Accurate and Efficient Selection of the Best Consumption Prediction Method in Smart Grids

Marc Frincu, Charalampos Chelmis, Muhammad Usman Noor and Viktor Prasanna
Department of Electrical Engineering
University of Southern California
Los Angeles, CA, USA
Email: {frincu,chelmis,mnoor,prasanna}@usc.edu

***Abstract*—Smart grids are becoming popular with the advent of sophisticated smart meters. They allow utilities to optimize energy consumption during peak hours by applying various demand response techniques including voluntary curtailment, direct control and price incentives. To sustain the curtailment over long periods of time of up to several hours utilities need to make fast and accurate consumption predictions on a large set of customers based on a continuous flow of real time data and huge historical data sets. Given the numerous consumption patterns customers exhibit, different prediction methods need to be used to reduce the prediction error. The straightforward approach of testing each customer against every method is unfeasible in this large volume and high velocity environment. To this aim, we propose a neural network based approach for automatically selecting the best prediction method per customer by relying only on a small subset of customers. We also introduce two historical averaging methods for consumption prediction that take advantage of the variability of the data and continuously update the results based on a sliding window technique. We show that once trained, the proposed neural network does not require frequent retraining, ensuring its applicability in online scenarios such as the sustainable demand response.**

***Keywords*-smart grid; consumption prediction method; neural network;**

I. Introduction

Recent technological advances have brought in households smart meters capable of bidirectional communication and remote control. They are part of the emerging *smart grid* which promises to optimize customer electricity consumption through dynamic price incentives [1], voluntary participation [2] or direct control [3] in a world in which energy costs are constantly rising. This helps utility providers increase the reliability of the power grid, while also benefiting customers by reducing their energy related costs. Demand Response (DR) [4] is a popular technique for consumption curtailment during peak hours by employing one of the three mentioned optimization techniques. Figure 1 shows a typical DR architecture for consumption curtailment, where real-time and historical consumption is used to predict future consumption. To avoid exceeding generation capacity and thus buying energy at high rates from the spot market a DR event may be triggered when consumption is predicted to

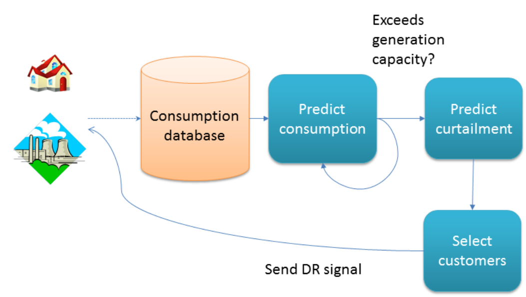

Figure 1: Demand Response architecture.

raise above a certain threshold. Before DR can take place, the potential for curtailment for each customer is analyzed and a subset of customers is selected for participation.

To perform DR, providers need to take into consideration several variables such as customer comfort, willingness to participate, and achievable curtailment during the targeted time interval. Personalized models tailored to the behavior of each customer would be ideal for accurate prediction of electricity consumption in real-time. The scale at which energy consumption data from residential and commercial customers have become available due to the pervasive deployment of smart meters however prohibits so far the development of such personalized solutions. Nonetheless, near real-time[1] (e.g., every 15 minutes) decision making is a requirement for some applications such as sustaining a target reduction throughout the duration of a DR event in smart grids (i.e., sustainable DR).

In the past, several machine learning and data mining

[1]Being distributed environments, smart grids exhibit network latency and processing delays. Hence real-time decision making in smart grids is considered to be in the order of seconds and minutes not milli or microseconds as in classic real-time computing. For this reason when referring to real-time we actually mean near real-time throughout the paper.

approaches have been applied to energy consumption data for prediction purposes. However, such approaches have certain limitations when real-time computational guarantees have to be met. In this paper we take a very pragmatic approach by intensively studying a real-world, large-scale and high resolution electricity consumption dataset, discussing the challenges associated with real-time predictions and evaluating the limitations of existing approaches on this task. In a real-world scenario, the rate at which smart meters produce data varies according to the utility cyber-infrastructure, ranging from seconds to several minutes. As a result, large amounts of continuously flowing data (e.g., a single month of 15-minute sampled electricity consumption data results in $\sim$4 billion data points for Los Angeles customers) need to be processed *fast*. Given the various consumption patterns of individual customers as well as the various methods used for prediction, more historical data may be necessary for training highly accurate models. To solve the volume and velocity aspect of this big data problem there is a need for scalable and agile solutions that require minimal training and that are capable of adapting to changes in real-time.

Our previous studies [2] have shown that prediction accuracy depends among others on costumer type. Since no algorithm is universally better, an exhaustive search among all possible available prediction methods is required to minimize the prediction error. Given real-time constraints, such a solution is unfeasible. The main challenge is therefore to properly identify the best performing prediction model for each individual customer in order to achieve good overall prediction performance, while keeping complexity low for a solution to have real-time applicability.

In this paper we address this issue by proposing an artificial neural network (ANN) based solution which achieves over 84% accuracy in selecting the best electricity consumption prediction method for each customer. To the best of our knowledge this is the first attempt to provide such an automated online method for accurate, real-time prediction of a large number of time series with different characteristics. Furthermore, to address the variability in customer consumption which is difficult to predict with the methods widely used by utility providers, we introduce two historical averaging models that use a sliding window of previous readings to forecast future electricity consumption values. The advantage of our historical averaging methods is their ability to provide accurate predictions with a low computational complexity. For experiments we rely on a representative real-life historical data set comprised of 190 industrial customers with consumption sampled at 15 minute intervals. The data set contains a wide range of consumption patterns as shown in Figure 2. Particularly, Figure 2 demonstrates few examples of electricity consumption (y-axis) variability over time (x-axis) in our dataset.

The main contributions of this paper are as follows:

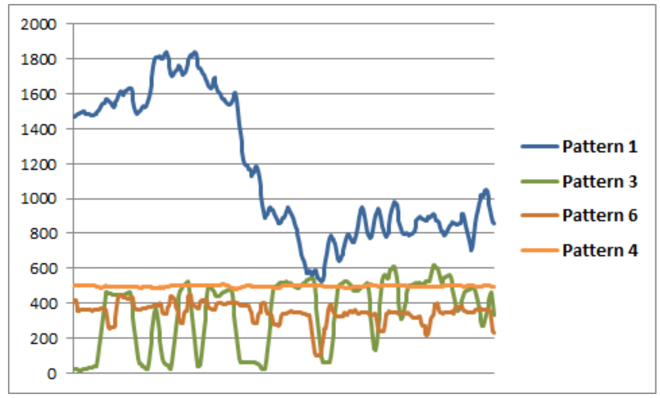

Figure 2: Electricity consumption (y-axis) as a function of time (x-axis).

- We analyze low complexity prediction approaches that are traditionally used by utilities on a real-world dataset. We introduce two historical averaging methods to further reduce dependencies on historical data, thus minimizing complexity.
- We provide a detailed evaluation of our methods on a real-world energy consumption dataset. We show that our proposed methods achieve remarkable improvements in terms of prediction accuracy as compared to traditional energy consumption prediction approaches, especially for customers exhibiting high variability consumption patterns.
- We propose a neural network based method to predict the best consumption prediction method given a small subset of the total customer data set. We show that by using simple information derived from the historical time series we can achieve high accuracy without having to retrain the network often. This makes our method suited for near real time scenarios such as DR.

The rest of the paper is structured as follows: Section II depicts some of the recent results in terms of consumption prediction and analysis of different methods; Section III presents four existing consumption prediction methods as well as the two methods we propose, while Sect. IV outlines the neural network method for prediction the best consumption method. Experimental results are analyzed in Sect. V. Finally conclusions and future work are presented in Sect. VI.

II. Related Work

Research on electricity demand forecasting considers long-term and medium-term prediction for utility planning and maintenance purposes, and short-term forecast for economic scheduling [6]. As utilities move towards sustainable DR, very short-term predictions are required for near real-time control (e.g., predictions 1 hour ahead at 15 minute granularity). Research on energy consumption prediction can be divided into three groups [6]: simple averaging models; statistical models (e.g., regression and time series); and artificial intelligence (e.g., Artificial Neural Networks (ANNs) and pattern matching [7]) approaches. Next, we briefly discuss some recent results in this area.

Averaging models: Utilities and Independent Service Operators (ISOs) use averaging models [8][9][10] based on recent consumption [11], due to their simplicity. Such models make predictions based on linear combinations of consumption values from "similar" days. Some of these methods used in practice are described in Sect. III.

Regression models: Regression models combine several independent features to form a linear function. This helps interpret the relationship between various factors more easily. Our prior work [12] builds regression tree models using weather and schedule data for energy prediction. It evaluates the effect of different feature combinations on the prediction accuracy.

Probabilistic linear regression and Gaussian process regression models for predicting the total kWh consumption as a function of building features, were proposed in [13]. A hybrid method for probabilistic short term load prediction was presented in [14], where a regression tree is used to cluster similar data, and then, a relevance vector machine is constructed for each cluster using Bayesian Inference. Support Vector Machines (SVM) were used for load forecasting in [15]. An on-line least-square SVM based method is introduced by Aung et al. [16]. In their small scale evaluation (only 2 smart meters were used), they argue their method outperforms the least-square SVM based method introduced in [17].

A multiple linear regression model for load prediction, where affecting factors are iteratively analyzed, was presented in [18]. A nonlinear and non-parametric regression model for next day half-hourly load prediction was employed in [19] for stochastic planning and operations decision making. The model contains a combination of maximum, minimum and average demand and temperature from the last 1-hour, 24-hours, and 48-hours. Day of the week, day of the year and holiday effects are also incorporated.

Time series: An overview of time series forecasting approaches for electricity price prediction was presented in [20]. In order to capture the market fundamentals at multiple time granularities (e.g., short, medium, and long-term), the price vector was split into components, which can be separately solved on different time horizons. One of the early reviews for time series based methods for load forecasting was given in [21]. Later, a time series method for short term load forecasting (few hours to few weeks ahead) of hourly loads was proposed in [22]. A comparison of time series methods for load forecasting with other methods was presented in [23]. Seasonal time series were investigated in [24].

Artificial Intelligence (AI) Approaches: Many common AI techniques, such as ANNs, expert systems and pattern matching techniques can be beneficial to demand forecasting [25]. An overview of AI methods applied to short-term electric load forecasting was provided in [26], [27]. Hourly electricity consumption forecasting for day-ahead prediction based on pattern sequence similarity was performed in [7]. An ensemble model based on this work, was later presented in [28].

All aforementioned methods give accurate predictions in specific scenarios regarding the consumption sampling granularity, consumption variation and number of used features.

Integrated software systems: Systems for analyzing, understanding and predicting the consumption behavior of customers have been proposed as well. In [29] a research platform tailored for real-time predictive analytics in a campus microgrid is presented. In [30] a management system for analyzing low latency time series analytics in smart cities is presented.

The difficulty of efficiently selecting customers given the large available pool smart grids operate has been addressed in [31]. In fact a scalable selection procedure combined with data analytics based on an approximate algorithm was proposed to cope with data volume [31].

Classifying Consumption Prediction Methods: The fact that the efficiency of prediction methods varies depending on the consumption pattern has recently gained attention [32]. In the paper the authors proposed a static sequence of preselected methods for various day periods. The major drawback of adopting such a static strategy is the inability to adapt to changes. Our analysis of several consumption datasets [2] including the 190 industrial customers (cf. Sect. V and Fig. 2) has shown that consumption patterns do change over time. As a result the approach we propose in this paper to perform dynamic re-selection of prediction methods when accuracy according to some performance metric as described in Sect. V deteriorates beyond a specified threshold.

III. Consumption Prediction Methods

In this section we briefly present four prediction methods used in our analysis and introduce two historical averaging models. The four methods are used by various utilities in US. While simple, they are preferred over more advanced because of their low compute requirements and the intuitive interpretation of their results.

ARIMA: Auto Regressive Integrated Moving-Average [5]. The ARIMA model predicts future electricity consumption values based on a linear combination of previous, equally spaced univariate time series data. Its advantage lies in the fact that it is simple to use, and that it does not require knowledge of the underlying domain. However, parameter estimation for ARIMA requires human expertise to examine the partial correlogram of the time series. ARIMA has been used to forecast real world time series data such as stock [33] or fuel prices [34], as well as electricity load [35]. In our experiments, ARIMA is trained on a 9 week window of preceding data sampled at 15 minute intervals, to make predictions for the following week.

NYISO: New York ISO [8]. This baseline is calculated from previous five days with the highest average kWh value. These days are chosen from a pool of ten previous days, which are selected starting two days prior to the event day, and excluding weekends, holidays, past DR event days or days on which there was a sharp drop in the energy consumption. In addition a day is included in the pool only if the average consumption on that day is more than 25% of the last selected day. The process repeats until all ten days have been placed in the pool of days for baseline calculation. Days are then ranked based on average hourly consumption and five days with the highest value are selected. Finally, the baseline is calculated by taking hourly averages across these days. For baseline calculation on a DR event day, a morning adjustment factor can also be calculated from the two hour values prior to the DR event by comparing calculated baseline consumption and actual measured data. The value of this adjustment factor cannot be less than 0.8 or more than 1.2 [11].

CASCE: Southern California Edison ISO [10]. This model estimates baseline consumption by averaging past ten days. These days cannot include weekends, holidays or past DR event days. Once ten days have been selected, the baseline is calculated as their hourly average. similar to NYISO, a morning adjustment factor is applied to the calculated baseline.

CAISO: California ISO [9]. According to this model, the baseline is the hourly average of three days with the highest average consumption value among a pool of ten selected previous days. Selected days cannot be weekends, holidays, past DR event days. CAISO's performance can be considerably improved by introducing a morning adjustment factor [11].

A. *Online Sliding Window Consumption Prediction*

Next, we detail an online consumption prediction method that performs averaging on a sliding window of historical data. Our hypothesis is that predictions can become faster and more accurate by employing running historical average values which are updated constantly with real-time data when these become available. To achieve this, we blend the averaging approach of ISO models with the moving average of ARIMA. Particularly, our method constructs an initial training matrix for the entire week made up of one month 15 minute averages totaling 672 values for each customer. Starting from Monday, we make one hour ahead predictions for the upcoming Monday by using a moving average technique. As the actual value for the predicted interval becomes available, the training matrix entry for the interval is updated by a weighted average as follows:

$$\bar{c^t} = \frac{\sum_{i=1}^{n} c_i^t}{n}, \tag{1}$$

where n the number of items in the sample and c_i are weekly measurements at the same time interval t. After the initial $\bar{c^t}$ values are computed for each customer, they need to be updated when a new observation c_{n+1}^t becomes available. The updated average can be computed as:

$$\hat{c^t} = \frac{\bar{c^t} * n * (1-w) + c_{n+1}^t * w}{n+1}, \tag{2}$$

where w is some arbitrary weight. Selecting a weighting strategy can be challenging and may require experimentation. For the purposes of this work, we experimented with numerous weighting schemes. In the end we selected to present two approaches that gave us the best prediction accuracy. Particularly, we derive our **conservative sliding window (CSW)** method by weighting new values by 20% and our **aggressive sliding window (ASW)** method by weighting new values by 50%. Our hypothesis is that CSW captures strong correlations (if any) among consumption values from one week to another. Contrary, ASW assumes recent observations to offer higher predictive power when the prediction horizon is short.

IV. Predicting the Best Consumption Forecasting Method

As explained in Sect. I the main issues that consumption prediction faces in the context of smart grids are the large historical data sets, continuous real-time data, and prediction accuracy that depends on a combination of data set and prediction method properties. In fact, we have shown in prior work [2] that customer characteristics affect the prediction accuracy of electricity forecasting models. Particularly, we concluded in that study that no single model gives the best performance in all cases and different models may perform differently for different customers. To validate this we conducted a pilot experiment to build an *Oracle* model that picks the best performing model for each prediction interval for each customer. The results (cf. Figure 7) show an improvement as compared to individual prediction methods. However, training and predicting the consumption for every customer across all available methods and picking the method that gives the smallest error at each time interval is computationally prohibitive and unfeasible for the real-time demands of sustainable DR.

In this section we propose an ANN forecasting model that selects for each customer the best prediction method to use for future electricity consumption prediction among a predefined set of forecasting models. Our ANN provides a computationally efficient solution that intelligently combines available models without violating the near real-time constraints imposed by the application. Specifically, we propose a feed forward back propagation network [36] that takes as input a vector of the daily consumption standard deviation values for the past 24 days. By using the consumption standard deviation as input to the network we want to capture the variations in user consumption since these are the ones that impact most the efficiency of the prediction methods as shown in our recent study [2]. Our experimental analysis (cf. Sect. V) suggests that our selection results in high prediction accuracy. During training, the ANN learns by comparing the estimated best method to the known actual best method for each customer. The error is fed back to the ANN and is used to adjust the inferred weights in the hidden layers so that the learned output matches the correct output as best as possible.

Figure 3 depicts the architecture behind the neural network method. The training set consists of a randomly picked subset of customers with consumption data for a given historical interval. On this dataset we first run an *Oracle method* which provides the prediction method with the smallest prediction error for each customer. Once the oracle finishes we feed as training set to the network the standard deviation of each customer together with the best method for that customer. After training is over we predict the best method for the remaining customers and periodically check whether or not the accuracy of the selected method decreased below a user defined threshold. If so, the network is retrained using the same random customer data set with updated consumption values for the recent past. From our initial experiments (cf. Sect. V) we observed that a random sample is as good as any other customer selection approach in terms of achieved accuracy. In future work, we plan to perform a careful study regarding the customer sample selection and its effect on overall accuracy.

V. Experiments

We have tested the efficiency of our approach on a real-world dataset consisting of 190 industrial customers from the LA area, each with 9 weeks of 96 daily consumption values, totaling 1,149,120 data points. Consumption data was recorded between May 31, 2012 and June 1, 2013. Figure 4 depicts the daily consumption values (y-axis) for each customer (x-axis) sorted by mean consumption, while Fig. 5 shows the daily consumption standard deviation values (y-axis) of each customer (x-axis). As it can be seen, approximately 1/3 of the customers demonstrate relatively stable consumption, while about 10% exhibit high variability.

To validate our ANN we used a three fold cross validation for training and testing. Particularly, a random subset containing 1/3 of the customers is used for training, and the rest for testing. We implemented our backpropagation network in R using the *neuralnet* package. The number of

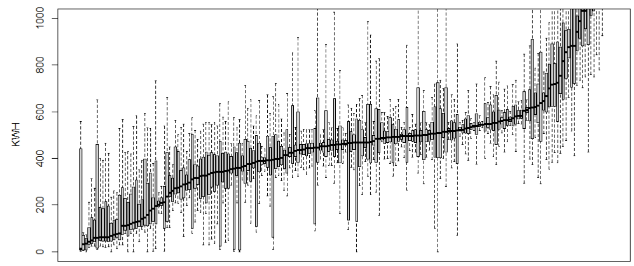

Figure 4: Consumption kWh.

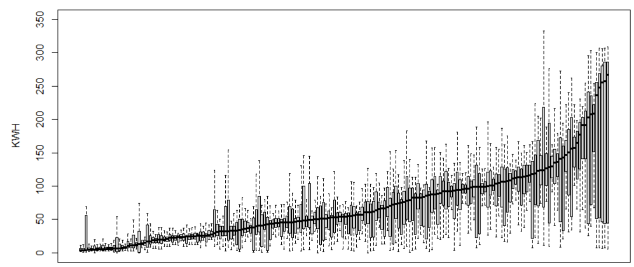

Figure 5: Consumption standard deviation.

hidden neurons was set to 70 and the threshold value for the stopping criterion was set to 0.01.

The performance metric we used to quantify accuracy in the task of electricity consumption prediction is Mean Absolute Percentage Error (MAPE):

$$MAPE = \left(\frac{1}{n} \sum_{i=1}^{n} \left| \frac{a_i - f_i}{a_i} \right| \right) * 100, \quad (3)$$

where a_i and f_i represent the actual and predicted values respectively, and n is the number of predictions. MAPE is a widely used metric in smart grids for determining prediction accuracy.

We compared the prediction accuracy of our neural network with the values given by the Oracle method. Based on the results we have also analyzed which of the 6 methods is the most resilient, i.e., the most versatile to variability in consumption values.

A. Experimental Results

Figure 6 shows the CDF of the MAPE values for all methods. Besides the Oracle method which always produces the best results, we also plot a random selection method according to which the “best” method is selected at random. Clearly, the neural network, which employs a combination of forecasting methods, achieves the best results, i.e., closest

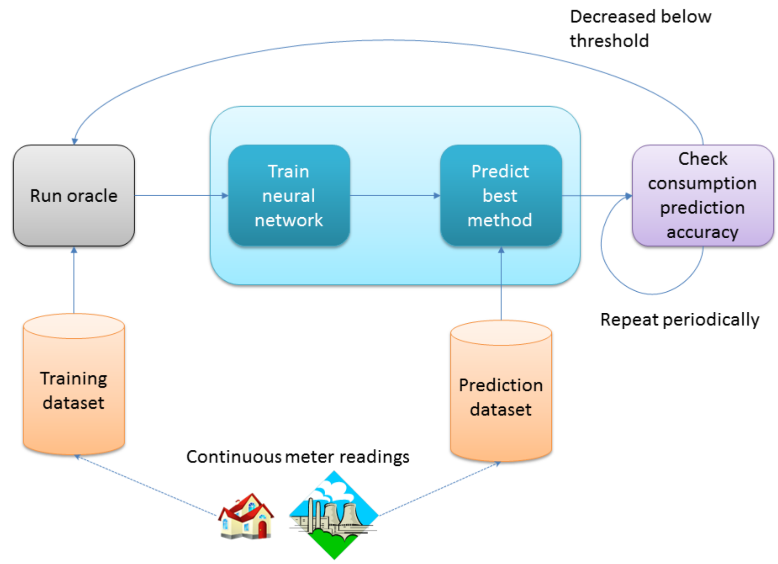

Figure 3: Architecture of the neural network method.

to the Oracle. Specifically, the neural network is able to predict the best method for each customer with an accuracy ranging from 0.8412 to 0.9448 for the three folds. The high accuracy is corroborated by the overall deviation from the best MAPE value, which we calculated to be -0.0318. We define deviation from the best MAPE value as follows:

$$\frac{MAPE_{predictedMethod} - MAPE_{bestMethod}}{MAPE_{bestMethod}} \tag{4}$$

The small deviation from the best MAPE value, combined with the large prediction window of one week for which it was achieved, leads us to conjecture that our proposed method is suitable for online scenarios where frequent retraining is not desirable.

Among individual forecasting methods, we found NYISO to be the worst, and our ASW to be the best. For almost 40% of the customers (customers exhibiting small variations in consumption as shown in Figures 4 and 5) however all methods produce comparable MAPE values. Evidently, each forecasting method is individually performing worse than the ANN approach, thus confirming empirically that the "sum is greater than its parts". Figure 7 shows the percent of times each method is the best versus the percent of times it is selected by the neural network. It can be seen that the dominant method is the aggressive sliding window with 45.79% of the time being the best. This points to the fact that giving a considerable weight to the latest values maximizes the accuracy and exceeds even that of ARIMA (32.11%). Surprisingly, the conservative method is outperformed by the rest of the baselines, even though Fig. 7 suggests that it is performing better in some cases than NYISO and

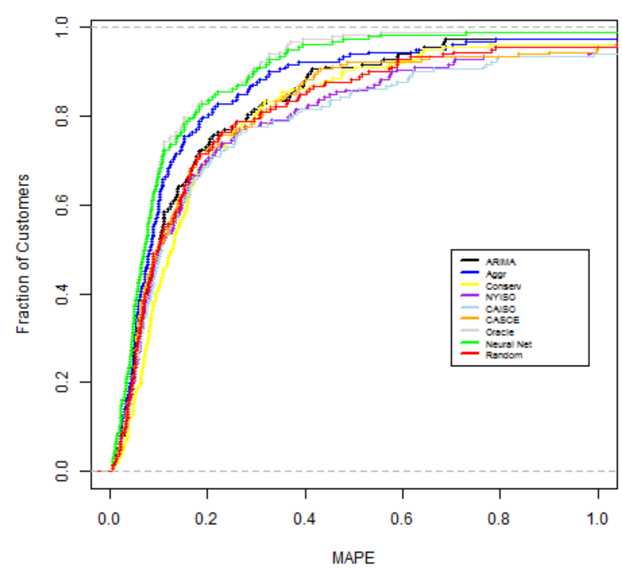

Figure 6: CDF of the MAPE values for all methods.

CASCE. Even though the observation is valid, CSW is always outperformed by some other method. We therefore conclude that weekly patterns are not strong in our dataset. Out of the ISO models, CASCE was dominant with 14.74% MAPE, whereas NYISO was the method which produced the fewest best MAPE values. This suggests that averages of consumption values for past similar days results in more accurate forecasts than averaging over high consumption

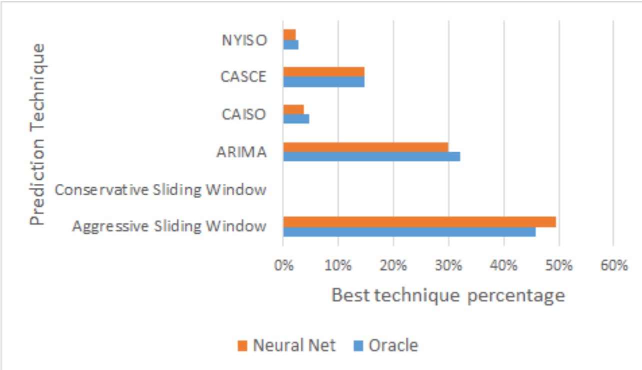

Figure 7: Accuracy of the neural network compared to the oracle.

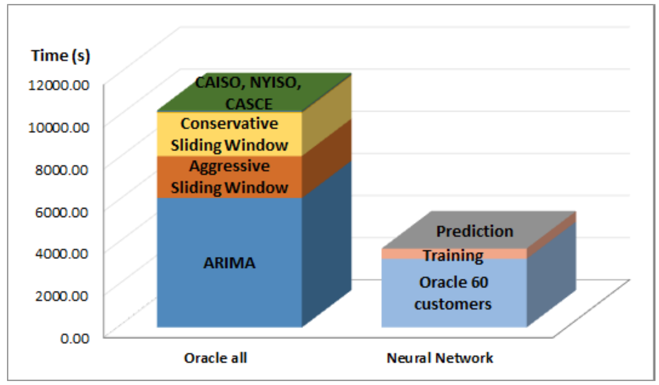

Figure 8: Execution times of the oracle method vs. the neural network.

values.

While our neural network achieves good results overall, it can also result in false positives. For example, as Fig. 7 shows, 3.68% of the time some other model is picked up instead of the ASW method. The confusion between ASW and other models may be an artifact of customer characteristics. Even though the impact on accuracy is small for this dataset, further examination is required to optimize performance.

Execution times[2]. We have compared the time needed to run the Oracle method against the time needed to train our neural network and make predictions. The main disadvantage of the Oracle is that it needs to run all methods for all customers in order to determine the best method per customer at each point in time. For our dataset this process takes about 170 minutes, which is too long for any near real-time decisions. In contrast the neural network takes 8 minutes to train and less than one second to predict on a subset of 60 customers. Compiling a training set for the ANN includes identifying the best method for each customer in the training data. In our experiments we measured this overhead to be $\sim$62 minutes. Figure 8 shows the results. Despite the substantial overhead, given the real-time requirements, we emphasize that the combined cost for running the Oracle on the subset and the training of the ANN is incurred infrequently as demonstrated by the small deviation from the best MAPE value for the one week ahead prediction. Nevertheless, to avoid affecting performance, retraining can be performed asynchronously once deterioration becomes substantial. The small prediction type of under a second makes it ideal for real-time use cases such as sustainable DR as compared to the Oracle method. Even though parallelizing the entire Oracle process by having one process per customer would ideally lead to a duration of $\sim$54 seconds per customer this would still be much longer than the time needed by the ANN to predict once trained.

[2]All experiments were performed on a commodity laptop.

VI. Conclusion

In this work we presented an efficient ANN based method for selecting the best consumption method in smart grids. Our method is suited for such environments where no single prediction method can be used due to different patterns in the time series. In the smart grid context, this is due to various consumption patterns among customers being detected. Our proposed ANN , including training and running the Oracle on the train dataset) is more than three times faster than the Oracle, which even though can result in better predictions is computationally prohibitive. We argued that our approach is applicable to online scenarios, such as sustaining consumption curtailment in smart grids during DR. To further reduce prediction error induced by traditional methods in use by utilities we introduced an online historical averaging method with different weighting schemes. We analyzed its efficiency with respect to the number of times it outperformed other methods in the task of energy consumption forecasting. In future work we plan to investigate the impact of customer sample selection as well as training dataset size on prediction error. We further plan to explore how to incrementally update the ANN without retraining it from scratch when required. Last but not the least, we intend to establish a mechanism which will keep track of prediction error and automatically identify the quickest point in time a retraining will be mandatory.

Acknowledgment

This material is based upon work supported by the United States Department of Energy under Award Number number DE-OE0000192, and the Los Angeles Department of Water and Power (LA DWP). The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof, the LA DWP, nor any of their employees.

REFERENCES

[1] C. Chen, J. Wang, Y. Heo, and S. Kishore, "Mpc-based appliance scheduling for residential building energy management controller," *Smart Grid, IEEE Transactions on*, vol. 4, no. 3, pp. 1401–1410, Sept 2013.

[2] S. Aman, M. Frincu, C. Charalampos, U. Noor, Y. Simmhan, and V. Prasanna, "Empirical comparison of prediction methods for electricity consumption forecasting," University of Southern California, Tech. Rep. 14-942, 2014, uRL: http://www.cs.usc.edu/assets/007/89887.pdf (accessed Mar 6, 2014).

[3] K.-H. Ng and G. Sheble, "Direct load control-a profit-based load management using linear programming," *Power Systems, IEEE Transactions on*, vol. 13, no. 2, pp. 688–694, May 1998.

[4] F. Rahimi and A. Ipakchi, "Demand response as a market resource under the smart grid paradigm," *Smart Grid, IEEE Transactions on*, vol. 1, no. 1, pp. 82–88, 2010.

[5] B. George E. P. and J. Gwilym M., *Time series analysis, forecasting and control*. Holden-Day, 1970.

[6] H. K. Alfares and M. Nazeeruddin, "Electric load forecasting: Literature survey and classification of methods," *International Journal of Systems Science*, vol. 33, no. 1, 2002.

[7] F. Martinez-Alvarez, A. Troncoso, J. Riquelme, and J. A. Ruiz, "Energy time series forecasting based on pattern sequence similarity," *IEEE Transactions on Knowledge and Data Engineering*, vol. 23, no. 8, 2011.

[8] NewYorkISO, "Emergency demand response program manual," Tech. Rep., 2013.

[9] CaliforniaISO, "Caiso demand response resource user guide," Tech. Rep., 2007, uRL: http://www.caiso.com/1ca6/1ca67a5816ee0.pdf (accessed Nov 23, 2013).

[10] SouthernCaliforniaEdison, "10 day average baseline and "day-off" adjustment," Tech. Rep., 2011.

[11] C. Katie, P. Mary Ann, G. Charles, and K. Sila, "Estimating demand response load impacts: Evaluation of baseline load models for non-residential buildings in california," Tech. Rep., 2008, uRL: http://drrc.lbl.gov/sites/drrc.lbl.gov/files/63728.pdf (accessed Nov 23, 2013).

[12] S. Aman, Y. Simmhan, and V. K. Prasanna, "Improving energy use forecast for campus micro-grids using indirect indicators," in *IEEE Workshop on Domain Driven Data Mining*, 2011.

[13] J. Z. Kolter and J. F. Jr, "A large-scale study on predicting and contextualizing building energy usage," in *AAAI Conference on Artificial Intelligence*, 2011.

[14] H. Mori and A. Takahashi, "Hybrid intelligent method of relevant vector machine and regression tree for probabilistic load forecasting," in *IEEE International Conference and Exhibition on Innovative Smart Grid Technologies (ISGT Europe)*, 2011.

[15] B.-J. Chen, M.-W. Chang, and C.-J. Lin, "Load forecasting using Support Vector Machines: a study on EUNITE competition 2001," *IEEE Transactions on Power Systems*, vol. 19, no. 4, pp. 1821–1830, 2004.

[16] Z. Aung, M. Toukhy, J. Williams, A. Sanchez, and S. Herrero, "Towards accurate electricity load forecasting in smart grids," in *DBKDA 2012, The Fourth International Conference on Advances in Databases, Knowledge, and Data Applications*, 2012, pp. 51–57.

[17] L. Ganyun, X. Wang, and Y. Jin, "Short-term load forecasting in power system using least squares support vector machine," in *Computational Intelligence, Theory and Applications*. Springer, 2006, pp. 117–126.

[18] T. Hong, M. Gui, M. Baran, and H. Willis, "Modeling and forecasting hourly electric load by multiple linear regression with interactions," in *Power and Energy Society General Meeting, 2010 IEEE*, 2010, pp. 1–8.

[19] S. Fan and R. Hyndman, "Short-term load forecasting based on a semi-parametric additive model," *IEEE Transactions on Power Systems*, vol. 27, no. 1, 2012.

[20] S. Chan, K. M. Tsui, H. C. Wu, Y. Hou, Y.-C. Wu, and F. Wu, "Load/price forecasting and managing demand response for smart grids: Methodologies and challenges," *Signal Processing Magazine, IEEE*, vol. 29, no. 5, pp. 68–85, 2012.

[21] M. Hagan and S. M. Behr, "The time series approach to short term load forecasting," *IEEE Transactions on Power Systems*, vol. 2, no. 3, 1987.

[22] N. Amjady, "Short-term hourly load forecasting using time-series modeling with peak load estimation capability," *IEEE Transactions on Power Systems*, 2001.

[23] J. W. Taylor, L. M. de Menezes, and P. E. McSharry, "A comparison of univariate methods for forecasting electricity demand up to a day ahead," *International Journal of Forecasting*, vol. 22, pp. 1–16, 2006.

[24] P. S. Kalekar, "Time series forecasting using holt-winters exponential smoothing," *Kanwal Rekhi School of Information Technology*, vol. 4329008, pp. 1–13, 2004.

[25] S. D. Ramchurn, P. Vytelingum, A. Rogers, and N. R. Jennings, "Putting the 'Smarts' Into the Smart Grid: A Grand Challenge for Artificial Intelligence," *Communications of the ACM*, vol. 55, no. 4, 2012.

[26] K. Metaxiotis, A. Kagiannas, D. Askounis, and J. Psarras, "Artificial intelligence in short term electric load forecasting: a state-of-the-art survey for the researcher," *Energy Conversion and Management*, vol. 44, no. 9, 2003.

[27] H. Hippert, C. Pedreira, and R. Souza, "Neural networks for short-term load forecasting: a review and evaluation," *IEEE Transactions on Power Systems*, vol. 16, no. 1, 2001.

[28] W. Shen, V. Babushkin, Z. Aung, and W. L. Woon, "An ensemble model for day-ahead electricity demand time series forecasting," in *International Conference on Future Energy Systems (ACM e-Energy)*, 2013.

[29] N. Balac, T. Sipes, N. Wolter, K. Nunes, B. Sinkovits, and H. Karimabadi, "Large scale predictive analytics for real-time energy management," in *Big Data, 2013 IEEE International Conference on*, 2013, pp. 657–664.

[30] A. Marascu, P. Pompey, E. Bouillet, O. Verscheure, M. Wurst, M. Grund, and P. Cudre-Mauroux, "Mistral: An architecture for low-latency analytics on massive time series," in *Big Data, 2013 IEEE International Conference on*, 2013, pp. 15–21.

[31] J. Kwac and R. Rajagopal, "Demand response targeting using big data analytics," in *Big Data, 2013 IEEE International Conference on*, 2013, pp. 683–690.

[32] A. Marinescu, C. Harris, I. Dusparic, V. Cahill, and S. Clarke, "A hybrid approach to very small scale electrical demand forecasting," in *Innovative Smart Grid Technologies Conference (ISGT), 2014 IEEE PES*, 2014, pp. 1–5.

[33] P.-F. Pai and C.-S. Lin, "A hybrid arima and support vector machines model in stock price forecasting," *Omega*, vol. 33, no. 6, pp. 497–505, December 2005. [Online]. Available: http://ideas.repec.org/a/eee/jomega/v33y2005i6p497-505.html

[34] C. Morana, "A semiparametric approach to short-term oil price forecasting," *Energy Economics*, vol. 23, no. 3, pp. 325–338, May 2001. [Online]. Available: http://ideas.repec.org/a/eee/eneeco/v23y2001i3p325-338.html

[35] J. Contreras, R. Espinola, F. Nogales, and A. Conejo, "Arima models to predict next-day electricity prices," *Power Systems, IEEE Transactions on*, vol. 18, no. 3, pp. 1014–1020, 2003.

[36] S. I. Gallant, *Neural Network Learning and Expert Systems*. Cambridge, MA, USA: MIT Press, 1993.