2023 IEEE PES Grid Edge Technologies Conference & Exposition (Grid Edge) | 978-1-6654-6012-5/23/$31.00 ©2023 IEEE | DOI: 10.1109/GridEdge54130.2023.10102724

A scalable method for probabilistic short-term forecasting of individual households consumption in low voltage grids

Lola Botman, Jesus Lago, Thijs Becker, Oscar Mauricio Agudelo, Koen Vanthournout, and Bart De Moor

***Abstract*—Short-term individual household load forecasting is relevant for several applications and low voltage grid (LVG) stakeholders, e.g., for grid simulations, operation planning, congestion anticipation or advance payments. Electrical consumption at the household level is highly stochastic, point forecasting cannot capture this efficiently. To have insights about the uncertainty of the prediction, probabilistic methods should be developed. We propose a method to predict the half-hourly consumption of individual households one day ahead, based on a neural network, enhanced with empirical quantiles based on the point forecasts errors. The method is scalable thanks to its low computational requirements. Additionally, it requires only historical data and calendar features. Finally, the method is evaluated in a case study where it achieves state-of-the-art accuracy.**

***Index Terms*—Short-Term Probabilistic Load Forecasting, Smart Meter, Household Consumption, Low Voltage Grid.**

I. Introduction

Electric load forecasting is a key task in the grid management and has been researched for over a century [1]. Load forecasting can be classified according to the horizon, i.e., how long in the future is being forecast (minutes, weeks, months etc.) and according to the spatial granularity, i.e., the level of aggregation of the load (household, city, country, etc.). Different applications require different horizon and spatial aggregation [2]. Long-term load forecasting at the country or province level is used for capacity planning, and for energy policies. For suppliers, monthly predictions are interesting to determine customers' advance payments [3] for example.

A. Related work

Residential consumption forecasting, at the household level, started receiving more attention recently, with the roll-out of smart meters. It is useful for day-to-day system operations, especially with the introduction of more renewable energy sources, electrical vehicles, and heat pumps which increase the complexity and uncertainty of the profiles [4]. Additionally, short-term household forecasting is needed to make the network smarter [5], and potentially influence customer behavior towards congestion avoidance systems. The main difference between aggregated profiles (substation or citywide level) and household profiles lies in the variability of the data. Individual household profile present a higher stochasticity than aggregated profiles, due to the customer (almost) unpredictable behavior. The literature has mainly been focused on point forecasting methods. However, probabilistic forecasting methods are more appropriate and meaningful in this case: they output relevant information about the uncertainty of the predictions. In 2016, the first review paper on probabilistic forecasting methods was published [4], reviewing all forecasting horizons from very short-term to long-term and several aggregation levels. However, there is only one reference to individual household consumption forecasting [6]. They insist there is a strong need for more research on probabilistic forecasting methods, reproducible case studies and evaluation methods. The latest review on LV load forecasting [7], published in 2021, also states the lack of research at the household level and recommends moving towards probabilistic methods.

In residential load forecasting, the peaks are often of interest: they might lead to congestions on the LVG. A regular point forecast method, and associated loss function (or error metric), tends to ignore peaks as they generate a large error, even when correctly predicted in terms of height and duration but with a short time delay. This is called the double peak penalty effect. A solution is to consider different error metrics [6]. Probabilistic forecasting is an alternative to overcome this.

In this paper, we propose a probabilistic forecasting method for individual household electrical consumption one day ahead, with half-hourly resolution. The method takes as input the historical load and generates time and calendar features based on the timestamps. First, deterministic predictions are generated using a neural network, which are then enhanced with empirical quantiles based on the point forecast error, to produce a probabilistic forecast. The method is inspired from successful applications in weather and price forecasting methods [8], [9]. The approach is applied on an open dataset [10]. We show that our model achieves the same accuracy as more complex and computationally expensive methods.

L. Botman, O. M. Agudelo and B. De Moor *Fellow IEEE, SIAM, IFAC* are with the research group STADIUS Center for Dynamical Systems, Signal Processing, and Data Analytics within the department of Electrical Engineering (ESAT), KU Leuven, Belgium (Corresponding author: lola.botman@kuleuven.be).

T. Becker and K. Vanthournout are from the Unit Energietechnologie (ETE), VITO/EnergyVille, Genk, Belgium.

J. Lago is with Amazon, Amsterdam, The Netherlands. He contributed to this work as an outside activity and not as part of his role at Amazon.

This research received funding from the Flemish Government under the "Onderzoeksprogramma Artificiële Intelligentie (AI) Vlaanderen" program, from the KU Leuven Research Fund (projects C16/15/059, C3/19/053, C24/18/022, C3/20/117, C3I-21-00316) and from the European Research Council under the European Union's Horizon 2020 research and innovation programme (ERC Adv. Grant, grant agreement N^{o}885682).

B. Motivation & Contributions

This paper aims to contribute to the literature by proposing an accurate probabilistic method to forecast residential electricity consumption based on empirical quantiles. Probabilistic methods are necessary, due to the high stochasticity of individual household consumption profiles, to avoid phenomenons such as the double peak penalty and to gain insights on the uncertainty of the predictions. The paper contributes to the literature by proposing a method that:

- can generate short-term probabilistic forecasting of individual household's electrical consumption, starting from any point forecasting method;
- does not have strong data requirements such as weather data or household-specific attributes;
- is accurate, i.e., performs as well as the state-of-the-art;
- is scalable, thanks to the low data requirement and low computational complexity.

Finally, for reproducibility, the dataset used is publically available and the code to is made available to the scientific community [11].

II. Proposed Method

The approach described in this paper consists of three main modules, applied sequentially, on each profile individually: the preprocessing, the point forecasting and the empirical quantiles. Each module is detailed in the respective subsections II-A, II-B and II-C. The approach is illustrated with the flowchart in Fig. 1:

- Preprocessing each profile individually;
- Training the point forecasting model on the training set;
- Applying the point forecasting model on the validation set. Computing and grouping the errors on the point forecasting per time step. Computing empirical quantiles per time step;
- Applying the point forecasting model on the test set and adding the empirical quantiles to the point forecasting to generate probabilistic forecasts.

A. Preprocessing

In order to maximize reproducibility, and allow for better benchmarking, the data preprocessing is kept to its minimum. Each household profile is normalized between $[0, 1]$ using MinMaxScaling [12].

B. Point Forecasting

The next step is to produce the point forecast one day ahead. Many methods have been proposed for short-term household forecasting [13]. However it is difficult to state which method performs best out of all published deterministic methods, as they use different datasets, different input features, different pre-processing techniques, etc. One method may perform better on a dataset and worse on another, and there is no evidence that the complex methods outperform the more simple ones [13]. We choose to use a Feed Forward Neural Network (FFNN), since it is widely known, it has similar performances to- and lower computational requirement than other machine learning methods, and is one of the easiest to implement. Artificial neural networks are also recommended by Kuster et al. [13] in the specific setting of short-term household load forecasting with sub-hourly resolution.

Inputs and outputs: The amount of time steps in one day depends on the resolution of the dataset, e.g., half-hourly profiles have 48 time steps in one day, while quarter-hourly profiles have 96 time steps in one day. The number of time steps in one day is denoted by k. The outputs of the neural network are the k consumption predictions one day ahead:

$$\hat{y}_d = [\hat{y}_{d_1}, \hat{y}_{d_2}, ..., \hat{y}_{d_k}] \in \mathbb{R}^k, \tag{1}$$

where d indicates the forecast day and k depends on the dataset time resolution. The input of the model consists of a number of days of the historical consumption $y_i \in \mathbb{R}^k$: (i) the two days of consumption values prior to the forecast day y_{d-2} and y_{d-1}, (ii) one full day seven days prior to the forecast day y_{d-7} and (iii) one full day fourteen days prior to the forecast day y_{d-14}. Indeed, one can expect a typical customers' behavior (and his subsequent electrical consumption) on a Wednesday to be similar to the previous Monday and Tuesday, as well as the Wednesday of the two previous weeks. Additionally, calendar and time features of the day to be predicted are added: (iv) hod_d the hour of the day, (v) dom_d the day of the month, and (vi) moy_d the month of the year, are cyclically encoded using a sine and cosine function, and a (vii) binary indicator wk_d is used for weekday versus weekend. The sine and cosine encoding allows the algorithm to understand that the features are cyclical, e.g., 23h00 in the evening is close to 1h00 in the morning [14], [15]. This means that in total, the input consists of four days of historical consumption values and seven calendar/time features. The input vector is denoted:

$$\begin{aligned}&[y_{d-2}, y_{d-1}, y_{d-7}, y_{d-14}, \sin(\text{hod}_d),\\ &\cos(\text{hod}_d), \sin(\text{dom}_d), \cos(\text{dom}_d), \qquad \in \mathbb{R}^{4k+7}.\\ &\sin(\text{moy}_d), \cos(\text{moy}_d), \text{wk}_d]\end{aligned} \tag{2}$$

Architecture: The architecture consists of three layers: (1) an input layer, with the number of neurons n_x equal to the input vector size $4k + 7$, (2) one dense hidden layer, with 200 neurons and (3) an output layer, with the number of neurons n_y equal to the output vector size k. The dimensions of the input and output layers thus depends on the dataset time resolution. The ReLU (rectified linear unit) activation function is used after the input layer and after the hidden layer, and a linear activation function is used after the output layer.

Training, validation & testing process: Due to the high stochasticity of the consumption profiles and large differences between them, linked with individual customer behavior, we choose to train one model per household profile. Each profile is split in time sequentially: the first 60% of the time steps are used for training, the next 20% for validation and the last 20% for testing. The Adam optimizer with default parameters from Keras [16] is used for the training, i.e., learning rate$= 0.001$, beta $1= 0.9$, beta $2= 0.999$, epsilon$= 1e-07$. Early stopping is used during the training, in order to avoid overfitting.

978-1-6654-6012-5/23/$31.00 ©2023 IEEE

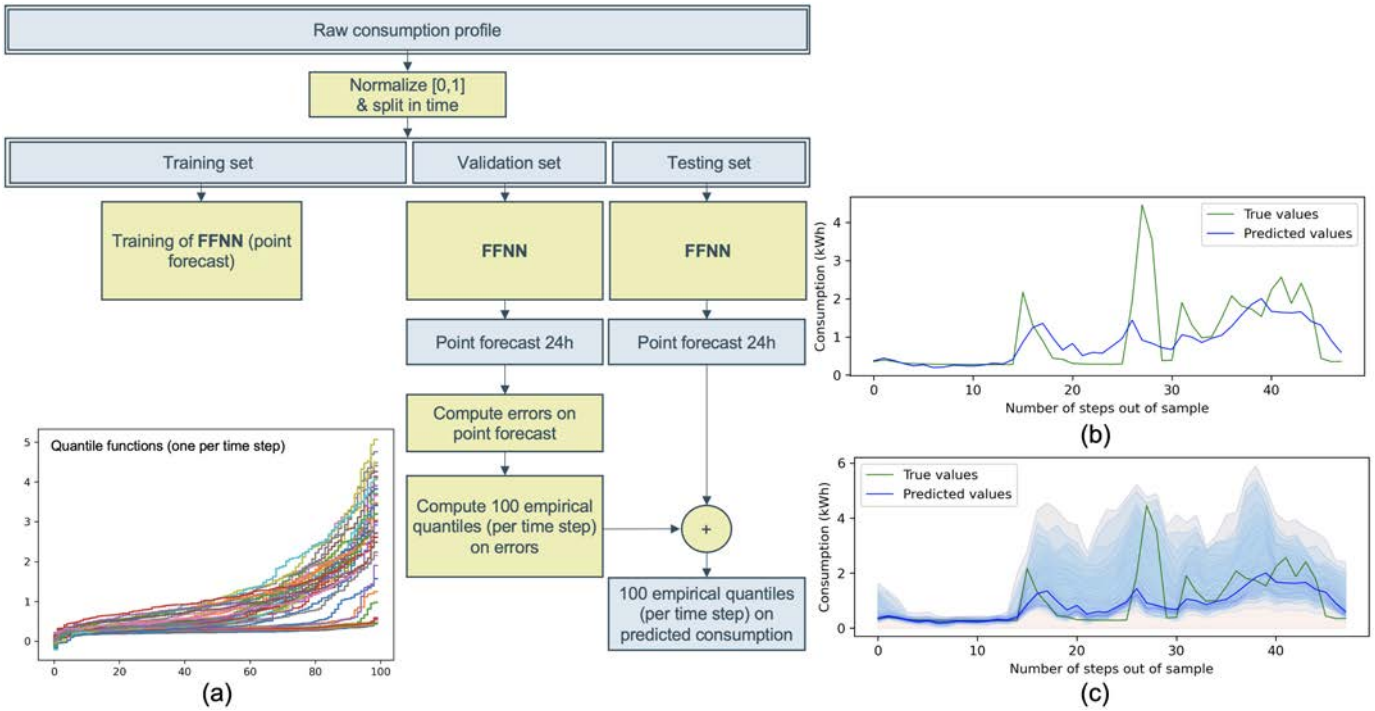

Fig. 1. Flowchart of the proposed method, applied individually on each profile. Each consumption profile is first preprocessed, then split in training, validation and testing set. The Feedforward Neural Network (FFNN) is trained on the training set. The FFNN is then applied on the validation set. The empirical quantiles are computed on the errors of the validation set. Finally, the FFNN is applied on the test set and the empirical quantiles are added to the point forecast in order to obtain a probabilistic estimate of the future consumption. (a) 100 quantiles are computed per time step, there is one quantile functions per time step in the day. (b) Point forecasting results one day ahead. (c) Point forecasting with empirical quantiles.

C. Empirical Quantiles

The trained point forecasting model is applied on the validation set. It generates one day ahead forecasts $\hat{y} = [\hat{y}_1, \hat{y}_2, ..., \hat{y}_k]$ for the m days in the validation set. The errors e (in dark yellow to khaki green on Figure 2) between the predictions $\hat{y}$ (in blue) and the true consumption y (in green) are computed for the m days. There are m error vectors of size k:

$$e = y - \hat{y} = [e_1, e_2, ..., e_k] \in \mathbb{R}^k. \tag{3}$$

The errors are then grouped per time steps, i.e., the first group g_1 consists of all prediction errors made at the first time step (00h00), the second group g_2 consists of all errors made at the second time step (00h30), and so on. In total, there are k sets of m errors, one for each time step of the day:

$$\begin{aligned} g_1 &= [e_1, e'_1, ..., e_1^{(m)}] \in \mathbb{R}^m \\ g_2 &= [e_2, e'_2, ..., e_2^{(m)}] \in \mathbb{R}^m \\ &\vdots \\ g_k &= [e_k, e'_k, ..., e_k^{(m)}] \in \mathbb{R}^m. \end{aligned} \tag{4}$$

Based on each set of error g_i, 100 empirical quantiles are computed. The 100 quantiles represent the error distribution of the point forecasting at each time of the day. In total, there are 100 quantiles for each of the k time steps, there are k quantile functions. An example is shown in Fig. 1(a).

Finally, the point forecasting model is applied to the test set, the output is one day ahead deterministic predictions, as illustrated on Fig. 1(b). The previously computed quantiles are added to these point predictions, this generates uncertainty bounds around the predictions, as shown in Fig. 1(c).

III. Case Study

A. Dataset

The dataset is provided by the Commission for Energy Regulation (CER) of Ireland and is available on the Irish Social Science Data Archive (ISSDA) website [10]. The data was collected in the context of the Smart Metering Project - Electricity Customer Behaviour Trial in 2009 and 2010. It consists of the half-hourly electrical consumption (in kWh) of more than 5,000 Irish homes and businesses from July 2009 until December 2010. In order to make the benchmark as consistent as possible, we use the exact same preprocessing pipeline as in [17], available on GitHub [18]. As Arpogaus et al. [17] state, they first remove non-residential profiles, since the interest and challenge of the model lies specifically in the stochastic behavior of the residential customers. They also remove all incomplete records. A subset of the remaining profile is selected randomly (10%, equivalent to 363 customers). As mentioned in II-A, the profiles are also normalized between $[0, 1]$ individually.

B. Evaluation

The conventional point forecasting performance metrics such as Mean Absolute Error (MAE) or Mean Square Error (MSE) can easily be computed for the method proposed,

978-1-6654-6012-5/23/$31.00 ©2023 IEEE

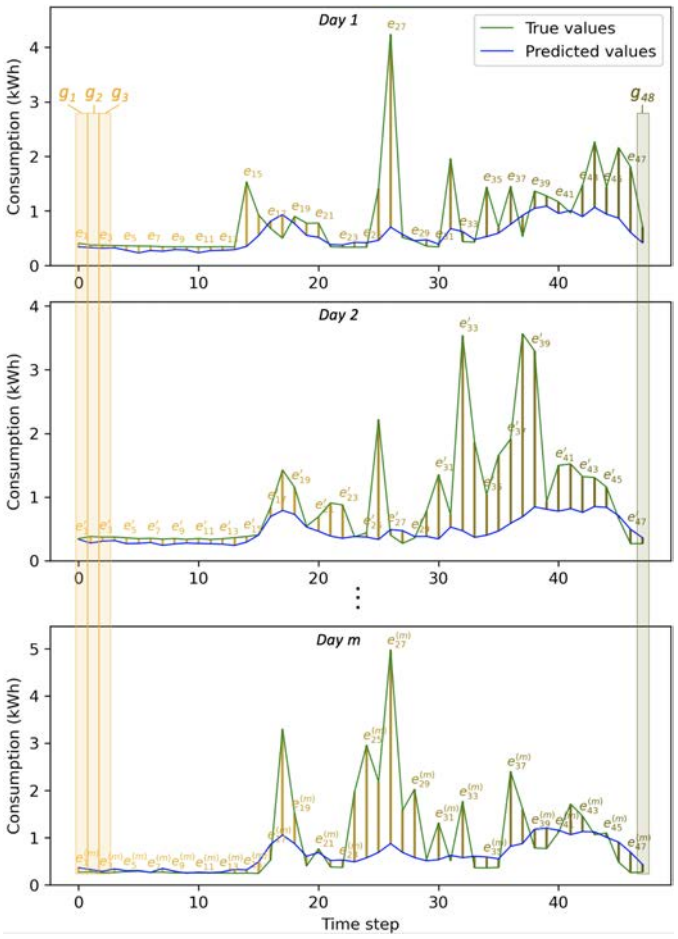

Fig. 2. The errors are computed between the predicted consumption (in blue) and the true consumption (in green). The errors are then grouped per time step to compute the empirical quantiles.

as the approach consists in point forecasting in a first step, enhanced with quantiles in a second step. Nevertheless, they don't take into account the reliability, nor the sharpness of the probabilistic forecast. A proper scoring rule, such as the Continuous Ranked Probability Score (CRPS) is recommended [9], [19], [20]. More details are provided in the supplementary materials [11]. We thus report the MAE, MSE and CRPS. Each metric is computed per profile and then averaged over all profiles.

C. Benchmark models

The method proposed in this paper is compared with two other methods: a probabilistic baseline and a complex state-of-the-art approach proposed in the recent literature using Bernstein-Polynomial Normalizing Flows (BNF) [17]. The authors claim that their method outperforms a simple Gaussian Model, a Gaussian Mixture Model and a Quantile Regression. In the point forecasting literature, the naive baseline is often a persistence model: the true values of the previous days are considered as the predictions for the next day:

$$\hat{y}_d = y_{d-1}, \tag{5}$$

assuming that without mathematical formulation or complex modeling, the best hypothesis that can be made is that the consumption of the current day will be similar to the consumption of the previous day. We propose a probabilistic baseline, grounded on the same assumption, using the persistence model in the prediction step and then applying the empirical quantiles' step as proposed in Section II-C.

The benchmark model uses BNF to transform a complex distribution to a more simple distribution. A neural network is trained to predict the parameters of the flows. These parameters yield the marginal Conditional Probability Density for all the time steps, one day ahead. Seven days of data prior to the predicted day are used as input and there are five additional exogenous inputs, i.e., the sine and cosine encoding of the day of the year, and the day of the week, as well as a binary indicator for the holidays. The reader is referred to the full paper for details [17], [18], [21]. As this approach outputs distribution functions per time steps, we computed quantiles from these distributions and then applied our own implementation of the discretized CRPS on the quantiles. Note that the BNF method results are slightly different from the results presented in [17], [21] as they have been re-run locally, with our own (not normalized) CRPS implementation, in order to allow for exact comparison with the other methods.

IV. Results & Discussion

The performance metric results are presented in Table I. The approach described in this paper is compared with a probabilistic baseline and a state-of-the-art technique, both detailed in Section III-C.

The point forecasting of the different methods can be compared with the MSE and MAE results. In both cases, smaller values mean better predictions. It is clear that the proposed method outperforms both benchmarks. However, this is expected, as the first method is a baseline making a strong assumption, and the second one is designed to predict a conditional probability directly, not a point forecast.

Regarding CRPS, there is an approximate 25% decrease in CRPS between the baseline and the two other methods, which means 25% improvement. The difference between both methods is rather small, but still shows a $1,75\%$ improvement.

The execution times of each method are summarized in Table II. It is relevant to take execution times into account as these algorithms are developed for applications with thousands of profiles, at the country level for example. It is crucial that the methods should be accurate and efficient. Each method is run on a machine with 9th Generation "Coffee Lake" 2.6 GHz 6-Core Intel Core i7 mobile processor (I7-9750H), MacOS operating system, Python version 3.8.3. The baseline takes less than two minutes to make the predictions for the 363 profiles and requires no training.

The BNF method requires 38 minutes for training and 7 minutes for the predictions of all 363 profiles. The BNF implementation [18] is provided by the authors and run locally. The prediction process could be parallelized, however, the training cannot as one model is trained on all profiles.

978-1-6654-6012-5/23/$31.00 ©2023 IEEE

	MSE	MAE	CRPS	CRPS' relative improvement to baseline
Baseline	0.6008	0.3944	0.3252	/
BNF [18]	1.0800	0.3335	0.2461	24.32%
Proposed method	0.3969	0.3080	0.2418	25.64%

TABLE I. Performance metrics of each method applied on the CER dataset.

	Training	Testing	Total
Baseline	/	1 min 15 s	1 min 15 s
BNF [18]	38 min	7 min	45 min
Proposed method	14 min	1 min	15 min

TABLE II. Execution time of the training and testing of each method.

The proposed method takes at most 15 minutes to run the whole pipeline, i.e., preprocessing, train-test split, model training, predictions and saving all results. More specifically, between 11 and 14 minutes are needed for training (it varies slightly between runs, the longest one is reported in Table II) and 1 minute for predicting all the 363 profiles. Each profile model is trained consecutively, and the predictions are made sequentially profile per profile as well. The run-time could be reduced by parallelizing the process.

It is clear that the baseline is faster, as there is no training and simple mathematics are applied. However, for similar performance metric achieved in terms of CRPS, the method proposed in this paper is three times faster than the BNF approach, and thus scalable for country wide applications.

V. Conclusion

This paper presents a fast and accurate approach for probabilistic short-term forecasting of household consumption based on historical data and calendar features inferred from the timestamps. The approach contributes to the limited literature in terms of probabilistic methods and lower aggregation level load. The method is compared to a baseline and to a state-of-the-art method. It presents strong advantages: (i) it has low data requirement, nor weather, nor households-specific attributes such as appliance information, or number of habitants are needed; (ii) it is fast and scalable thanks to the low computational complexity and low data requirements; (iii) it is accurate as it achieves state-of-the-art performances.

In future work, the method could be improved with other point forecasting methods in the first step, although it might impact scalability if more complex models are used. The error measure proposed in [6] could be used in the point forecasting model to reduce the double penalty effect. The model presented in this paper could also be applied on new datasets in order to assess the generalization of the method.

Acknowledgment

The work of this paper has been carried out as part of a collaboration between VITO/EnergyVille and the research group STADIUS from KU Leuven.

References

[1] T. Hong and S. Fan, "Probabilistic electric load forecasting: A tutorial review," *International Journal of Forecasting*, vol. 32, no. 3, pp. 914–938, 2016.

[2] P. E. McSharry, S. Bouwman, and G. Bloemhof, "Probabilistic forecasts of the magnitude and timing of peak electricity demand," *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 1166–1172, 2005.

[3] L. Botman, J. Soenen, K. Theodorakos, A. Yurtman, J. Bekker, K. Vanthournout, H. Blockeel, B. De Moor, and J. Lago, "A scalable ensemble approach to forecast the electricity consumption of households," *IEEE Transactions on Smart Grid*, pp. 1–12, 2022.

[4] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu, and Y. Zhang, "Short-Term Residential Load Forecasting Based on LSTM Recurrent Neural Network," *IEEE Transactions on Smart Grid*, vol. 10, no. 1, pp. 841–851, 2019.

[5] A. F. Ebrahim and O. Mohammed, "Household load forecasting based on a pre-processing non-intrusive load monitoring techniques," *IEEE Green Technologies Conference*, vol. 2018-April, pp. 107–114, 2018.

[6] S. Haben, J. Ward, D. Vukadinovic Greetham, C. Singleton, and P. Grindrod, "A new error measure for forecasts of household-level, high resolution electrical energy consumption," *International Journal of Forecasting*, vol. 30, no. 2, pp. 246–256, 2014.

[7] S. Haben, S. Arora, G. Giasemidis, M. Voss, and D. Vukadinović Greetham, "Review of low voltage load forecasting: Methods, applications, and recommendations," *Applied Energy*, vol. 304, pp. 1–33, 2021.

[8] A. Gerossier, R. Girard, A. Bocquet, and G. Kariniotakis, "Robust day-ahead forecasting of household electricity demand and operational challenges," *Energies*, vol. 11, no. 12, 2018.

[9] J. Nowotarski and R. Weron, "Recent advances in electricity price forecasting: A review of probabilistic forecasting," *Renewable and Sustainable Energy Reviews*, vol. 81, no. June 2017, pp. 1548–1568, 2018.

[10] Commission for Energy Regulation (CER), "CER Smart Metering Project - Electrical Customer Behaviour Trial," 2012. [Online]. Available: https://www.ucd.ie/issda/data/commissionforenergyregulationcer/

[11] L. Botman and J. Lago, "Python Implementation: Short-term probabilistic forecasting of individual households' electrical consumption," 2022. [Online]. Available: https://gitlab.esat.kuleuven.be/Lola.Botman/psthl-forecasting

[12] F. Pedregosa, G. Varoquaux, L. Buitinck, G. Louppe, O. Grisel, and A. Mueller, "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, no. 1, pp. 2825–2830, 2011.

[13] C. Kuster, Y. Rezgui, and M. Mourshed, "Electrical load forecasting models: A critical systematic review," *Sustainable Cities and Society*, vol. 35, no. June, pp. 257–270, 2017.

[14] G. Petneházi, "Recurrent neural networks for time-series forecasting," *arXiv*, 2019.

[15] A. van Wyk, "Encoding cyclical features for deep learning." [Online]. Available: https://www.avanwyk.com/encoding-cyclical-features-for-deep-learning/

[16] TensorFlow, "Keras Adam Optimizer." [Online]. Available: https://www.tensorflow.org/api_docs/python/tf/keras/optimizers/Adam

[17] M. Arpogaus, M. Voß, B. Sick, M. Nigge-Uricher, and O. Dürr, "Probabilistic Short-Term Low-Voltage Load Forecasting using Bernstein-Polynomial Normalizing Flows," *ICML 2021 Workshop on Tackling Climate Change with Machine Learning*, pp. 1–6, 2021.

[18] M. Arpogaus, "Short-Term Probabilistic Load Forecasting using Conditioned Bernstein-Polynomial Normalizing Flows (STPLF-BNF)," 2021. [Online]. Available: https://github.com/MArpogaus/stplf-bnf

[19] T. Gneiting and A. E. Raftery, "Strictly proper scoring rules, prediction, and estimation," *Journal of the American Statistical Association*, vol. 102, no. 477, pp. 359–378, 2007.

[20] T. Gneiting and M. Katzfuss, "Probabilistic forecasting," *Annual Review of Statistics and Its Application*, vol. 1, pp. 125–151, 2014.

[21] M. Arpogaus, M. Voss, B. Sick, M. Nigge-Uricher, and O. Dürr, "Short-Term Density Forecasting of Low-Voltage Load using Bernstein-Polynomial Normalizing Flows," *arXiv*, pp. 1–10, 2022.