Residential Energy Consumption Prediction Using Inter-Household Energy Data and Socioeconomic Information

Pascal A. Schirmer
Comm. & Intelligent Systems Group
School of Eng. & Computer Science
University of Hertfordshire
Hatfield AL10 9AB, UK
p.schirmer@herts.ac.uk

Christian Geiger
Institute for Machine Tools
Faculty of Mechanical Engineering
Technical University of Munich
85748, Garching, Germany
christian.geiger@tum.de

Iosif Mporas
Comm. & Intelligent Systems Group
School of Eng. & Computer Science
University of Hertfordshire
Hatfield AL10 9AB, UK
p.schirmer@herts.ac.uk

***Abstract*—Previous studies have shown that residential energy consumption prediction accuracy can be improved when households energy data are fused with residents' socioeconomic information. In this article we propose an architecture for the prediction of residential energy consumption using past energy consumption from other/neighboring households in combination with socioeconomic information of the corresponding residents. The architecture is based on a Long Short Term Memory model and was evaluated using a large-scale dataset monitoring households of London. The proposed approach significantly improves the accuracy of the energy consumption predictor reducing the mean absolute error up to 25.2% with prediction error rate equal to 5.4%.**

***Index Terms*—Energy Consumption Prediction, Load Forecasting, Socioeconomic Features.**

I. Introduction

Over the last years the world energy demand has increased due to the population growth and economic development [1]. It is expected that the energy demands will further increase in the next two decades [2]. Urban development heavily relies on electric energy systems, which require energy consumption at the same time with energy generation [3]. Energy generation below the actual needs will result in blackouts while generation of electric energy above the actual needs makes the electric energy system financially unsustainable [4], [5]. Thus, short-time prediction of the energy needs is crucial for electric energy systems and given that the 27% of the global energy consumption corresponds to the residential sector [6], prediction of residential energy consumption is essential.

Residential energy consumption prediction is a time series prediction problem [7], [8]. Data can be collected from smart-meters, one per customer, which in most cases is equivalent to per household data. The energy consumption data, which are usually collected with sampling period from 1 second to 1 hour [9], [10], in most cases consist of active power samples and less often reactive power, load angle and current harmonics [11], [12]. For the residential sector the interest is mostly in short-term prediction (sub-hourly, hourly) for the prevention of blackouts, but also in long-term prediction (monthly, yearly) for national planning and investment [13], [14]. The prediction of energy consumption is a difficult task because except the periodic patterns (e.g. daily and weekly routine) irregular components appear in the energy consumption signal as well [15].

Several approaches have been proposed in the literature for residential energy consumption prediction, most of which are based on machine learning algorithms for regression, e.g. Linear Regression (LR) [15], Support Vector Regression (SVR) [16], Decision Trees (DTs) [15] and deep learning methods like Convolutional Neural Networks ($CNNs$) [2], Recurrent Neural Networks ($RNNs$) [15] and Long Short Term Memory ($LSTM$) [2]. In most proposed approaches the short or long term history of energy consumption is used to predict future energy load demands.

Except energy demand forecasting approaches based on energy data, use of non-energy data that can affect load demand has been evaluated. Such data are the weather conditions [17], calendar information [15] and socioeconomic factors [4], [18]. The use of non-energy consumption data was shown to generally improve the electricity demand forecast. In this paper we present a regression based method for short-term prediction of residential electricity load demand using socioeconomic information of the customers and energy consumption information from neighboring households. In detail a unified model of several houses is used to integrate socioeconomic information (e.g. age and number of the residents), which otherwise could not be used as it is constant for one single household.

The remainder of this paper is organized as follows. In Section II the proposed architecture for energy consumption prediction is presented. In Sections III and IV the experimental setup and evaluation results are described, respectively. In Section V we conclude this work.

II. Inter-Household Energy Consumption Prediction

Residential energy consumption relies on the characteristics and needs of each consumer (families, house-mates or individuals). The prediction of the energy consumption of a target house can be formalized as:

$$\hat{h}_{target}(t+w) = f(h_{target}(t_0:t)) \tag{1}$$

where $[t_0:t]$ is the previous time window used to predict the energy consumption at $(t+w)$, $h_{target}(t_0:t) \in \mathbb{R}^{(t-t_0+1)}$ is the energy consumption of the previous time window, $\hat{h}_{target}(t+w) \in \mathbb{R}^1$ its w-step ahead prediction and $f()$ an arbitrary regression model (e.g. LR, SVR, $LSTM$, etc.).

We assume that across different households there are common energy consumption trends and motifs or even inter-dependencies due to potential socioeconomic relationships, which potentially have time lags between them or appear simultaneously [19]. This motivates us to use the energy consumption history of $N-1$ other households as an additional input channel of information to enhance the prediction of energy load demand of the target house. In that case the formalization of the problem is expressed as:

$$\begin{gathered}\hat{h}_{target}(t+w) = f(h_{target}(t_0:t), h_i(t_0:t)) \\ with\ 1 \leq i < (N-1)\end{gathered} \tag{2}$$

where $h_i(t_0:t)$ is the energy consumption signal in the time window $[t_0:t]$ for the i^{th} neighboring household.

Given that prediction models are trained from several households' data, the use of socioeconomic information of the consumers of the target house would result to load demand forecasting models adapted to the specifications of each socioeconomic group of consumers. Socioeconomic dependent models are expected to predict more precisely the energy consumption behavior of a house [18] and the prediction can be formalized as:

$$\begin{gathered}\hat{h}_{target}(t+w) = f(h_{target}(t_0:t), h_i(t_0:t), s_{target}) \\ with\ 1 \leq i < (N-1)\end{gathered} \tag{3}$$

where $s_{target} \in \mathbb{R}^K$ is the K-dimensional socioeconomic information of the target house. The proposed architecture using inter-household energy data and socioeconomic information for prediction of energy consumption using a regression model is shown in Fig. 1.

As can be seen in Fig. 1 the architecture consists of a data fusion stage, fusing the energy consumption signal of the different houses and the socioeconomic information to one feature vector, and a regression stage for predicting the energy consumption. $LSTM$ was chosen as it was proven to outperform LR, DTs, $DNNs$ and $RNNs$ on a similar architecture [20].

In detail the architecture in Fig. 1 consists of an input layer of $M^{(0)} = (t-t_0+1)$ nodes, a number of L $LSTM$ layers with $M^{(L)}$ nodes, a feedforward output layer consisting of $M^{(L+1)}$ nodes and an input sequence of previous energy values $h(t_0:t) = [h(t), h(t-1), ..., h(t-t_0+1)]^T \in \mathbb{R}^{(M^0)}$. In detail each $LSTM$ layer l maps its input sequence $h^{(l-1)}(t) \in \mathbb{R}^{N^{(l-1)}}$ to an output sequence $h^{(l)}(t) \in \mathbb{R}^{N^{(l)}}$, $1 \leq l \leq L$, while the final layer uses a feedforward mapping thus the output of the baseline $LSTM$ is given as in [21]

$$\hat{h}_{target}(t+w) = \sigma^{(L+1)}(W^{(L+1)} h^{(L)}(t_0:t) + b^{(L+1)}) \tag{4}$$

where $\sigma^{(L+1)}(\cdot)$ is the element-wise non-linear activation function of the output layer, W is the weighting matrix and b is the bias vector [21].

Furthermore considering the inter-household information the input signal becomes a $(M^0 \times N)$ dimensional matrix $H(t_0:t) = [H(t), H(t-1), ..., H(t-t_0+1)]^T \in \mathbb{R}^{(M^0 \times N)}$ thus the output of the $LSTM$ is given as:

$$\hat{h}_{target}(t+w) = \sigma^{(L+1)}(\boldsymbol{W}^{(L+1)} H^L(t_0:t) + \boldsymbol{b}^{(L+1)}) \tag{5}$$

where $W_i^{(L+1)}$ and $b_i^{(L+1)}$ is the weighting matrix and the bias vector for the i^{th} household in H.

Moreover considering the socioeconomic information of the target house s_{target} the input signal becomes a $(M^0 \times N + K)$ dimensional matrix extending the energy inputs $H(t_0:t)$ to $H = \{H(t_0:t), s_{target}\}$.

III. Experimental Setup

The energy prediction architecture utilizing inter-household energy data and socioeconomic information presented in Section II was evaluated using the dataset, features and regression algorithm presented below

A. Dataset

To evaluate the proposed architecture the publicly available dataset "Smart Meters in London" ($SMinL$) [22] was used. The dataset consists of a refactored version of the London data store containing the energy readings from 5567 households in London measured between November 2011 and February 2014 at a sampling rate of 1 sample per 30 min [23]. In addition, the dataset contains annotations regarding the socioeconomic background of the residents. Specifically a set of 17 social groups, called $ACRON$ groups, was formed out of the total number of households. Each of these groups is characterized by a set of 825 socioeconomic features clustered in 84 subcategories and 15 main categories. For our evaluation the year 2013 (1^{st} of January 2013 – 31^{st} of December 2013) was used, since year 2012 has several gaps in the measurements, using 50 households per $ACRON$ group. Furthermore we excluded $ACRON$-$\{B, K, M\}$ as they have missing samples for the evaluated duration. The average properties for the 50 households in each $ACRON$ group for the year 2013 are displayed in Table I.

As can be seen in Table I the $ACRON$ groups cover a wide spectrum of different household characteristics with significant variations in average energy consumption, different number of residents, average age of residents, etc., thus making it

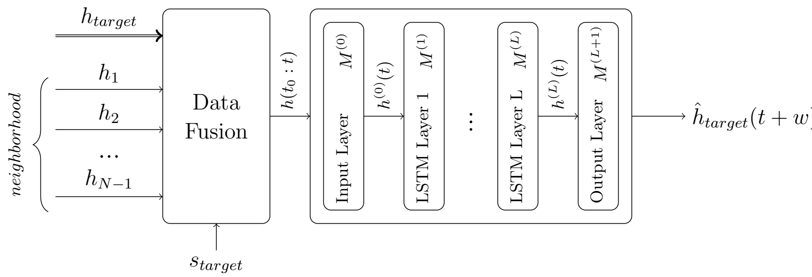

Fig. 1. Proposed unified regression model using inter-household socioeconomic information.

TABLE I
LIST OF EVALUATED DATASETS AND THEIR PROPERTIES FOR 2013. THE VALUES ARE AVERAGE VALUES AS EACH ACRON-X DATASET CONSISTS OF 50 HOUSEHOLDS

Dataset	Parameters					
	Energy (kWh)	**Avg. People**	**Avg. Age**	**Avg. Income (tsd)**	**Avg. Beds**	**Avg. Value (tsd)**
ACRON-A	4215	3.4	42.3	195	5.2	1321
ACRON-C	4772	2.7	46.5	117	3.9	599
ACRON-D	5200	3	32.7	148	3.1	1163
ACRON-E	4251	3.1	32.6	126	3.2	606
ACRON-F	3207	2.8	43.8	103	3.8	425
ACRON-G	3614	3.2	39.2	118	3.8	449
ACRON-H	3671	3.2	38.7	106	3.7	414
ACRON-I	3785	2.2	51.4	75	2.8	401
ACRON-J	3743	2.9	33.9	107	3.2	396
ACRON-L	3208	3.1	36.2	81	3.1	294
ACRON-N	3203	2.2	43.3	46	1.8	270
ACRON-O	2966	2.7	34	71	2.4	331
ACRON-P	2290	3.6	30.5	65	2.8	362
ACRON-Q	2671	2.6	33.7	46	1.9	312

suitable for training the proposed $LSTM$ architecture with inter-household energy data and socioeconomic information.

B. Feature Extraction

The importance of each of the 15 main groups of socioeconomic features was evaluated using the $ReliefF$ algorithm [24]. The results summarized for the 15 main categories are illustrated in Fig. 2

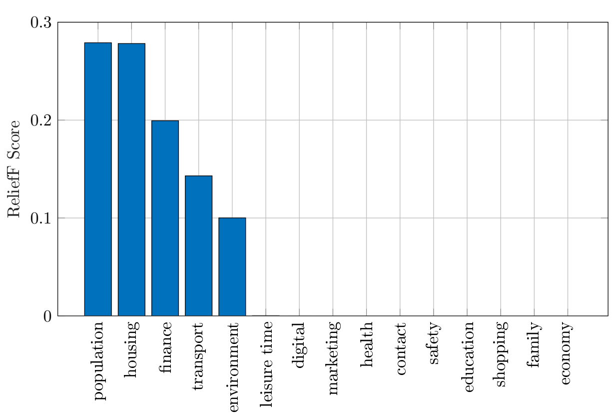

Fig. 2. Feature ranking for 15 different socioeconomic feature categories out of the $SMinL$ database

As it can be seen in Fig. 2 there are only 5 main feature categories carrying significant information about the energy consumption. In detail these are the population (e.g. number and age of residents), the house itself (e.g. size and value), the financial situation (e.g. income and number of people working), its transport (e.g. public transport) and the attitude towards the environment (e.g. how important sustainability is for the residents).

Based on the above three experimental protocols were designed, with the Baseline protocol (BL) using the previous energy values of the target household only to predict energy consumption, the Inter-Household (IH) protocol using additionally the past energy consumption of other (neighbouring) houses and the Inter-Household Socioeconomic (IHS) protocol using additionally social information as determined from the feature ranking, i.e. the 5 feature categories with the highest $ReliefF$ score. Specifically for each of the three protocols one unified $LSTM$ for all houses was trained so that the socioeconomic information, which is a constant for each house itself (e.g. number of residents in one house), could be included as an additional feature. Specifically, the same network architecture (e.g. number of layers and nodes) was used for each protocol, but was trained separately for each protocol respectively.

C. Regression Model

For the regression stage a $LSTM$ model was used, with the architecture of the $LSTM$ illustrated in Fig. 1, similar as in [21] however using a unified model for all houses as described in Section III-B. Specifically, the output of the $LSTM$ was realized as multi-step ahead prediction using all previous available samples, thus sequence-to-sequence operation was used. The free parameters of the $LSTM$, namely the number of layers L and the number of nodes M per layer were optimized after grid search on a bootstrap subset from the $SMinL$ database, using only the energy consumption data of the target house. The free parameters optimization of the $LSTM$ model with respect to the mean absolute error are shown in Table III.

As can be seen in Table III the optimized model consists of two layers with 16 nodes per layer. The state activation function of all $LSTM$ layers is the hyperbolic tangent ($tanh$), which was found to converge faster when being compared to other activation functions. Each model was trained for 50 epochs.

TABLE II

PREDICTION ERROR MAE (%) FOR THE THREE EXPERIMENTAL SETUPS UTILIZING DIFFERENT LENGTHS OF STEP-AHEAD PREDICTIONS W

Data	w=1			w=2			w=12			w=24			w=48		
	BL	IH	IHS	BL	IH	IHS	BL	IH	IHS	BL	IH	IHS	BL	IH	IHS
ACRON-A	17.2	16.8	**16.7**	19.0	18.3	**18.2**	23.5	22.3	**21.5**	23.7	22.6	**21.8**	20.4	19.9	**19.8**
ACRON-C	4.1	4.2	**3.8**	5.9	5.6	**5.5**	15.3	12.6	**11.2**	16.3	12.6	**11.0**	6.6	6.7	**6.1**
ACRON-D	4.1	3.8	**3.5**	6.1	5.1	**4.5**	17.0	12.8	**11.1**	17.9	13.8	**11.0**	6.7	7.1	**6.1**
ACRON-E	3.4	3.6	**3.3**	4.8	4.8	**4.5**	10.4	9.2	**8.4**	11.3	8.5	**7.9**	5.1	5.3	**5.0**
ACRON-F	2.8	2.8	**2.5**	3.7	3.5	**3.3**	7.5	6.5	**6.6**	7.7	6.2	**6.6**	4.4	4.5	**4.0**
ACRON-G	3.1	3.0	**2.9**	4.1	3.9	**3.8**	8.9	7.7	**7.0**	9.6	8.1	**6.9**	5.1	5.1	**4.7**
ACRON-H	3.4	3.2	**3.1**	4.7	4.3	**4.2**	10.7	9.9	**9.4**	11.4	9.2	**8.5**	5.0	5.0	**4.8**
ACRON-I	3.1	3.2	**3.0**	4.1	4.3	**4.1**	8.4	7.6	**6.8**	9.1	7.1	**6.6**	4.5	4.4	**4.1**
ACRON-J	3.5	3.4	**3.2**	4.5	4.3	**4.0**	7.1	7.1	**6.7**	7.6	6.5	**6.4**	4.2	4.3	**4.0**
ACRON-L	2.7	2.5	**2.3**	3.6	3.3	**2.8**	7.7	6.7	**6.4**	7.8	6.6	**6.1**	4.1	4.2	**3.7**
ACRON-N	3.0	2.9	**2.7**	4.0	3.8	**3.6**	8.4	7.0	**6.9**	8.7	7.2	**6.4**	4.5	4.5	**4.1**
ACRON-O	2.7	2.9	**2.6**	3.6	3.7	**3.2**	7.0	5.9	**5.5**	6.9	5.6	**5.3**	3.8	4.2	**3.4**
ACRON-P	4.3	4.4	**4.2**	6.2	6.2	**5.6**	8.4	7.8	**5.4**	8.6	8.0	**6.9**	3.3	3.8	**3.0**
ACRON-Q	2.9	2.9	**2.6**	3.5	3.7	**2.9**	6.5	5.9	**5.4**	6.5	5.0	**5.2**	3.7	3.7	**3.1**

TABLE III

PARAMETRIZATION (MAE (%)) FOR THE $LSTM$ MODEL FOR DIFFERENT NUMBERS OF LAYERS L AND NODES M

L/M	Nodes						
	2	**4**	**8**	**16**	**32**	**64**	**128**
1	4.89	4.89	4.92	4.89	5.12	4.94	4.90
2	4.98	5.03	5.10	**4.49**	4.85	5.58	5.07
3	5.36	5.00	5.85	5.52	5.67	5.85	5.51
4	5.40	5.26	4.91	5.48	5.06	5.23	5.26

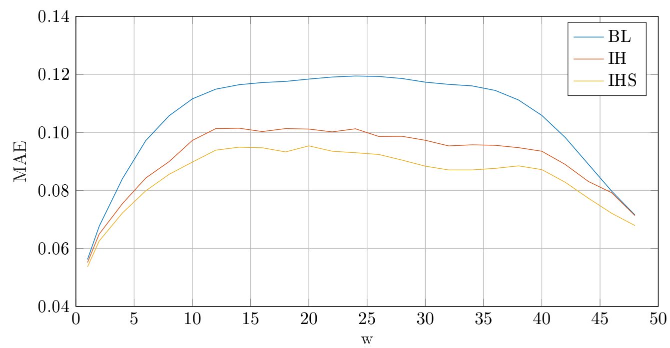

Fig. 3. Average MAE (%) for all $ACRON$ groups for different number of step-ahead predictions W

IV. EXPERIMENTAL RESULTS

The architecture presented in Section II was evaluated according to the setup described in Section III. The performance was evaluated in terms of Mean Absolute Error (MAE), i.e.

$$MAE = \frac{\sum_{t=1}^{T} |h_{target}(t) - \hat{h}_{target}(t)|}{T} \tag{6}$$

The experimental protocols were evaluated for each $ACRON$ group, for a discrete number of step-ahead predictions $W = [1, 2, 12, 24, 48]$, using 10-fold cross validation. The prediction results are tabulated in Table II.

As can be seen in Table II the proposed architecture using inter-household (IH) information outperforms the baseline system (BL) across all $ACRON$ groups. Moreover the protocol using both inter-household and socioeconomic (IHS) information further improves the energy prediction for all $ACRON$ groups. To further compare the proposed architecture the prediction error MAE is averaged over all $ACRON$ groups and illustrated in Fig. 3 for $W = [1, ..., 48]$.

As can be seen in Fig. 3 the MAE curves show similar shape for all three protocols. In detail the MAE is relatively low for small step-ahead values w, and becomes maximum for approximately $w = 24$ (i.e. prediction after 12 hours). MAE is decreasing for $w = 48$ (i.e. prediction after 24 hours) is due to the repetitive behaviour of energy consumption in a step-ahead window of 24 hours. This is in agreement with [15] where same daily, weekly, monthly or even seasonal energy consumption patterns were investigated. Furthermore it can be seen that adding the inter-household information improves significantly the energy prediction performance with a maximum performance improvement of 18.2% for $w = 32$. Moreover incorporating social information in the predictive model further improves the prediction accuracy resulting in an absolute improvement of 25.3% for $w = 32$ when compared to the baseline system.

V. CONCLUSION

An energy consumption prediction architecture was proposed. The architecture is using the load demand history of the target house as well as the load demand history of other neighboring households for energy consumption forecasting. The addition of socioeconomic information to the predictive unified model further improves the accuracy. The improvement of energy load forecasting using inter-household energy consumption is owed to common habits between consumers, trends and motifs that appear either simultaneously or with time lags across different households. Further research should focus on evaluation of the proposed architecture using incomplete data, e.g. utilizing $ACRON$ groups B,K and M.

References

[1] G.Y. Zhao et all, "Energy consumption in machining: Classification, prediction, and reduction strategy," *Energy*, vol. 133, pp. 142–157, 2017.

[2] Sung-BaeCho Tae-Young Kim, "Predicting residential energy consumption using cnn-lstm neural networks," *Energy*, vol. 182, pp. 72–81, 2019.

[3] H. Ibrahim, A. Ilinca, J. Perron, "Energy storage systems—characteristics and comparisons," *Renewable and Sustainable Energy Reviews*, vol. 12, no. 5, pp. 1221–1250, 2008.

[4] M. Erdem Günay, "Forecasting annual gross electricity demand by artificial neural networks using predicted values of socio-economic indicators and climatic conditions: Case of turkey," *Energy Policy*, vol. 90, no. C, pp. 92–101, 2016.

[5] Hans W. Sinn, "Buffering volatility: A study on the limits of germany's energy revolution," *European Economic Review*, vol. 99, pp. 130–150, 2017.

[6] Payam Nejat et all, "A global review of energy consumption, co2 emissions and policy in the residential sector (with an overview of the top ten co2 emitting countries)," *Renewable and Sustainable Energy Reviews*, vol. 43, pp. 843–862, 2015.

[7] Chinthaka Dinesh, Stephen Makonin, and Ivan V. Bajic, "Residential power forecasting using load identification and graph spectral clustering," *IEEE Transactions on Circuits and Systems II: Express Briefs*, p. 1, 2019.

[8] Chirag Deb et all, "A review on time series forecasting techniques for building energy consumption," *Renewable and Sustainable Energy Reviews*, vol. 74, pp. 902–924, 2017.

[9] Pascal A. Schirmer and Iosif Mporas, "Improving energy disaggregation performance using appliance-driven sampling rates," in *2019 27th European Signal Processing Conference (EUSIPCO)*. 2019, pp. 1–5, IEEE.

[10] Pascal A. Schirmer, Iosif Mporas, and Michael Paraskevas, "Energy disaggregation using elastic matching algorithms," *Entropy*, vol. 22, no. 1, pp. 71, 2020.

[11] Pascal A. Schirmer and Iosif Mporas, "Statistical and electrical features evaluation for electrical appliances energy disaggregation," *Sustainability*, vol. 11, no. 11, pp. 3222, 2019.

[12] W. Wichakool, A.-T. Avestruz, R. W. Cox, and S. B. Leeb, "Modeling and estimating current harmonics of variable electronic loads," *IEEE Transactions on Power Electronics*, vol. 24, no. 12, pp. 2803–2811, 2009.

[13] Kadir Amasyali, Nora M. El-Gohary, "A review of data-driven building energy consumption prediction studies," *Renewable and Sustainable Energy Reviews*, vol. 81, pp. 1192–1205, 2018.

[14] Huiling Cai, Shoupeng Shen, Qingcheng Lin, Xuefeng Li, and Hui Xiao, "Predicting the energy consumption of residential buildings for regional electricity supply-side and demand-side management," *IEEE Access*, vol. 7, pp. 30386–30397, 2019.

[15] Peter Lusis, Kaveh Rajab Khalilpour, Lachlan L. H. Andrew, Ariel Liebman, "Short-term residential load forecasting: Impact of calendar effects and forecast granularity," *Applied Energy*, vol. 205, pp. 654–669, 2017.

[16] Petra Vrablecová, Anna Bou Ezzeddine, Viera Rozinajová, Slavomír Šárik, and Arun Kumar Sangaiah, "Smart grid load forecasting using online support vector regression," *Computers & Electrical Engineering*, vol. 65, pp. 102–117, 2018.

[17] Yu Shimizu, Taichiro Sakagami, and Hiraoki Kitano, "Prediction of weather dependent energy consumption of residential housings," in *The 6th IEEE International Conference on Renewable Energy Research and Applications (ICRERA 2017)*, [Piscataway, New Jersey], 2017, pp. 967–970, IEEE.

[18] Changwan Kim Hyojoo Son, "Short-term forecasting of electricity demand for the residential sector using weather and social variables," *Resources, Conservation and Recycling*, vol. 123, pp. 200–207, 2017.

[19] Shirantha Welikala, Chinthaka Dinesh, Mervyn Parakrama B. Ekanayake, Roshan Indika Godaliyadda, and Janaka Ekanayake, "Incorporating appliance usage patterns for non-intrusive load monitoring and load forecasting," *IEEE Transactions on Smart Grid*, vol. 10, no. 1, pp. 448–461, 2019.

[20] Pascal A. Schirmer, Iosif Mporas, and Ilyas Potamitis, "Evaluation of regression algorithms in residential energy consumption prediction," in *2019 3rd European Conference on Electrical Engineering and Computer Science (EECS)*. 2019 (in press), pp. 1–5, IEEE.

[21] Lukas Mauch and Bin Yang, "A new approach for supervised power disaggregation by using a deep recurrent lstm network," in *2015 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Piscataway, NJ and Piscataway, NJ, 2015, pp. 63–67, IEEE.

[22] Jean-Michel D., *Smart meters in London*, https://www.kaggle.com/jeanmidev/smart-meters-in-london, 2017.

[23] UK Power Networks, *SmartMeter Energy Consumption Data in London Households*, https://data.london.gov.uk/dataset/smartmeter-energy-use-data-in-london-households, 2014.

[24] Igor Kononenko, Marko Robnik-Sikonja, Marko Robnik, Uros Pompe, "Relieff for estimation and discretization of attributes in classification, regression, and ilp problems," 1996.