggregated consumers.

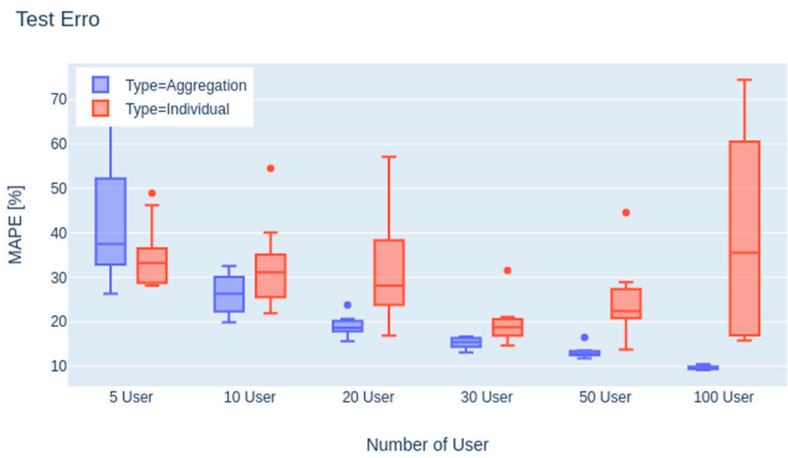

Figure 7. MAPE in the test of individual and aggregated consumers.

The MAPE of individual forecasting is comparable with aggregated forecasting when small number of consumers are considered, less than 10. Therefore, in small micro-grids the individual curve could be more useful than the aggregated one. The MAPE for individual forecasting is due to atypical consumers that tend to affect the individual forecast more than the aggregate. A way to improve the individual forecasting would be to work with consumers with similar consumption profiles. Thus, grouping would be by cluster and not random as was done in this work.

Another paper [36] also works with the CER database, uses a control group with 782 consumers, and his MAPE error is 6%. This work presents a training error of 6.88% when training uses 100 consumers. According to Figures 6 and 7, the error decreases with the increasing number of consumers. With 256 consumers, a 5% training error is found. The error decrease with the increase in consumers is due to the fact that with 256 consumers the demand curve is smoother than with 100 consumers. The error decrease is stable, after 100 consumers, with errors smaller than 3% for 782 consumers.

6. Conclusions

This paper presented a comparison of load forecasting considering an aggregated and an individual demand curve. The load forecasting was performed by the LSTM RNNs that was designed to work with a series of data such as energy demand curves. The analysis was performed with 5, 10, 20, 30, 50 and 100 consumers. The selection of consumers was randomized, and the experiment was repeated 10 times. The aggregated 100-consumer forecasting presented the lowest MAPE. The decrease in the MAPE was due to the smoothed demand curve of the aggregated consumers. The aggregated forecasting reduced the volatility of the electricity consumption. On the other hand, the individual forecast was more susceptible to atypical consumers which produce larger differences in the forecast for small groups (below 20 consumers). However, for small micro-grids, individual consumption is better than aggregate.

Author Contributions: Conceptualization, R.C.L. and A.A.P.; methodology, A.A.P.; validation, A.B.; writing—original draft preparation, A.A.P.; writing—review and editing, R.C.L. and A.B.; All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior-Brasil (CAPES)–Finance Code 001.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Accessed through the Irish Social Science Data Archive https://www.ucd.ie/issda/data/commissionforenergyregulationcer (accessed on 17 June 2022).

Acknowledgments: This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior-Brasil (CAPES)–Finance Code 001. The authors would like to thank the Irish Social Science Data Archive for making the database used in this paper available through "CER Smart Metering Project-Electricity Customer Behaviour Trial, 2009–2010" and also to ISSDA, as follows: "Accessed through the Irish Social Science Data Archive- https://www.ucd.ie/issda/data/commissionforenergyregulationcer (accessed on 17 June 2022).

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Deng, Z.; Wang, B.; Xu, Y.; Xu, T.; Liu, C.; Zhu, Z. Multi-scale convolutional neural network with time-cognition for multi-step short-Term load forecasting. *IEEE Access* **2019**, *7*, 88058–88071. [CrossRef]
2. Kong, W.; Dong, Z.Y.; Hill, D.J.; Luo, F.; Xu, Y. Short-term residential load forecasting based on resident behaviour learning. *IEEE Trans. Power Syst.* **2018**, *33*, 2017–2018. [CrossRef]
3. Kelly, J.; Knottenbelt, W. The UK-DALE dataset, domestic appliance-level electricity demand and whole-house demand from five UK homes. *Sci. Data* **2015**, *2*, 1–14. [CrossRef] [PubMed]
4. Sevlian, R.; Rajagopal, R. Short Term Electricity Load Forecasting on Varying Levels of Aggregation. 2017, 1–19. Available online: http://arxiv.org/abs/1404.0058 (accessed on 17 June 2022).
5. Elvers, A.; Vos, M.; Albayrak, S. Short-term probabilistic load forecasting at low aggregation levels using convolutional neural networks. In Proceedings of the 2019 IEEE Milan PowerTech, Milan, Italy, 23–27 June 2019. [CrossRef]
6. Kong, W.; Dong, Z.Y.; Jia, Y.; Hill, D.J.; Xu, Y.; Zhang, Y. Short-Term Residential Load Forecasting Based on LSTM Recurrent Neural Network. *IEEE Trans. Smart Grid* **2019**, *10*, 841–851. [CrossRef]
7. Gholizadeh, N.; Musilek, P. Federated learning with hyperparameter-based clustering for electrical load forecasting. *Internet Things* **2022**, *17*, 100470. [CrossRef]
8. Kong, W.; Dong, Z.Y.; Luo, F.; Meng, K.; Zhang, W.; Wang, F.; Zhao, X. Effect of automatic hyperparameter tuning for residential load forecasting via deep learning. In Proceedings of the 2017 Australasian Universities Power Engineering Conference (AUPEC), Melbourne, VIC, Australia, 19–22 November 2017. [CrossRef]
9. Amarasinghe, K.; Marino, D.L.; Manic, M. Deep neural networks for energy load forecasting. *IEEE Int. Symp. Ind. Electron.* **2017**, 1483–1488. [CrossRef]

10. Al Mamun, A.; Hoq, M.; Hossain, E.; Bayindir, R. A hybrid deep learning model with evolutionary algorithm for short-term load forecasting. In Proceedings of the 2019 8th International Conference on Renewable Energy Research and Applications (ICRERA), Brasov, Romania, 3–6 November 2019; pp. 886–891. [CrossRef]
11. Marino, D.L.; Amarasinghe, K.; Manic, M. Building energy load forecasting using Deep Neural Networks. In Proceedings of the IECON 2016-42nd Annual Conference of the IEEE Industrial Electronics Society, Florence, Italy, 23–26 October 2016; pp. 7046–7051. [CrossRef]
12. Somu, N.; MR, G.R.; Ramamritham, K. A hybrid model for building energy consumption forecasting using long short term memory networks. *Appl. Energy* **2020**, *261*, 114131. [CrossRef]
13. Jiao, R.; Zhang, T.; Jiang, Y.; He, H. Short-term non-residential load forecasting based on multiple sequences LSTM recurrent neural network. *IEEE Access* **2018**, *6*, 59438–59448. [CrossRef]
14. Hossen, T.; Nair, A.S.; Chinnathambi, R.A.; Ranganathan, P. Residential Load Forecasting Using Deep Neural Networks (DNN). In Proceedings of the 2018 North American Power Symposium (NAPS), Fargo, ND, USA, 9–11 September 2018; pp. 1–5. [CrossRef]
15. Jiang, Z.; Lin, R.; Yang, F. A hybrid machine learning model for electricity consumer categorization using smart meter data. *Energies* **2018**, *11*, 2235. [CrossRef]
16. Viegas, J.L.; Vieira, S.M.; Sousa, J.M.C. Fuzzy clustering and prediction of electricity demand based on household characteristics. In Proceedings of the 2015 Conference of the International Fuzzy Systems Association and the European Society for Fuzzy Logic and Technology, Asturias, Spain, 30 June–3 July 2015. [CrossRef]
17. Rahman, H.; Selvarasan, I.; Jahitha Begum, A. Short-term forecasting of total energy consumption for India-a black box based approach. *Energies* **2018**, *11*, 3442. [CrossRef]
18. Estebsari, A.; Rajabi, R. Single residential load forecasting using deep learning and image encoding techniques. *Electronics* **2020**, *9*, 68. [CrossRef]
19. Hsiao, Y.H. Household electricity demand forecast based on context information and user daily schedule analysis from meter data. *IEEE Trans. Ind. Inform.* **2015**, *11*, 33–43. [CrossRef]
20. Yan, X.; Abbes, D.; Francois, B. Uncertainty analysis for day ahead power reserve quantification in an urban microgrid including PV generators. *Renew. Energy* **2017**, *106*, 288–297. [CrossRef]
21. Beccali, M.; Cellura, M.; Lo Brano, V.; Marvuglia, A. Short-term prediction of household electricity consumption: Assessing weather sensitivity in a Mediterranean area. *Renew. Sustain. Energy Rev.* **2008**, *12*, 2040–2065. [CrossRef]
22. Moradzadeh, A.; Zakeri, S.; Shoaran, M.; Mohammadi-Ivatloo, B.; Mohammadi, F. Short-term load forecasting of microgrid via hybrid support vector regression and long short-term memory algorithms. *Sustainability* **2020**, *12*, 7076. [CrossRef]
23. Syed, D.; Abu-Rub, H.; Ghrayeb, A.; Refaat, S.S. Household-Level Energy Forecasting in Smart Buildings Using a Novel Hybrid Deep Learning Model. *IEEE Access* **2021**, *9*, 33498–33511. [CrossRef]
24. Peng, Y.; Wang, Y.; Lu, X.; Li, H.; Shi, D.; Wang, Z.; Li, J. Short-term Load Forecasting at Different Aggregation Levels with Predictability Analysis. In Proceedings of the 2019 IEEE Innovative Smart Grid Technologies-Asia (ISGT Asia), Chengdu, China, 21–24 May 2019; pp. 3385–3390. [CrossRef]
25. Hou, T.; Fang, R.; Tang, J.; Ge, G.; Yang, D.; Liu, J.; Zhang, W. A Novel Short-Term Residential Electric Load Forecasting Method Based on Adaptive Load Aggregation and Deep Learning Algorithms. *Energies* **2021**, *14*, 7820. [CrossRef]
26. Tian, C.; Ma, J.; Zhang, C.; Zhan, P. A deep neural network model for short-term load forecast based on long short-term memory network and convolutional neural network. *Energies* **2018**, *11*, 3493. [CrossRef]
27. Yudantaka, K.; Kim, J.S.; Song, H. Dual deep learning networks based load forecasting with partial real-time information and its application to system marginal price prediction. *Energies* **2019**, *13*, 148. [CrossRef]
28. Wang, Z.; Zhao, B.; Guo, H.; Tang, L.; Peng, Y. Deep ensemble learning model for short-term load forecasting within active learning framework. *Energies* **2019**, *12*, 3809. [CrossRef]
29. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016.
30. Aggarwal, C.C. *Neural Networks and Deep Learning*, 1st ed.; Springer: New York, NY, USA, 2018; ISBN 978-3-319-94462-3.
31. Chung, J.; Gulcehre, C.; Cho, K.; Bengio, Y. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. *arXiv* **2014**, arXiv:1412.3555.
32. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef] [PubMed]
33. Skansi, S. *Introduction to Deep Learning*; Springer: Berlin/Heidelberg, Germany, 2018; ISBN 9783319730035.
34. Exchange, T.; North, B.S. Electricity Smart Metering Customer Behaviour Trials (CBT) Findings Report DOCUMENT TYPE: REFERENCE: DATE. *Trial* **2011**, 1–58. Available online: https://www.cru.ie/wp-content/uploads/2011/07/cer11080ai.pdf (accessed on 17 June 2022).
35. CER (Commission for Energy Regulation). Available online: hhttps://www.ucd.ie/issda/data/commissionforenergyregulationcer (accessed on 17 June 2022).
36. Humeau, S.; Wijaya, T.K.; Vasirani, M.; Aberer, K. Electricity load forecasting for residential customers: Exploiting aggregation and correlation between households. In Proceedings of the 2013 Sustainable Internet and ICT for Sustainability (SustainIT), Palermo, Italy, 30–31 October 2013; pp. 1–6. [CrossRef]