Prediction Models for Dynamic Demand Response

Saima Aman*, Marc Frincu[y], Charalampos Chelmis[y], Muhammad Noor[y], Yogesh Simmhan[z], Viktor K. Prasanna[y]

*Computer Science Department, University of Southern California, Los Angeles, CA
[y]Electrical Engineering Department, University of Southern California, Los Angeles, CA
[z]Supercomputer Education and Research Centre, Indian Institute of Science, Bangalore, India
Email: saman@usc.edu, frincu@usc.edu, chelmis@usc.edu, mnoor@usc.edu, simmhan@serc.iisc.in, prasanna@usc.edu

Acknowledgement

This material is based upon work supported by the Department of Energy under Award Number DE-OE0000192.

Disclaimer

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, the Los Angeles Department of Water and Power, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Prediction Models for Dynamic Demand Response

Requirements, Challenges, and Insights

Saima Aman[*], Marc Frincu[†], Charalampos Chelmis[†], Muhammad Noor[†], Yogesh Simmhan[‡], Viktor K. Prasanna[†]
[*]Computer Science Department, University of Southern California, Los Angeles, CA
[†]Electrical Engineering Department, University of Southern California, Los Angeles, CA
[‡]Supercomputer Education and Research Centre, Indian Institute of Science, Bangalore, India
Email: saman@usc.edu, frincu@usc.edu, chelmis@usc.edu, mnoor@usc.edu, simmhan@serc.iisc.in, prasanna@usc.edu

Abstract— **As Smart Grids move closer to dynamic curtailment programs, Demand Response (DR) events will become necessary not only on fixed time intervals and weekdays predetermined by static policies, but also during changing decision periods and weekends to react to real-time demand signals. Unique challenges arise in this context vis-a-vis demand prediction and curtailment estimation and the transformation of such tasks into an automated, efficient dynamic demand response (D^2R) process. While existing work has concentrated on increasing the accuracy of prediction models for DR, there is a lack of studies for prediction models for D^2R, which we address in this paper. Our first contribution is the formal definition of D^2R, and the description of its challenges and requirements. Our second contribution is a feasibility analysis of very-short-term prediction of electricity consumption for D^2R over a diverse, large-scale dataset that includes both small residential customers and large buildings. Our third, and major contribution is a set of insights into the predictability of electricity consumption in the context of D^2R. Specifically, we focus on prediction models that can operate at a very small data granularity (here 15-min intervals), for both weekdays and weekends - all conditions that characterize scenarios for D^2R. We find that short-term time series and simple averaging models used by Independent Service Operators and utilities achieve superior prediction accuracy. We also observe that workdays are more predictable than weekends and holiday. Also, smaller customers have large variation in consumption and are less predictable than larger buildings. Key implications of our findings are that better models are required for small customers and for non-workdays, both of which are critical for D^2R. Also, prediction models require just few days' worth of data indicating that small amounts of historical training data can be used to make reliable predictions, simplifying the complexity of big data challenge associated with D^2R.**

I. Introduction

Electricity consumption optimization is critical to enhance electric grid reliability and to avoid supply-demand mismatches. Utilities have long used demand response (DR) for achieving customer-driven curtailment during peak demand periods to maintain reliability [1]. Traditionally, planning and notification for DR is done a day ahead of the day when curtailment is to be performed [2]. However, the Smart Grid is transitioning towards dynamic demand response [3], in which the utility provider needs to perform DR at a few hours' advance notice whenever necessitated by dynamically changing conditions of the grid. We formally define Dynamic Demand Response as follows:

TABLE I: Comparison of DR and D^2R characteristics and challenges

	DR	D^2R
Goal	advance planning	dynamic adaptation
Horizon	day ahead	hours ahead
Data/control granularity	coarse	fine
Data rate	monthly billing	real-time data from smart meters
Timing and duration	fixed and pre-defined	flexible, dynamically determined
Extent of curtailment	fixed	dynamically determined/adjustable
Customer selection	selected a-priori	dynamically selected
Challenges	labor intensive, data unavailability, inability to adapt	small latency requirements, computational complexity, data deluge

Definition 1: Dynamic Demand Response (D^2R) is the process of balancing supply and demand in real-time and adapting to dynamically changing conditions by automating and transforming the demand response planning process.

Several factors drive the transition towards D^2R; most notable the integration of renewable energy sources, which due to their intermittent, non-dispatchable, and uncertain nature result in supply instability [2]. The need to curtail at time periods which were traditionally considered non-peak periods, such as weekends, as a result of such instabilities is beyond existing DR policies. The need to curtail any time as a result of such instabilities is beyond existing DR policies, which are traditionally defined for workdays, and usually in hot summer afternoons [4], [1]. Besides, Plug-in electric vehicles (PEVs) can introduce spikes in consumption at arbitrary times during the course of a day [5], whereas special events can result in increased load on weekends. The key differences between DR and D^2R are summarized in Table I.

In DR, the focus has been on large industrial and commercial customers [2], selected a-priori, who are expected to contribute large-sized curtailment. With increasing adoption of smart meters [6], [3], and home energy management and automation systems [7], [2] however, the participation of small customers in demand side management is increasing. The electricity demand of such small customers might be easier to regulate (i.e. shift or shave) as compared to the load of commercial entities, however, consumption prediction for small customers and at high temporal granularity is challenging [8], [9]. One of the key implications of involving small customers would be to dynamically and optimally select customers for participation in curtailment [2] and request only the minimum curtailment required to avoid fatigue and loss of interest in the customers [1].

While existing work has focused on improving consumption prediction models[1], [9], to date, there has been little study on the differences in consumption characteristics of various customer types and their impact on prediction models' accuracy. Existing studies have shown that consumption prediction accuracy is high when consumption values of individual customers are aggregated together [6], [10], [11]. This is attributed to the law of large numbers such that larger the number of customers in an aggregated group, the lower the prediction error for the group [12]. Predictions for aggregated groups make it impossible to discover curtailment potential of individual customers, which is necessary for wider adoption of D^{2}R. Models that work well for large commercial customers with smaller consumption variability over time, could be less efficient for small residential customers, whose consumption pattern fluctuates significantly. Thus, it is necessary to identify effective methods for predicting demand for diverse customers.

In this paper, we compare six short-term electricity consumption prediction models. Our study differs from previous work and focus specially to D^{2}R challenges, in that:

- it deals with small, highly variable, individual customer consumption, as well as relatively larger and more stable building consumption;
- consumption data granularity is very small (i.e., 15-min interval) for appropriately timing the requests for dynamic demand response (D^{2}R) [1], as opposed to prior work on hourly or higher granularity predictions;
- it focuses on short-term predictions (hours ahead) required for (D^{2}R) [1], as opposed to most prior work on day-ahead predictions;
- it evaluates the relationship between prediction accuracy and day type, i.e., workday versus weekends or holidays.

These distinctions make the insights we draw greatly useful for researchers and practitioners in the smart grid domain. Our goal is to get a comprehensive understanding of the performance of prediction models for D^{2}R. *Prediction models used for D^{2}R should balance conflicting requirements of* ***high prediction accuracy****,* ***low compute time for training and prediction****, and* ***reliability*** *at any time of the week and for diverse customers.* This paper can be considered as first attempt in studying prediction models specifically from this perspective.

II. Prediction Models for D^{2}R

Our work advances previous research on analyzing performance of prediction models, such as [8], [2], [13], but differs from them in experiments and analysis focused specifically on D^{2}R. Electricity consumption prediction models can be broadly categorized into three groups [14]: 1) simple averaging models; 2) statistical models like regression and time series models; and 3) artificial intelligence and machine learning models (AI/ML) like neural networks and support vector machines [14] [10], [15], [8]. As many AI/ML methods involve longer training times, they may not be suitable for near real time predictions required for D^{2}R, and hence not considered here.

In the following, we describe the models considered in our study. While the models used in our analysis are not exhaustive, they represent the most commonly used algorithms in demand-response systems [1], [8] and meet the requirements for D^{2}R predictions as mentioned previously.

A. Averaging Models

Averaging models are popular among utilities and ISOs [16][17][18] due to their simplicity [19]. Averaging models make predictions based on linear combinations of consumption values from limited historical data. Averaging models have been shown to perform as well as advanced machine learning and time-series models [8] while considerably reducing the computational need for large-scale predictive analysis of home energy data. In our study, we consider three popular averaging models and a Time of Week (ToW) model, as described below:

1) New York ISO Model (NYISO): It predicts for the next day by taking hourly averages of the five days with highest average consumption value among a pool of ten previous days, starting from two days prior to prediction [16]. It excludes data from weekends, holidays, past DR event days or days with sharp drop in the energy consumption.

2) California ISO Model (CAISO): It predicts for the next day by taking hourly averages of the three days with highest average consumption value among a pool of ten previous days, excluding weekends, holidays, and past DR event days [17].

3) Southern California Edison Model (CASCE): It predicts for the next day by taking hourly averages across past ten immediate or similar days, excluding weekends, holidays, and past DR event days [19], [18].

4) Time of Week Average Model (ToW) : It predicts for each 15-min interval in a week by taking average over all weeks in the training dataset. It captures consumption variations over the duration of a day , i.e., from day to night, and across different days of the week. Time related features are important for electricity consumption [20] as it is closely tied to human schedules and activities.

B. Regression Models

Regression models combine several independent features to form a linear function. Commonly used regression models for electricity consumption prediction are regression tree models [21], probabilistic linear regression, and gaussian process regression models [22]. Hybrid methods that combine regression-based models with other models have also been used for short term load prediction [23]. A multiple linear regression model for load prediction was presented in [26]. A non-linear and non-parametric regression model for next day half-hourly load prediction was employed in [27] for stochastic planning and operations decision making. In other studies, Support Vector Machines have also been used for load forecasting [24], [25].

We use regression trees [28] in this study. A regression tree recursively partitions data into smaller regions until each region can be represented by a constant or a linear regression model. Its key advantage is its flowchart or tree representation that enables domain users to interpret the impact of different features on predicted values [21]: Also, once trained, predictions are fast to compute by a tree look-up [1].

C. Time Series Models

A Time Series model predicts future values based on recent observations. One of the early reviews for time series based models for load forecasting is given in [29]. A comparison of time series methods for load forecasting with other methods is presented in [30]. A time series method for short to medium term load forecasting (few hours to few weeks ahead) of hourly loads was proposed in [31].

In this study, we use Auto-Regressive Integrated Moving Average (ARIMA) [29]. ARIMA is defined in terms of three parameters: d, the number of times a time series needs to be differenced to make it stationary; p, the auto-regressive order, that denotes the number of past observations included in the model; and q, the moving average order that denotes the number of past white noise error terms included in the model. These parameters are derived from the Box-Jenkins test [32].

III. Experimental Setup

A. Dataset Description

Electricity consumption data: Our data for small customers is drawn from a major California power utility. It comprises of 15-min kWh values from 89 household customers, collected between Feb 2013 and Apr 2013. The data[1] for the building-level large customers comes from USC campus microgrid [33], [3]. It comprises of 15-min kWh values from 170 USC campus buildings, collected between Jul 2009 and Jun 2013 [33], [3]. It represents large customers of diverse type: teaching and office spaces, residential, and administrative buildings. For both datasets, we excluded customers with major discontinuities in data, and used linear interpolation for minor gaps. Key properties of the datasets are summarized in Table II, and their distribution is shown in Figure 1. For more details on the datasets, the readers are referred to [34].

Weather data: We obtained curated weather data from NOAA [35], [36] We used hourly temperature observations, which were interpolated to 15-min values.

Schedule Data: It was obtained for the campus dataset comprised of information on working days, holidays, and semester durations (for campus dataset). We used this information to compare the performance of workday versus non-workday performance of the models.

[1]Available from the USC Facility Management Services (FMS).

TABLE II: Description of microgrid and utility datasets.

	Campus Microgrid	Utility Dataset
Number of participants	170	89
Data collection period	4 years	3 months
Data points	4 years $\times$ 365 days $\times$ 96 intervals $\approx 140 * 10^3$ points per building $\approx 16 * 10^6$ points total	3 months $\times$ 30 days $\times$ 96 intervals $\approx 8.5 * 10^3$ points per customer $\approx 0.73 * 10^6$ points total
Client type	buildings	households
Mean consumption (kWh)	large (30.52±7.65)	small (0.22±0.15)
Average variance (kWh)	122.56	0.026

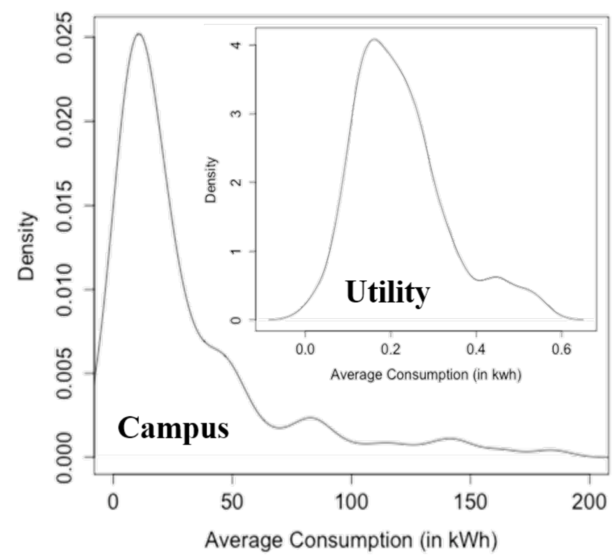


Fig. 1: Probability density function (PDF) of average kWh consumption per 15-min interval of campus buildings. Embedded: the PDF of utility area customers.

B. Prediction Models' Configuration

For the averaging models and regression tree models, we spilt both datasets in 2:1 where 2 parts were used for training and 1 part for testing. For both datasets, we build one prediction model per customer. For the regression tree models, we selected the feature combination that offered the best prediction accuracy based on our previous work [21]: day of week, semester, temperature, and holiday/working day flag. The time series models are trained using a sliding window of 8 weeks preceding the prediction period to predict for three horizons: 1, 4, and 24 hours. For the time series ARIMA model, the parameters were found to be $(8, 1, 8)$ for the campus dataset and $(4, 1, 4)$ for the utility dataset.The prediction models' accuracy was compared using the Mean Absolute Percentage Error (MAPE) [1].

IV. Performance Analysis

Observation 1: **Prediction accuracy is higher for customers with high consumption.** We compare how accuracy varies with customer size, which is defined in terms of the average consumption value in a 15-min interval. Due to space

TABLE III: Average MAPE (with standard deviation) for TS-1hr for groups of buildings/customers.

	avg. kWh	workdays	all days
Utility	kWh ≤ 5	0.3075 ± 0.1309	0.3054 ± 0.1269
Campus	kWh ≤ 5	0.1204 ± 0.0649	0.1186 ± 0.0623
	$5 <$ kWh ≤ 15	0.0743 ± 0.0382	0.0750 ± 0.0355
	$15 <$ kWh ≤ 50	0.0610 ± 0.0308	0.0617 ± 0.0293
	kWh > 50	0.0392 ± 0.0147	0.0416 ± 0.0199

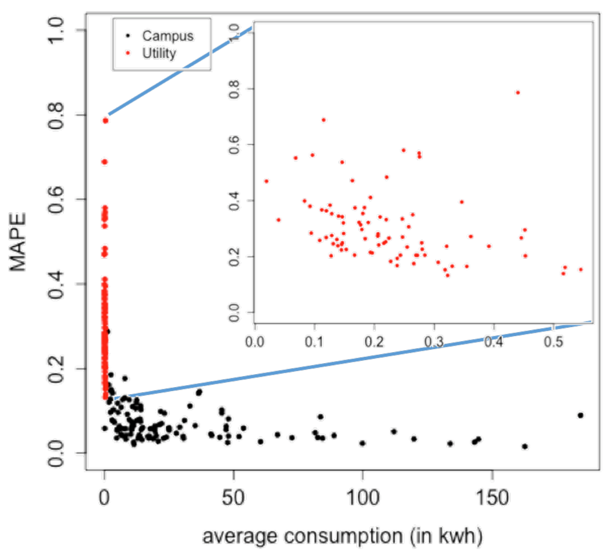


Fig. 2: TS-1hr MAPE as a function of average kWh.

limitations we report here results for the Time Series (1-hr) model in Fig 2. Results for the rest of the models are available at [34]. Higher errors for small sized customers can be observed in both datasets. TS-1hr, the best performing method achieves 30.74% average MAPE (Figure 3c), whereas, CASCE, the best performing averaging method, is far worse, with average MAPE 45.91% for utility customers. We further quantify the relationship of customer size and prediction error by dividing campus buildings in four groups according to average consumption and calculating average MAPE for each group. Table III summarizes the results. Evidently, the average MAPE drops significantly with consumption from 12.04% for smaller buildings to 3.92% for large consumers. This result corroborates previous observations of higher accuracy for larger aggregated consumption prediction [6], [10], [11].

The discrepancy in prediction accuracy between small and large customers can be explained by two factors. First, there is higher variability in small customers [34], i.e., households, where even switching on or off a light bulb can cause a noticeable change in consumption. Second, activities in campus buildings are expected to be periodic, governed by pre-defined schedules and hence expected to be less variable. However, for smaller customers, even a small offset in the predicted kWh value results in a higher percentage error value as a result of the difference between predicted and actual value being high proportionally to large customers.

Insight 1: *D^2R requires higher accuracy models for small customers.*

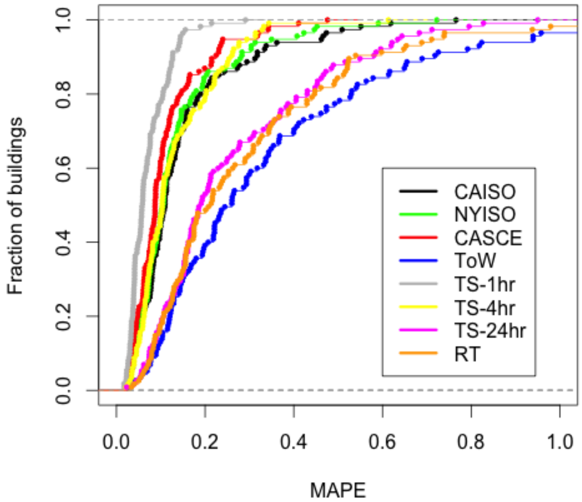
Observation 2: **Few recent observations are better predictors than large sets of historical observations.** Averaging models perform well while using only a small set of recent historical data (i.e., 2-3 weeks). We found CASCE (Figure 3) to be particularly effective for workdays, while ARIMA achieves the best performance while only requiring few training data. According to Figure 3, even simple heuristics such as the Time of Week model perform reasonably well with more recent data (i.e., ToW performs better when trained on data spanning 2 months than when using data covering a period of 2 years). We conclude that historical data of extended time spans enclose consumption patterns that change over time introducing "noise" and deteriorating prediction accuracy. Other researchers have also found that increasing the training data did not improve accuracy [8].

Insight 2: *Prediction models for D^2R relying on few data can maintain high short-term prediction accuracy while significantly reducing storage requirements and computational complexity associated with training and latency (i.e., predictions can be made in real-time).* It also implies that reliable predictions for new buildings or customers can be initiated sooner without waiting to accumulate large training data.

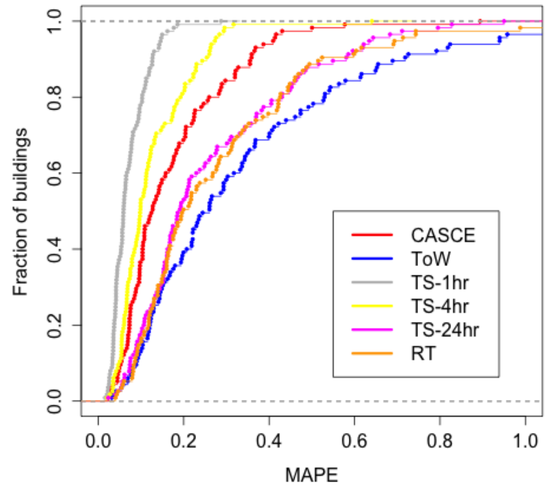
Observation 3: **Simple averaging models are inadequate for DR during weekends.** In our experiments, we evaluated models' performance with respect to all days versus just workdays (Figure 3b). For workdays, the three ISO/utility models achieve lower than 20% MAPE for over 80% of the campus buildings. CASCE is the best among them, with an average MAPE of 10.93%. However, when including all days, CASCE's performance is affected the most. Its average MAPE increases from 10.93% to 17.29%, indicated by a shift of the CDF line to the right in Figure 3d. For experiments involving weekends, we used a modified version of CASCE, which was trained on all days of the week. The degradation in CASCE's performance when including weekends can be attributed to weekend loads being different than weekdays for both campus buildings and utility datasets due to different schedules. Contrary, ARIMA's accuracy deteriorates only slightly for weekends. For TS 1-hour model, MAPE increases by 1.13% (from 7.05% to 7.13%). Time series model benefits from temporal locality and thus does not distinguish between workdays and weekends. The regression tree and time of week models are unaffected as both models inherently capture the workday/weekend information. Specifically, our regression tree model uses day of week and workday/holidays as features, whereas time of week prediction is done by taking averages individually for each day of the week.

Insight 3: *D^2R requires the development of accurate models for all days of the week.* Traditional DR involved industrial customers [2] with weekends considered as non-peak. Instead D^2R can be initiated at any time involving both industrial and residential customers.

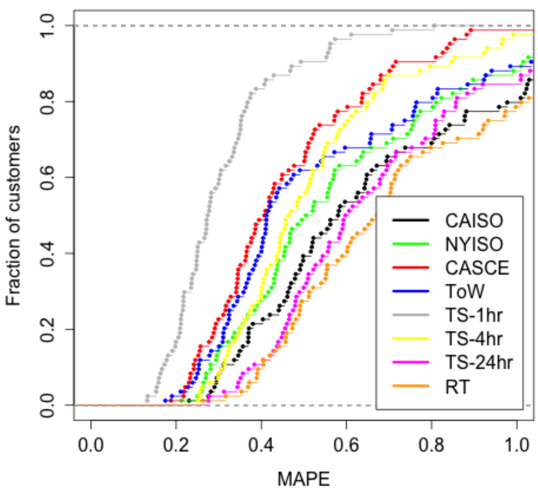
Observation 4: **ARIMA achieves the best prediction accuracy for very-short-term predictions.** For both datasets, the time series 1-hour model achieves the best performance. It's accuracy however deteriorates for longer horizons. While



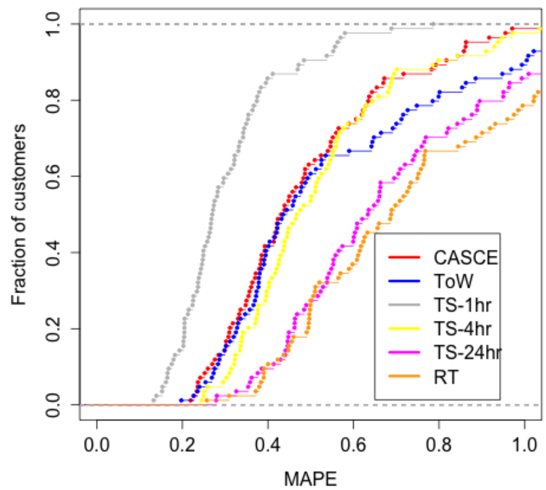
(a) Campus dataset – Workdays



(b) Campus dataset – All days



(c) Utility dataset – Workdays



(d) Utility dataset – All days

Fig. 3: CDF of MAPE values for campus and utility datasets.

this result is well known [1], the confirmation in the context of near-real-time consumption prediction makes it a good candidate for D^{2}R. For campus buildings, average MAPE increases from $7.05 \pm 4.4\%$ to 13.05% for 4-hour horizon, and down to 26.38% for 24-hour horizon for workdays (Figure 3a). For utility customers, TS 1-hour outperforms other prediction techniques for workdays, achieving MAPE below 30% for 80% of the customers (Figure 3c). Combining autoregression with moving averaging, ARIMA can approximate temporal locality in electricity consumption. However, as the prediction horizon increases so does the volatility in the consumption time series data. Therefore, continuous re-training of ARIMA models is necessary for high accuracy to be maintained.

Insight 4: *While TS 1-hour provides best results, its higher training cost [1] makes it problematic for D^{2}R.*

Observation 5: **Models that try to capture global patterns over long time periods are not suitable for D^{2}R.** We found both the regression tree and Time of the Week models to be ineffective for D^{2}R, even though we have demonstrated their usefulness for medium and long-term predictions previously [21], [1]. Also, the results indicate that using additional features, as in regression tree model, did not improve prediction performance.

Insight 5: *Regression tree model is not suited for short term prediction required in D^{2}R, though it has been found useful for medium and long-term predictions [21], [1].*

V. Conclusions

We described how dynamic demand response (D^{2}R) requires very-short-term consumption prediction to make real-time adaptive decisions about curtailment. Prediction models used for D^{2}R should balance conflicting requirements of high prediction accuracy, low compute time for training and prediction, and reliability at any time of the week and for diverse customers. We analyzed six prediction models leading to key insights relevant for D^{2}R. 1) Our results indicate that there is an inherent randomness associated with small customers, which makes it harder to reliably predict their energy consumption compared to larger customers. Thus, D^2R requires higher accuracy models for small customers. 2) Prediction models for D^2R relying on few data can maintain high short-term prediction accuracy while significantly reducing storage requirements and computational complexity associated with training and latency. 3) D^2R requires the development of accurate models for all days of the week. 4) While Time Series 1-hour provides best results, its higher training cost makes it problematic for D^2R. 5) Regression tree model is not suited for short term prediction required in D^2R, though it has been found useful for medium and long-term predictions. For future energy management systems, researchers need to design better models, personalized for individual customers that leverage big data available in D^2R environments to overcome inherent randomness in consumption profiles of individual customers.

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] S. Aman, Y. Simmhan, and V. K. Prasanna, "Holistic measures for evaluating prediction models in smart grids," *IEEE Transactions on Knowledge and Data Engineering (to appear)*, 2014.

[2] H. Ziekow, C. Goebel, J. Struker, and H.-A. Jacobsen, "The potential of smart home sensors in forecasting household electricity demand," in *IEEE International Conference on Smart Grid Communications*, 2013.

[3] Y. Simmhan, S. Aman, A. Kumbhare, R. Liu, S. Stevens, Q. Zhou, and V. Prasanna, "Cloud-based software platform for data-driven smart grid management," *IEEE/AIP Computing in Science and Engineering*, 2013.

[4] S. Park, S. Ryu, Y. Choi, and H. Kim, "A framework for baseline load estimation in demand response: Data mining approach," in *IEEE International Conference on Smart Grid Communications*, 2014.

[5] K. Clement-Nyns, E. Haesen, and J. Driesen, "The impact of charging plug-in hybrid electric vehicles on a residential distribution grid," *IEEE Transactions on Power Systems*, vol. 25, no. 1, 2010.

[6] Y. Simmhan and M. Noor, "Scalable prediction of energy consumption using incremental time series clustering," in *Workshop on Big Data and Smarter Cities, 2013 IEEE International Conference on Big Data*, 2013.

[7] S. Aman, Y. Simmhan, and V. Prasanna, "Energy Management Systems: State of the Art and Emerging Trends," *IEEE Communications Magazine, Ultimate Technologies and Advances for Future Smart Grid (UTASG)*, 2013.

[8] D. Lachut, N. Banerjee, and S. Rollins, "Predictability of energy use in homes," in *International Green Computing Conference*, 2014.

[9] H. Y. Noh and R. Rajagopal, "Data-driven forecasting algorithms for building energy consumption," in *SPIE 8692, Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems*, 2013.

[10] 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.

[11] 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.

[12] T. Hossa, A. Filipowska, and K. Fabisz, "The comparison of medium-term energy demand forecasting methods for the need of microgrid management," in *IEEE International Conference on Smart Grid Communications*, 2014.

[13] A. Veit, C. Goebel, R. Tidke, C. Doblander, and H.-A. Jacobsen, "Household electricity demand forecasting - benchmarking state-of-the-art methods," in *ACM e-Energy Conference*, 2014.

[14] 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.

[15] 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.

[16] "Emergency Demand Response Program Manual, Sec 5.2: Calculation of Customer Baseline Load (CBL)," New York Independent System Operator (NYISO), Tech. Rep., 2010.

[17] "CAISO Demand Response User Guide, Guide to Participation in MRTU Release 1, Version 3.0," California Independent System Operator (CAISO), Tech. Rep., 2007.

[18] "10-Day Average Baseline and Day-Of Adjustment," Southern California Edison, Tech. Rep., 2011.

[19] K. Coughlin, M. A. Piette, C. Goldman, and S. Kiliccote, "Estimating demand response load impacts: Evaluation of baseline load models for non-residential buildings in California," Lawrence Berkeley National Lab, Tech. Rep. LBNL-63728, 2008.

[20] E. A. Feinberg and D. Genethliou, "Load forecasting," *Applied Mathematics for Restructured Electric Power Systems: Optimization, Control, and Computational Intelligence*, pp. 269–285, 2005.

[21] 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.

[22] 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.

[23] 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.

[24] 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.

[25] 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.

[26] 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.

[27] 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.

[28] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and Regression Trees*. Chapman and Hall, 1984.

[29] 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.

[30] 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.

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

[32] G. E. P. Box and G. M. Jenkins, *Time series analysis, forecasting and control*. Holden-Day, 1970.

[33] Y. Simmhan, V. Prasanna, S. Aman, S. Natarajan, W. Yin, and Q. Zhou, "Towards data-driven demand-response optimization in a campus microgrid," in *ACM Workshop On Embedded Sensing Systems For Energy-Efficiency In Buildings (BuildSys)*, 2011.

[34] S. Aman, M. Frincu, C. Chelmis, M. U. Noor, Y. Simmhan, and V. Prasanna, "Empirical Comparison of Prediction Methods for Electricity Consumption Forecasting," *CS Department Technical Report 14-942, University of Southern California*, 2014.

[35] NOAA, "Quality Controlled Local Climatological Data Improvements/Differences/Updates." 2013. [Online]. Available: http://cdo.ncdc.noaa.gov/qclcd/qclcdimprovements.pdf

[36] N. Addy, S. Kiliccote, J. Mathieu, and D. S. Callaway, "Understanding the effect of baseline modeling implementation choices on analysis of demand response performance," in *ASME 2012 International Mechanical Engineering Congress and Exposition*, 2013.