Evaluating Forecasting Techniques for Integrating Household Energy Prosumers into Smart Grids

Teodor Petrican, Andreea Valeria Vesa, Marcel Antal, Claudia Pop, Tudor Cioara, Ionut Anghel, Ioan Salomie
Email: teodor.petrican@student.utcluj.ro, {andreea.vesa, marcel.antal, claudia.pop}@cs.utcluj.ro, {tudor.cioara, ionut.anghel, ioan.salomie}@cs.utcluj.ro

Department of Computer Science, Technical University of Cluj-Napoca
Distributed Systems Research Laboratory

***Abstract*—This paper tackles the problem of integrating household energy prosumers in Smart Energy Grids by analyzing a set of state-of-the-art energy forecasting techniques that allow individual or aggregated prosumers to evaluate their future energy demand and inform the Distributed System Operator (DSO) about potential grid imbalances. Thus, the DSO can perform a proactive strategy to manage the grid and avoid problems before they appear. The key element of this approach is the prediction technique, that must be accurate enough such that the resulting grid imbalances can be compensated in real-time. The paper evaluates a set of state-of-the-art statistical and Machine Learning (ML) prediction techniques, such as SARIMA, feed-forward and recurrent neural networks, support vector regression or ensemble prediction models, on real household historical energy demand logs by performing a feature selection process for each ML algorithm as to identify the best elements that influence the energy demand of a house. A set of experiments are performed on the REFIT Electrical Load Measurements data set evaluating each model's performance with respect to the selected features. Among the evaluated algorithms, the Ensemble Prediction Model gives best prediction accuracy, showing a Mean Absolute Percentage Error (MAPE) of 14.4% followed by the SVM model with a MAPE of 15.4%.**

***Index Terms*—Smart Grids, Energy Demand Forecasting, Short-Term Load Forecasting, Multi-layer Neural Network, Recurrent Neural Networks, Support Vector Regression**

I. INTRODUCTION

Energy demand forecasting has become of real interest nowadays, since more and more focus is put on designing efficient systems that are able to control and optimize energy consumption and production worldwide. The key benefits of having such a prediction system are both environmental and economical. During the last decades energy consumption has increased exponentially therefore a proper management has become necessary to keep a sustainable and secure environment. As electric energy cannot be stored for future use, it has become crucial to produce as much energy as needed, and this needs to happen almost in real-time. On the global electricity market, a more accurate hourly/daily prediction is of great interest to obtain in real-time the best purchase prices. Moreover the adoption of different pricing methods is encouraged such as time-of-use billing, demand-based pricing etc. [1]

Smart Grids are placing the energy industry in a new era meant to offer a better integration of producer-consumer power generation systems, with major benefits in the economical, environmental and security directions. Being a two-way communication between the utility (usually called the Distributed System Operator (DSO)) and its consumers, an optimization of the energy consumption is critical for maintaining the power grid reliability and avoiding supply-demand mismatches. As presented in Figure 1 the Smart Grid is a distributed system where each consumer (households in our case) is bidirectionally connected with the DSO. As the available power in the grid varies over time, the DSO may suggest consumers energy curves to reduce their consumption when rates are lower. This is done through a Demand Response (DR) signal. The consumers may accept or not the proposal. In case of acceptance the consumers need to reschedule their consumption according to the DR signal. A new energy curve will be sent to the utility that represents the actual day ahead plan.

A crucial aspect of the system described above is the Energy Consumption Prediction Module, as the DSO will make adjustment decisions based on this module's output, the decisions quality being tightly related to the predictions quality. The focus of our study is to evaluate several models for short-term energy consumption forecasting. The goal is to obtain forecasts of energy consumption for the next day at a granularity of one hour, thus having a 24 steps ahead prediction.

The main contribution of this paper is a thorough investigation of state-of-the-art machine learning algorithms used for forecasting and evaluating various instances of their models with different features on the REFIT Electrical Load Measurement data set [2] aiming to determine the best algorithm configuration for energy prediction.

The rest of the paper is structured as follows: Section II shows related work, Section III presents the models of the machine learning algorithms evaluated for forecasting, Section IV presents the evaluation of the predictions on a real-world dataset, while Section V concludes the paper.

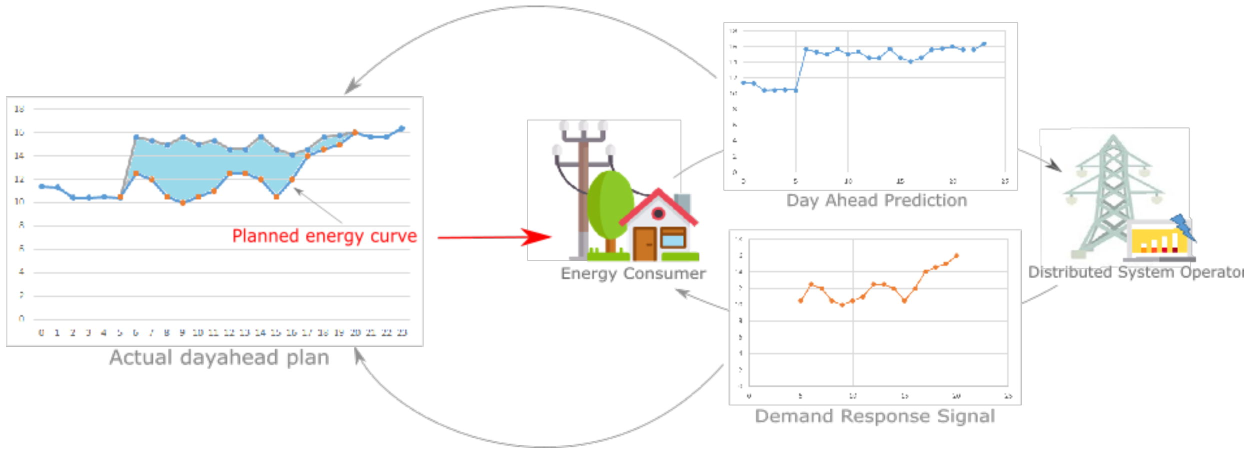

Fig. 1. Day ahead energy scheduling plan in the context of Smart Grids

II. Related Work

During the last decades, various energy demand forecasting models have been examined by several researchers, at different levels, varying from the very short range to long term. This paper focuses on short-term load forecasting (STLF) techniques. Srivatstava et al. [3] have classified the forecasting models into four categories: statistical approaches (multiple regression [4], iterative reweighted least square, adoptive load forecasting, stochastic time series (ARIMA [5]) etc.), artificial intelligent techniques (fuzzy logic [6], neural networks etc.), genetic algorithms [7] and knowledge based expert systems [8].

Recently a study made by Alani et al. [9] emphasizes the strengths and weaknesses of most popular algorithms used for STLF. While ARIMA (Autoregressive Integrated Moving Average) has been widely used as a baseline method for evaluating other STLF methods, lately a special focus has been put on evaluating neural networks [10]. ARIMA works under the assumption that the observed and the future values in the time series are linearly related, while neural networks are more capable of capturing nonlinearities in the data. However certain variations of ARIMA models, Seasonal ARIMA (SARIMA) have been used in combinations with other models to perform energy load prediction: with SVM [11], with regression [12], with neural networks [13].

A set of studies that use neural networks to predict the energy demand have been carried out in various countries [14]–[20]. Neural networks are described by several authors as having a greater capability in predicting energy consumption rather than traditional methods such as ARIMA or regression methods. The authors of [18] show a short-term prediction model based on artificial neural networks that feeds back part of its outputs and demonstrates high precision. Another solution proposed in [11], where authors develop a residential building energy consumption model based on a Back Propagation ANN model, shows that the model is in line with the energy trends.

Certain studies have proposed Suport Vector Regression (SVR) technique as a feasible alternative to ANNs for predicting energy consumption [1]. Paudel et al. [21] present a SVM based solution that uses beside energy consumption measurements climate data and the approximate occupancy profile of the building. Relevant historical data is also chosen based on a dynamic time warping technique. The latter seems to reduce significantly the training time, yet preserving a high precision. The advantage of SVR models over ANNs relies in their ability to mitigate the problem of getting stuck into a local minima while training [22], still having a good performance. Fux et al. [23] compared a SVM model against a neural network model for one day ahead prediction, concluding that SVM based models generally perform better having less parameters to be configured and being easier to manipulate.

In the last years, deep neural networks gained more and more interest in the research area, being used in various domains, from computer vision to modeling and forecasting. A novel methodology for load prediction based on Deep Neural Networks is presented in [24]. Two LSTM algorithms are compared: LSTM standard and LSTM-based Sequence to Sequence (S2S). Experiments performed at various time granularities show that the standard LSTM fails at one-minute time granularity, while the LSTM S2S performs well at both one minute and half hour time granularities. Since then several papers have tackled the problem of STLF using this type of networks [25]–[28].

It seems that there is not a consensus regarding which prediction model is better, some claiming that complex regression models fail to beat the autoregressive model [29]. Also, as already noted, some authors claim SVM as yielding the best results, while others claim that ANNs are better. This suggests that models' accuracy is dependent on a variety of factors, an important one being the underlying dataset and its characteristics.

Many studies have concentrated on commercial or industrial buildings, that have strong consumption patterns such as seasonality, low power consumption during weekends versus weekdays etc., however little study has been carried in terms of individual household consumption prediction [30], [31]. The latter problem proves to be more challenging as several uncertainty factors are present, such as human activity, holidays, various appliances, climate etc.

The aim of this paper is to implement and inspect several prediction models for aggregated and individual household consumption for STLF to identify challenges and issues and set up the expectations from prediction modules for integrating household energy prosumers into the Smart Grid.

III. Prediction Models

The prediction module defines a parametric mathematical model of the household energy consumption pattern, and can be viewed as a black box mathematical function that receives an input containing historical data and computes the predicted values for the future. In order to determine the parameters of the model, a training process is performed on the historical data. The training process is an optimization process that aims to determine the best parameters of the model that minimize the error between the predicted values and the training values. From a machine learning point of view, the inputs of the function are called features and the output is the forecast or predicted value. A key step in developing machine learning algorithms is feature selection.

To achieve a relevant feature selection, the prediction problem must be analyzed. In our case, the prediction module has to predict the hourly energy consumption values for the next day, i.e. 24 steps, such that the households can be integrated with the day-ahead energy marketplace. This falls under the category of STLF or Short-Term Load Forecasting. The most relevant elements that might influence the household energy consumption pattern over a 24-hour period must be identified.

As is generally the case with any prediction model, we make use of the historical consumption monitoring in order to determine the consumption pattern for next 24 hours. This can be further enhanced by adding temporal information as input, as we expect different consumption behaviors at different times of the day, week and year respectively. Thus, we add the following additional temporal inputs: current month, day of week and hour of day. The energy consumption patterns might also be influenced by the climate, so we incorporate weather data, such as the temperature and the amount of rain. A conceptual black box view of this module can be seen in Figure 2.

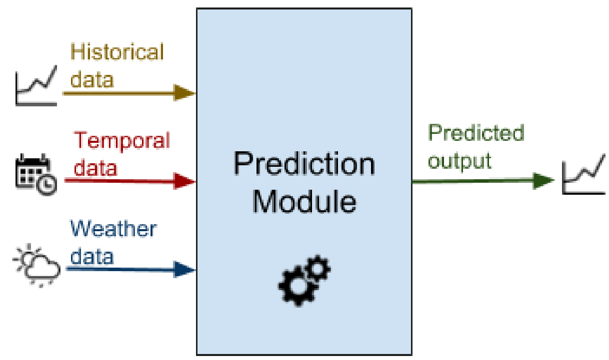

Fig. 2. Conceptual representation of the prediction module

We implement five different predictions models: one time-series based approach (SARIMA), two neural network approaches (MLP and LSTM), a support vector regression model and an ensemble model. We test the behavior of said models both at individual household level and aggregated consumption level. We expect to obtain better results at aggregated level due to the individual's behavior playing a smaller role in the variability of data.

A. SARIMA Prediction Model

The first model we employ is a time-series based linear regression model, specifically SARIMA (Seasonal Autoregressive Integrated Moving Average), due to its simplicity and ubiquitous nature in literature. We use this model's results as a baseline for other more advanced models. If linear regression leads to better results, the more advanced models would be unjustified.

The ARIMA model has 3 parameters, p, d and q which specify the size of the autoregressive window, the differencing amount and the size of the moving-average window respectively. In addition to those, SARIMA model has four extra parameters, (P, D, Q and S) the first three having the same meaning, but operating on the seasonal component, an the last (S) specifying the time span of the seasonal pattern. The model can be shortly written as $(p, d, q) \times (P, D, Q)S$. There are various approaches in literature about how to determine the most appropriate parameters for a given dataset. In this paper, we have used a grid-search approach that tries different values for the parameters and use the Akaike information criterion [32] to check the best model. The S parameter value was set to 24 in this case, since our data is hourly. Indeed, there is a strong daily seasonality on the dataset presented in section IV which can be checked with decomposition. We have added a one-step seasonal and non-seasonal differencing, thus $d = D = 1$. The other parameters were determined via grid-search, reaching to the following model:

$$SARIMA(1, 1, 1) \times (2, 1, 0)24 \tag{1}$$

The last 2 weeks of historical consumption data were used as input, with no other additional information given.

B. MLP Prediction Model

The second model we employed moves away from the time-series approach and uses a deep-learning model, namely a fully-connected multi-layer perceptron (MLP) which is a type of feedforward neural network.

Although from a black box perspective the model operates in the same way as the linear regression, it is conceptually different. Historical data is no longer treated as a time-series, as it is all seen as distinct features with no temporal dependency between them. The MLP architecture will have a number of 24 outputs, the size of the forecasting window.

We have trained and tested various MLP configurations on the dataset described in Section IV-A, to determine how the number of layers, neurons, epochs, optimizers influence the prediction. We concluded that the following configuration works best in our case: MLP with one hidden layer with 100 neurons, $ReLu$ activation functions for the hidden layer, $tanh$ activation function for the output layer, ADAM optimizer with MSE as loss function and a He uniform variance scaling initializer [33]; we have trained the model with a batch size of 500 and 300 epochs.

The input features for the network consisted of the last 3 days of historical data (72 features) and calendar and temperature features as described previously (hourly temperature in Celsius and rainfall in mm).

As a preliminary step, we have applied a one-step differencing to the data and a min-max scaling to bring the data in the interval $[-1, 1]$ before feeding it into the network. Min-max scaling is a linear transformation that preserves the relationships among the original data. This is a necessary step as we have heterogeneous features and scaling brings them in the same range. After forecasting, the results are scaled back to their original size, using the inverse of the scaling function initially used.

C. LSTM Prediction Model

Another network model we used is a recurrent Long Short-Term Memory network [34], that has recently gained increased popularity for time series forecasting and is well regarded in coping with irregularities of data. The idea behind the LSTM network is that a module contains a feeedback loop, like all recurrent networks and, in addition, a *memory cell* achieved with the help of a series of gates [35]. This way, the network is able to capture long term dependencies in the time series.

As opposed to the MLP approach, the LSTM network captures temporal dependencies among data with the help of the memory cell. There are multiple approaches: either use a seq2seq model which receives inputs for a window of 24 timesteps and offers 24 outputs, or use a model which receives features corresponding to 1 timestep and issues one output, letting the model build the "window" internally. In the latter case, in order to predict the entire forecasting window one can either feed the output back as input to the network, or train 24 independent networks, each of them trained to predict one of the 24 hours in the future.

We have tried both approaches with inputs for one timestep, and found that training 24 networks leads to better results: feeding the output back to input cascades the error at each timestamp cumulating in dramatic errors towards the end of the forecasting frame. Thus, we have 24 networks, each of them the same input, i.e. the consumption for the last hour, calendar and temperature data.

The same scaling as in the case of the MLP network has been applied, but no differencing, as we have observed that differencing impacts negatively the quality of the model.

D. SVR Prediction Model

Another approach we propose is based on Support Vector Regression (SVR). Recently, SVRs have been widely used as an alternative for energy consumption forecasting solutions due to the fact they are easy to configure and expressive enough to capture non-linearities in the data. The only major step required in the configuration phase for SVR is choosing an appropriate kernel function.

For our problem, 24 individual SVR models have been trained to predict every step of the forecasting window. That is, each i-th model has been trained to predict the energy consumption for the i-th hour from a 24 hour window, similar to the approach taken for the LSTM-based model described above. The final result represents a concatenation of all the results obtained from the individual models. Several kernel functions (Gaussian, Linear, Polynomial) have been evaluated, out of which Gaussian kernels was found to give the best results in our case. Prior to training, the data has been standardized to have 0 mean and a variance of 1.

E. Ensemble Prediction Model

We also explored the possibility of ensemble forecasting. The ensemble model can be viewed as a box containing all the other models and an additional module at the end which combines the outputs from all the models to generate an aggregated one.

There are different possible approaches to combine the outputs of the models, such as median, average or weighted average [36]. We combine them using a weighted average of the outputs, where the weights are given by the performance of each model, as in the following formula:

$$E(t) = \frac{\sum_k fibo(rank(k) + 1) * E_k(t)}{\sum_k fibo(rank(k) + 1)} \qquad (2)$$

where:

- $rank(k)$ - position of the k^{th} model in the list of models sorted decreasingly by $RMSE(k)$
- $fibo(x)$ - the x^{th} Fibonacci number

IV. Experimental Results

This section starts with a description of the dataset and the experimental setup used in our work. Next we present the evaluation results of the forecasting methods described in Section III.

A. Experimental Setup and Dataset

Our experiments are conducted on the REFIT Electrical Load Measurements data set [2]. The data set contains 1,194,958,790 readings of household electrical consumption measurements taken over a period of two years from 20 households in the United Kingdom at aggregate and appliance level. The measurements represent the power in Watts timestamped at irregular time intervals of about 6-8 seconds, our goal being to forecast at the energy level, which is an integral of power over time. The following preprocessing steps were needed:

TABLE I
DAY-AHEAD ENERGY FORECASTING PREDICTION RESULTS ON AGGREGATED DATASET

Prediction Model	Prediction on entire dataset			Prediction on entire dataset with Weather		
	RMSE [Wh]	MAPE [%]	MAE [Wh]	RMSE [Wh]	MAPE [%]	MAE [Wh]
SARIMA	115.99	20.68	80.12	-	-	-
MLP	120.27	23.04	79.68	108.38	20.76	77.03
LSTM	112.16	19.29	78.58	105.89	18.09	73.48
SVR	94.33	15.40	62.83	94.19	17.16	65.91
Ensemble	-	-	-	86.36	14.40	57.16

TABLE II
DAY-AHEAD ENERGY FORECASTING PREDICTION RESULTS ON INDIVIDUALS

Prediction Model	Prediction on individual household		
	RMSE [Wh]	MAPE [%]	MAE [Wh]
SARIMA	519.51	101.67	332.82
MLP	560.69	158.06	427.41
LSTM	518.93	100.84	381.54
SVR	467.33	53.45	297.10
Ensemble	476.28	68.51	315.59

1) Transforming the data from power (W) to energy (kWh), with one measurement per hour
2) Overcoming the problem of missing measurements
3) Aggregating the electric consumption from all households

Due to the discrete nature of the data and the irregularity in timestamping, we have estimated the energy by aggregating the measurements in a weighted average manner, where the weights correspond to the amount of time elapsed between each measurement and the one before, as in the following formula:

$$E(x) = \frac{1}{3.6 * 10^3} \sum_{i=1}^{N_x} w(i) * P(i) \tag{3}$$

$$w(i) = \begin{cases} t_x(i) & \text{if } i = 1 \\ t_x(i) - t_x(i-1) & \text{if } i \in [2, N_x - 1] \\ 3600 - t_x(i-1) & \text{if } i = N_x \end{cases} \tag{4}$$

where:

- $E(x)$ - energy consumption at hour x (kWh)
- $P(i)$ - power at i-th measurement (W)
- $t_x(i)$ - no. of seconds passed for the i-t measurement from start of hour x (s)
- N_x - no. of measurements that occurred during hour x

For the second step of the preprocessing, we took two approaches. There were period where a great amount of data was missing, specifically towards the beginning and end of the dataset. For our purpose, these periods were excluded from the dataset. For the remaining cases, we have filled the missing entries with the aggregated values from the other houses.

Finally, the last step of preprocessing is to compute an aggregated data set and is done as an average of the consumption of all the households.

The dataset also offers a wide array of climate information from a station near the households at a 15-minute granularity. From this, we have used the temperature and the amount of rainfall aggregated to one entry per hour to match the energy consumption entries.

The models described in Section III were implemented and evaluated using Python 3. The neural networks were described with the help of Keras library [37] working with the underlying TensorFlow library [38]. All the training and evaluations were done on a computer with the following configuration: Intel Core i7-7700K, 32 GB RAM DDR4 and nVidia GeForce GTX 1070 with CUDA.

B. Prediction Results

Three error metrics are used to evaluate the performance of the approaches discussed in Section III: root mean square error (RMSE), mean absolute percentage error (MAPE) and mean absolute error (MAE). In literature, there is a lack of consensus regarding the evaluation metric, although these three errors are frequently used. MAPE error, while it has some major drawbacks [39], is useful as it provides a dimensionless metric that can be used to compare models across different datasets.

In order to compute the results, we have divided the dataset in two parts, one for training and one for testing, comprising of 4 weeks of data.

With the exception of SARIMA, the models are also tested both with and without weather data as features, to verify whether these additional inputs bring anything valuable to the performance of the models.

The results for predictions on the aggregated consumption are presented in Table I, while the results for the individual households are presented in Table II. Also, the predictions of all the models for two consecutive days are presented in Figure 3 and, to make it easier to visualize, the prediction for one day ahead for the best model is presented in Figure 4.

As it can be seen in Figure 3, all the models follow the trends and amplitude of the data quite well; this holds true over the test set. A surprising result is that SARIMA without

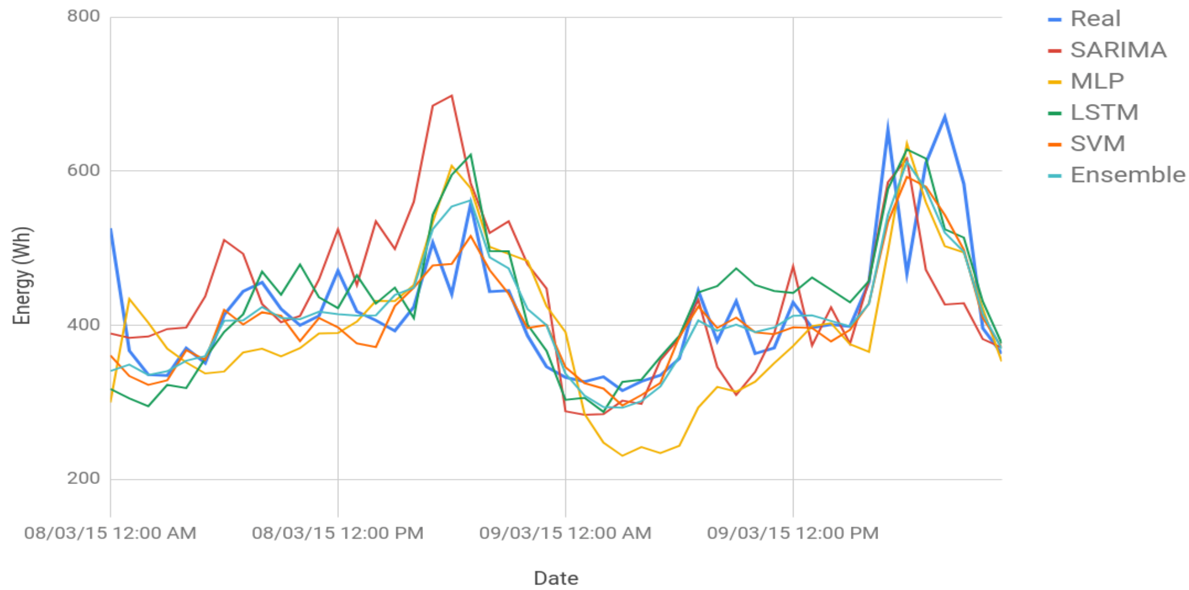

Fig. 3. Forecasting comparison for the aggregated electricity consumption data set on two consecutive days

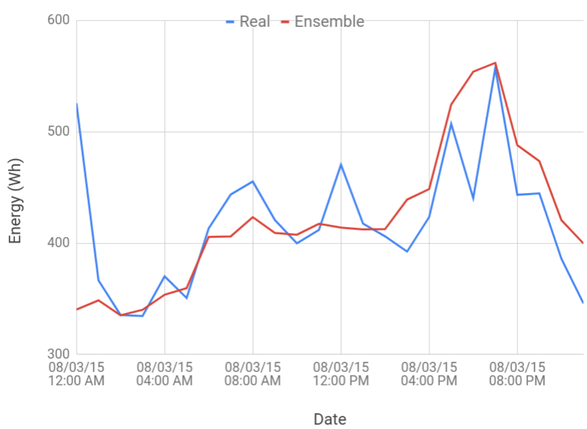

Fig. 4. Next day prediction using the Ensemble Model

any weather data is giving better average results than MLP. Thus, SARIMA is still a valid approach which might justify its use in certain contexts, as it is easy to set up and works with only historical data. Introducing weather features for the MLP model definitely helps, reducing all of the three metrics.

LSTM is giving better results than SARIMA and MLP, which is promising. This type of neural network is well known for requiring a large amount of training data, so we might expect better results given a larger amount of training data. It has to be noted that the training time for this model is quite long. While this is not a major problem for 24 step-ahead STLF, the feasibility of the model needs to be evaluated should there be the need for a greater number of forecasting steps. Though, given the nature of this neural network we expect that re-trainings are seldom required.

SVM gives the best results among the individual models. Interestingly, it performs better without using the weather features as input, though only marginally (RMSE is actually a bit worse).

The ensemble model comprising of a weighted average of all the models reduces the error even more, to an average MAPE of 14.4% which is the best result we obtained. This is somewhat expected, as the errors of the individual models are uncorrelated and this approach helps canceling out some of them.

The results for individual households, as it can be seen from the tables, are far worse than the aggregated ones, SVR giving the best results. This confirms our expectations, as there is a lot of variability introduced by the residents' behavior. Also, the consumption is too easily influenced by any household device which might be used sporadically.

V. Conclusions

This paper presented a thorough evaluation of several statistical and machine learning algorithms used for household energy consumption prediction aimed to aid the integration of households with Smart Energy Grids. The SARIMA algorithm, two neural network architectures, namely Multi-Layer Perceptron (MLP) and Long Short-Term Memory (LSTM) cells, Support Vector Regression (SVR) and Ensemble models were compared and evaluated against each other on a trace of real household historical monitored energy consumption values. Several models of each algorithms were instantiated, each having different features, such as historical data, temporal and weather data. The experimental results show that forecasting accuracy increases as the aggregation level increases, results being already much better for 20 households than an individual one. We expect this phenomenon to maintain as the aggregation level gets higher. These findings should be taken into account when computing the demand response signals.

On aggregated load forecasting, the best results for a single model are obtained by the SVR algorithm, yielding a MAPE error of 15.4%, while the best overall results are obtained with an ensemble model, yielding a MAPE error of 14.4%.

ACKNOWLEDGMENT

This work has been conducted within the eDREAM project Grant number 774478, co-funded by the European Commission as part of the H2020 Framework Programme (H2020-LCE-2017-SGS).

REFERENCES

[1] J. Campillo, F. Wallin, D. Torstensson, and I. Vassileva, "Energy demand model design for forecasting electricity consumption and simulating demand response scenarios in sweden," in *4th International Conference in Applied Energy 2012, July 5-8, 2012. Suzhou, China.*, 2012.

[2] D. Murray, L. Stankovic, and V. Stankovic, "An electrical load measurements dataset of United Kingdom households from a two-year longitudinal study," *Scientific data*, vol. 4, p. 160122, 2017.

[3] A. Srivastava, A. S. Pandey, and D. Singh, "Short-term load forecasting methods: A review," in *Emerging Trends in Electrical Electronics & Sustainable Energy Systems (ICETEESES), International Conference on*. IEEE, 2016, pp. 130–138.

[4] I. Moghram and S. Rahman, "Analysis and evaluation of five short-term load forecasting techniques," *IEEE Transactions on power systems*, vol. 4, no. 4, pp. 1484–1491, 1989.

[5] M. Cho, J. Hwang, and C. Chen, "Customer short term load forecasting by using ARIMA transfer function model," in *Energy Management and Power Delivery, 1995. Proceedings of EMPD'95., 1995 International Conference on*, vol. 1. IEEE, 1995, pp. 317–322.

[6] K. Liu, S. Subbarayan, R. Shoults, M. Manry, C. Kwan, F. Lewis, and J. Naccarino, "Comparison of very short-term load forecasting techniques," *IEEE Transactions on power systems*, vol. 11, no. 2, pp. 877–882, 1996.

[7] D. Srinivasan, "Evolving artificial neural networks for short term load forecasting," *Neurocomputing*, vol. 23, no. 1-3, pp. 265–276, 1998.

[8] K.-L. Ho, Y.-Y. Hsu, C.-F. Chen, T.-E. Lee, C.-C. Liang, T.-S. Lai, and K.-K. Chen, "Short term load forecasting of Taiwan power system using a knowledge-based expert system," *IEEE Transactions on Power Systems*, vol. 5, no. 4, pp. 1214–1221, 1990.

[9] A. Y. Alani and I. O. Osunmakinde, "Short-term multiple forecasting of electric energy loads for sustainable demand planning in smart grids for smart homes," *Sustainability*, vol. 9, no. 11, p. 1972, 2017.

[10] S. Yao, Y. Song, L. Zhang, and X. Cheng, "Wavelet transform and neural networks for short-term electrical load forecasting," *Energy conversion and management*, vol. 41, no. 18, pp. 1975–1988, 2000.

[11] M. Braun, T. Bernard, O. Piller, and F. Sedehizade, "24-hours demand forecasting based on sarima and support vector machines," *Procedia Engineering*, vol. 89, pp. 926–933, 2014.

[12] D. Chikobvu and C. Sigauke, "Regression-sarima modelling of daily peak electricity demand in south africa," *Journal of Energy in Southern Africa*, vol. 23, no. 3, pp. 23–30, 2012.

[13] Y. Yang, J. Wu, Y. Chen, and C. Li, "A new strategy for short-term load forecasting," vol. 2013, 05 2013.

[14] M. Kankal, A. Akpınar, M. İ. Kömürcü, and T. Ş. Özşahin, "Modeling and forecasting of Turkeys energy consumption using socio-economic and demographic variables," *Applied Energy*, vol. 88, no. 5, pp. 1927–1939, 2011.

[15] G. Oğcu, O. F. Demirel, and S. Zaim, "Forecasting electricity consumption with neural networks and support vector regression," *Procedia-Social and Behavioral Sciences*, vol. 58, pp. 1576–1585, 2012.

[16] Ş. Birim and A. Tümtürk, "Modeling and forecasting Turkeys electricity consumption by using artificial neural network," *American Scientific Research Journal for Engineering, Technology, and Sciences (ASRJETS)*, vol. 25, no. 1, pp. 192–208, 2016.

[17] M. J. Kargar and D. Charsoghi, "Predicting annual electricity consumption in iran using artificial neural networks (NARX)," *Indian J. Sci. Res*, vol. 5, no. 1, pp. 231–242, 2014.

[18] L. Ekonomou, "Greek long-term energy consumption prediction using artificial neural networks," *Energy*, vol. 35, no. 2, pp. 512–517, 2010.

[19] K. Kandananond, "Forecasting electricity demand in Thailand with an artificial neural network approach," *Energies*, vol. 4, no. 8, pp. 1246–1257, 2011.

[20] B. M. AbuAl-Foul, "Forecasting energy demand in Jordan using artificial neural networks," *Topics in Middle Eastern and North African Economies*, vol. 14, 2012.

[21] S. Paudel, P. H. Nguyen, W. L. Kling, M. Elmitri, B. Lacarrière, and O. L. Corre, "Support vector machine in prediction of building energy demand using pseudo dynamic approach," *arXiv preprint arXiv:1507.05019*, 2015.

[22] X. Li, D. Gong, L. Li, and C. Sun, "Next day load forecasting using svm," in *International Symposium on Neural Networks*. Springer, 2005, pp. 634–639.

[23] S. F. Fux, M. J. Benz, A. Ashouri, and L. Guzzella, "Short-term thermal and electric load forecasting in buildings," in *CISBAT (International Conference on Clean-tech for Smart Cities & Buildings: From Nano to Urban Scale)*, no. EPFL-CONF-211899, 2013.

[24] D. L. Marino, K. Amarasinghe, and M. Manic, "Building energy load forecasting using deep neural networks," in *Industrial Electronics Society, IECON 2016-42nd Annual Conference of the IEEE*. IEEE, 2016, pp. 7046–7051.

[25] H. Zheng, J. Yuan, and L. Chen, "Short-term load forecasting using EMD-LSTM neural networks with a Xgboost algorithm for feature importance evaluation," *Energies*, vol. 10, no. 8, p. 1168, 2017.

[26] 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*, 2017.

[27] W. He, "Load forecasting via deep neural networks," *Procedia Computer Science*, vol. 122, pp. 308–314, 2017.

[28] L. Di Persio and O. Honchar, "Analysis of recurrent neural networks for short-term energy load forecasting," in *AIP Conference Proceedings*, vol. 1906, no. 1. AIP Publishing, 2017, p. 190006.

[29] P. Dagnely, T. Ruette, T. Tourwé, E. Tsiporkova, and C. Verhelst, "Predicting hourly energy consumption. can regression modeling improve on an autoregressive baseline?" in *DARE*, 2015.

[30] K. Gajowniczek and T. Zabkowski, "Short term electricity forecasting using individual smart meter data," *Procedia Computer Science*, vol. 35, pp. 589–597, 2014.

[31] F. Javed, N. Arshad, F. Wallin, I. Vassileva, and E. Dahlquist, "Forecasting for demand response in smart grids: An analysis on use of anthropologic and structural data and short term multiple loads forecasting," *Applied Energy*, vol. 96, pp. 150–160, 2012.

[32] H. Akaike, *Information Theory and an Extension of the Maximum Likelihood Principle*. New York, NY: Springer New York, 1998, pp. 199–213.

[33] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 1026–1034.

[34] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997.

[35] H. Sak, A. Senior, and F. Beaufays, "Long short-term memory recurrent neural network architectures for large scale acoustic modeling," *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, pp. 338–342, 01 2014.

[36] P. Valtonen, S. Honkapuro, and J. Partanen, "Improving short-term load forecast accuracy by utilizing smart metering," in *CIRED Workshop*, Lyon, France, 2010.

[37] F. Chollet *et al.*, "Keras," https://keras.io, 2015.

[38] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, "TensorFlow: Large-scale machine learning on heterogeneous systems," 2015, software available from tensorflow.org. [Online]. Available: https://www.tensorflow.org/

[39] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," *International journal of forecasting*, vol. 22, no. 4, pp. 679–688, 2006.