On Hourly Home Peak Load Prediction

Rayman Preet Singh, Peter Xiang Gao and Daniel J. Lizotte
School of Computer Science
University of Waterloo
{rmmathar, x33gao, dlizotte}@uwaterloo.ca

***Abstract*—The Ontario electrical grid is sized to meet peak electricity load. A reduction in peak load would allow deferring large infrastructural costs of additional power plants, thereby lowering generation cost and electricity prices. Proposed solutions for peak load reduction include *demand response* and *storage*. Both these solutions require accurate prediction of a home's peak and mean load. Existing work has focused only on mean load prediction. We find that these methods exhibit high error when predicting peak load. Moreover, a home's historic peak load and occupancy is a better predictor of peak load than observable physical characteristics such as temperature and season. We explore the use of *Seasonal Auto Regressive Moving Average (SARMA)* for peak load prediction and find that it has 30% lower root mean square error than best known prior methods.**

I. Introduction

Peak load is the highest aggregate demand (or *load*) for electric power in a certain area within a certain time frame. Generally, the peak load is observed for a relatively small duration of time. In 2009, the annual peak load in Ontario, Canada occurred less than 1% of the time [2]. To meet this peak load, additional power plants have to be put into operation. These are more flexible but less efficient than conventional plants, and bear higher infrastructural and carbon costs. Hence if the peak load were reduced, it would lower the cost of generating electricity and its prices. The Government of Ontario plans to spend $12 billion over the next 20 years on peak load reduction [3].

In Ontario, the peak load of the residential sector is the largest contributor to the province's total peak load [5]. The adoption of electric vehicles by residential consumers threatens to further increase the peak load [7]. Consequently, aided by wide-scale smart meter deployment, the following methods are being proposed to reduce peak load. *Distributed storage* involves residential consumers storing electricity using in-home batteries during off-peak periods and consuming it during peak periods [9], [23]. In another method, known as *demand-response*, consumers' appliances are made to change their time of operation (e.g. defer use of dishwashers to off-peak periods). However, to be useful, both these approaches require an accurate model to predict each home's load. Further, as most grid operations such as generation scheduling and pricing are conducted on an hourly scale [25], a model for predicting home load on a timescale of an hour or less is required.

To develop a parsimonious predictive model for home load, we investigate the use of machine learning techniques. Our goal is to *build a model that uses observable characteristics such as a home's past load values, day, time, season and weather associated with their occurrence, to predict the next hour's peak load*. No prior work has considered prediction of hourly home peak load.

We make the following key contributions:

- We quantify the effectiveness of *regression-based* techniques in predicting hourly home peak load.
- We demonstrate that for hourly prediction *history-based* observables are more significant than *physical* observables such as temperature.
- We show that *Seasonal Auto-Regressive Moving Average (SARMA)* can be used to model both the intrinsic load pattern and consumer activity in a home, and that it incurs 30% lower error than regression-based techniques.

In Section II we provide relevant background and an overview of related work, followed by a description of our dataset and methodology in Section III and IV. We evaluate our approach in Section V followed by a discussion of its implications in Section VI, and conclusion in Section VII.

II. Background, Motivation and Related Work

A. Background

Energy consumption of a home is driven by consumers' activity. Peak load of a home bounds its energy consumption (or *mean load*) for any given timescale. Thus applications like demand response can use peak load prediction to better understand and limit demand without affecting consumer comfort. Similarly, if peak demand can be predicted, home storage can be charged/discharged intelligently, leading to better utilization.

Microgrid initiatives aim to build a sustainable subsystem for energy consumption using distributed sources like in-home wind turbines, solar panels, and batteries, in conjunction with conventional *generators*. A home peak load prediction method can then be used for automated coordination amongst these different sources. These wide ranging applications motivate us to build a home peak load prediction mechanism.

The 95th percentile, denoted p_{95}, is widely used to measure *capacity* in various applications such as transformer sizing, Internet bandwidth calculations, and ISP billings. Hence we define the hourly peak load y_t of a home as the 95th percentile of all measured load values during the hour t.

B. Applications of Peak Load Prediction

We now elaborate on the potential applications of home peak load prediction and their significance.

Demand Response and Storage Demand response involves peak load reduction by changing consumption patterns by varying pricing policies, or by transferring control of home appliances to the grid. Examples include *time of use pricing* [4] and the **PeakSaver**[1] program. Hence, the grid can employ home peak load prediction to understand each home's future peak load and use it for dynamic pricing and scheduling appliance loads.

Storage aims to reduce peak load by allowing homes to consume stored energy during peak hours. Using home peak load prediction, energy can be stored and consumed in a way that serves the grid's interests while maximizing consumer comfort.

Microgrids In contrast to the centralized grid, microgrids are localized groups of generation, storage and loads, predominantly employing renewables with on-site generators as backup. Examples include Fort Bragg[2], NC. Since the generating capacity in micro grids is limited, peak load prediction can be used to store energy whenever a larger peak is anticipated the next day.

Abnormal Consumption Detection Abnormal events are the leading cause of wasteful energy consumption in a home [12]. For example, the oven is left turned on after dinner, or the refrigerator door is left open. Home peak load prediction in coordination with machine learning methods can be used to detect such events. For instance, whenever the actual peak load exceeds the predicted peak, the consumer can be alerted and her response can be used to train the system.

C. Related Work

Previous research has focused on predicting energy consumption or mean load of a geographical region, city block or a commercial building [11], [14], [19], [20], [24], [27]. Likewise, researchers have studied peak load prediction for geographical areas or cities. The problem of predicting peak load on a relatively small scale such as a home, has largely been overlooked as its applications are only now beginning to appear. Similarly, peak load prediction on a relatively large time-scale such as a day, week or year has been explored by previous work [14]–[17], [19] but such mechanisms do not suffice for the applications discussed above.

Weather, temperature, and seasonality have been shown to significantly affect energy consumption of a region [19], [24], which motivates us to explore its effect on hourly peak load. Edwards et al. [13] investigate the accuracy of various machine learning techniques including *Support Vector Regression (SVR)*, *Least Squares Support Vector Machines (LS-SVM)*, and *Artificial Neural Networks (ANN)* for predicting hourly home energy consumption (mean load). Edwards et al. report that SVR and LS-SVR are most accurate, but they do not reveal the features used. This motivates us to explore possible features for hourly peak load prediction.

[1] https://www.peaksaver.com/

[2] http://www.army.mil/article/68234/

III. Dataset

Our dataset comprises aggregate load values of 24 homes in the Kitchener-Waterloo area, measured for a period of one year (February 2011-January 2012), sampled every six seconds. The Current Cost Envi Device [1] is used for measurement. The measured values are stored temporarily on a netbook and uploaded daily to our data collection server. We record ambient air temperature values published by the University of Waterloo Weather Station [6] and compute hourly average values.

IV. Prediction Models

We now the describe the two types of prediction models that we employ: *Regression* based and *Time Series* based.

A. Regression

We define a set of 5 relevant features of each measured peak load value y_t, to capture different properties that we believe drive peak load. Their motivation, physical meaning and definition are as follows:

1) Time of Day (x_t^1): In a home, consumers' occupancy and activities typically follows an underlying routine. For instance, typical consumers cook food during morning and evening hours, use air conditioning in the afternoon, and run their dishwasher in the evening. This intrinsic pattern is likely to repeat across different days. In relation to hourly peak load y_t, we define time of day (x_t^1) as a feature, where $x_t^1 \in \{1, 2 \dots 24\}$.
2) Day of week (x_t^2): Consumer occupancy and activity patterns vary vastly on weekend days (Saturday, Sunday) as compared to weekdays (Monday- Friday). Hence, in addition to considering the time t of the peak load y_t, we define the day of the week (x_t^2) as a feature, where $x_t^2 \in \{1, 2, \dots 7\}$.
3) Ambient temperature (x_t^3): Weather and seasonality has been shown to affect energy consumption and has been used to model it. This is due to consumer use of air conditioning and electric heaters in warm and cold weather respectively. Extending this approach, we define x_t^3 as the average ambient air temperature, and use it as a feature for the $(t+1)th$ hour because x_t^3 is not observable until the $(t+1)th$ hour.
4) Variance (x_t^4): Most consumer appliances cycle through different modes of operation and have varying consumption. Examples include washers, dryers, microwave ovens and refrigerators. When in operation these different modes of operation cause a large variation in the appliance's and the home's load. Hence, to capture consumer activity we define the variance of measured load values during the $(t-1)th$ hour as a feature x_t^4 for hour t.
5) Last peak load (x_t^5): A consumer's activity period typically spans across hours. Hence, a consumer is more likely to cause a *high* hourly peak load if the previous hour's peak load was also high due to her activity. This is supported by our dataset: we find a high correlation (0.52) between consecutive hours' peak load values.

We define the following encoding to encode the time of day (x_t^1) and day of week (x_t^2) features. Day of week (x_t^2) is encoded as $\tilde{\mathbf{x}}_\mathbf{t}^\mathbf{2} = (\tilde{x}_{t,1}^2, \tilde{x}_{t,2}^2, \tilde{x}_{t,3}^2, \tilde{x}_{t,4}^2, \tilde{x}_{t,5}^2, \tilde{x}_{t,6}^2, \tilde{x}_{t,7}^2)$ where

$$\tilde{x}_{t,i}^2 = \begin{cases} 1, & \text{if } i = x_t^2, \\ 0, & \text{otherwise} \end{cases}$$

Hence, Tuesday is represented as $\tilde{\mathbf{x}}_\mathbf{t}^\mathbf{2} = (0, 1, 0, 0, 0, 0, 0)$. The hour of day ($x_t^1$) is encoded similarly. Such encoding decorrelates consecutive hours and days, allowing the model to predict from a wider range of values. Throughout the remainder of the paper, we shall use x_t^1 and x_t^2 to denote their respective encoded forms.

Given the feature vector, $\mathbf{x_t} = \{x_t^1, x_t^2, x_t^3, x_t^4, x_t^5\}$ of hour t, we use the following nonlinear regression based techniques to find a function $f(\cdot)$, such that, $y_t \approx f(\mathbf{x_t})$. To obtain f, we use the following techniques:

1) Support Vector Regression (SVR): SVR expresses f as a nonlinear function of the input $\mathbf{x_t}$ together with a subset of *support vectors* taken from the dataset [22]. The nonlinear mapping is defined by a *kernel function* and its parameters. The loss criterion of SVR is *ϵ-insensitive*, meaning that the function is not penalized for training data that are predicted within ϵ of their correct value.
2) Least Squares Support Vector Regression (LS-SVR): This approach operates in a fashion similar to SVR but has two important differences. First, *all* of the training data are used as support vectors. Second, the loss criterion is the sum of squared differences between *all* observed and predicted peak loads. LS-SVR is more commonly known as *Gaussian process regression* [21].
3) Artificial Neural Networks (ANN): This non-linear regression approach learns a function expressed in terms of *hidden units* that transform the input features [8]. In our experiments we use 10 hidden units.

To evaluate the contribution of the physical features (x_t^1, x_t^2, x_t^3) and the history-based features (x_t^4, x_t^5) towards predictive accuracy, we measure their cross-validation errors. As an example, we illustrate their impact on the accuracy of SVR. Figure 1 shows the feature analysis for peak load prediction for one home, using a polynomial kernel (degree 3) and varying γ the *regularization* parameter. Increasing γ corresponds to fitting the training data more closely. We defer the detailed evaluation on our entire dataset to Section V. From Figure 1 we see that physical features alone perform much worse than history-based features alone at the optimal γ value of 5. Moreover, when combined, the improvement over using history-based features alone (now achieved at γ=0.45) is slight. This is in sharp contrast to previous work which has focused only on daily, weekly or monthly load prediction for cities, blocks or regions. Further, it hints at the possibility of improvement in prediction accuracy by making better use of recent history of home load. We accomplish this using *time series* methods, which are designed to take advantage of structure in historical observations to make good predictions.

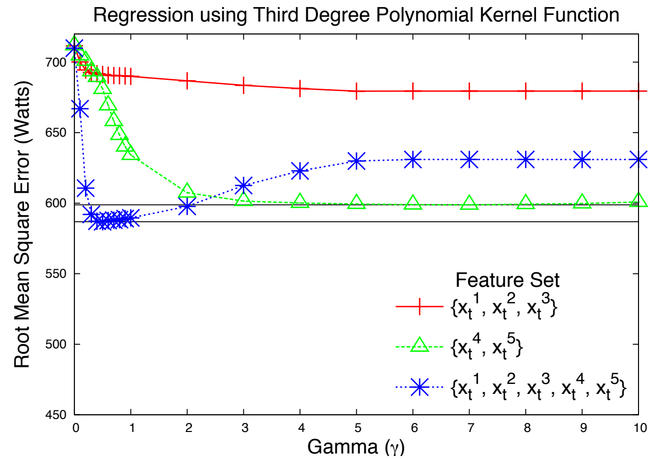

Fig. 1. Variation of prediction *Root Mean Square Error (RMSE)* with varying γ, which measures the trade-off between training error minimization and smoothness of the prediction function. Estimates use 10-fold cross-validation.

B. *Seasonal Auto-Regressive Moving Average Model*

Although there is an observed hidden pattern in home load owing to routine consumer activity, the lack of knowledge of consumers' exact behaviour makes peak load prediction difficult. This is evident from the low prediction accuracy obtained by regression based methods. For instance, on some occasions consumers do their laundry in the morning, causing under-prediction, or skip cooking dinner, leading to over-prediction. To model both the routine and stochastic consumer activity, we use the Auto-Regressive Moving Average (ARMA) model [18]. We denote the measured peak load as a time series $\{y_t\}_{t=1}^T$, where y_t is the peak load of hour t (p_{95} of 600 data points in an hour) and T is the total duration of measurement. The ARMA model decomposes this time series into two parts: Auto-Regressive (AR) and Moving Average (MA). Routine consumer activity is captured by the Auto-Regressive part whereas stochastic activities are captured by the Moving Average (MA). Essentially, MA is a white noise process, which traces abrupt fluctuations in the time series caused by stochastic activities. Further, because consumer routines follow a daily pattern, we use the Seasonal ARMA (SARMA) to model the time series.

1) ARMA Model: Given $\{y_t\}_{t=1}^T$, consider a time series of white noise $\{\epsilon_t\}_{t=1}^T$, the current observation y_t can then be represented as the linear combination of the current white noise ϵ_t, previous observations $\{y_{t-1}, ..., y_{t-p}\}$ and white noise values $\{\epsilon_{t-1}, ..., \epsilon_{t-q}\}$. This forms the $ARMA(p, q)$ model, defined as:

$$y_t = \underbrace{\delta + \sum_{i=1}^{p} \phi_i y_{t-i}}_{AR\ (routine\ activities)} + \underbrace{\epsilon_t - \sum_{i=1}^{q} \theta_i \epsilon_{t-i}}_{MA\ (stochastic\ activities)} \tag{1}$$

where p and q represent the *degree* or number of previous timepoints in the model. Variables δ, ϕ_i and θ_i are model parameters. We define the *backshift operator* B as

$$B^i y_t = y_{t-i}. \tag{2}$$

Rewriting Eq. (1) using Eq. (2), we get

$$y_t = \delta + \left(\sum_{i=1}^{p} \phi_i B^i\right) y_t + \left(1 - \sum_{i=1}^{q} \theta_i B^i\right) \epsilon_t. \quad (3)$$

Let

$$\Phi(B) = \sum_{i=0}^{p} \phi_i B^i,\ \Theta(B) = 1 - \sum_{i=1}^{q} \theta_i B^i. \quad (4)$$

Rewriting Eq. (3) using Eqs. (4), we get

$$\Phi(B) y_t = \delta + \Theta(B)\epsilon_t \quad (5)$$

where ϵ_t should satisfy the stationary assumption, which means ϵ_t is independently and identically distributed with zero mean and constant variance σ^2.

2) SARMA Model: Because the hourly peak load follows a daily periodic pattern, we use the Seasonal Auto-Regressive Moving Average model to capture the "seasonality", i.e. periodic variation, in the peak load time series. Note that in our case seasonality is caused by daily routine rather than weather. We decompose the time series $\{y_t\}_{t=1}^{T}$ into two parts

$$y_t = S_t + N_t \quad (6)$$

where S_t denotes peak load with periodicity s, and N_t denotes peak load without periodicity. Let s denote the period of S_t (24 hours). Hence, $S_t = S_{t+s}$. Using Eq. (2), we get

$$S_t - S_{t-s} = (1 - B^s) S_t = 0 \quad (7)$$

Multiplying Eq. (6) on both sides by $1 - B^s$,

$$(1 - B^s) y_t = (1 - B^s) S_t + (1 - B^s) N_t = (1 - B^s) N_t \quad (8)$$

Since $(1 - B^s) N_t$ lacks periodicity, we model $(1 - B^s) y_t$ using Eq. (4). We have:

$$\Phi(B)(1 - B^s) y_t = \sigma + \Theta(B)(1 - B^s)\epsilon_t \quad (9)$$

Parameters δ, ϕ_i and θ_i are found by minimizing ϵ_t^2 for each value of t, over the dataset.

V. Experimental Evaluation

We use the dataset (described in Section III) comprising load values of 24 homes in the Kitchener-Waterloo area, to evaluate the prediction models.

A. Regression

We first extract the feature set (described in Section IV-A) from the dataset. Each feature is then scaled to the range $[-1, 1]$. To evaluate prediction accuracy, we divide each home's dataset into 10 subsets using randomized sub-selection and perform a 10-fold cross validation. We use the following metrics to measure prediction error: Root Mean Square Error (RMSE) is defined as:

$$\text{RMSE} = \sqrt{\frac{\sum_{t=1}^{N} (y_t - \hat{y}_t)^2}{N}}$$

where y_t, $\hat{y}_t$ are measured and predicted peak load for hour t.

Since RMSE values are not comparable across datasets, we also define the Normalized Mean Square Error (NMSE):

$$\text{NMSE} = \frac{\frac{1}{N} \sum_{t=1}^{N} (y_t - \hat{y}_t)^2}{\text{Var}(y_t)} = \frac{\text{MSE}}{\text{Var}(y_t)}$$

NMSE consists of the ratio of the Mean Square Error (MSE) to the variance of a given test set.

We study four different kernel functions: a Linear kernel, a Polynomial kernel, a Radial Basis Function (RBF) kernel, and a Sigmoid kernel function using the LIBSVM library [10]. We use the ϵ-insensitive Vapnik loss function [22], where the penalty incurred is 0 for error in the range $[-\epsilon, \epsilon]$ and increases linearly otherwise. We find that the Polynomial kernel function with degree 3 at $\epsilon = 100$ provides the most accuracy for peak load prediction using SVR and LS-SVR in all 24 homes. However, different optimal γ values are observed in the various homes. We also compare the performance of SVR and LS-SVR, with ANN using the same feature set.

Figure 2 shows a comparison between these techniques using RMSE over all 24 homes' dataset. We find that LS-SVR outperforms ANN in 17 homes' datasets and LS-SVR in 21 homes' datasets. As shown in Figure 3, similar results are observed when using NMSE as the metric for prediction error. When averaged over all homes, LS-SVR provides a RMSE of 900 W and a NMSE of 0.64.

To find the effectiveness of these techniques in predicting hourly mean load (or energy consumption), we test them over our dataset using the 10-fold cross-validation technique described above. Figures 4 and 5 show (using RMSE and NMSE respectively) a comparison of these techniques in predicting mean load. As in case of peak load prediction, LS-SVR outperforms SVR and ANN for mean load prediction. This validates the finding by Edwards et al. [13] and quantifies the accuracies of the three methods. However, for each home both RMSE and NMSE for mean load prediction are lower than that for peak load prediction. When averaged over all homes, LS-SVR provides a RMSE of 390 W and a NMSE of 0.48 when predicting mean load, but a RMSE of 900 W and NMSE of 0.64 when predicting peak load. These values establish the ineffectiveness of mean load predictors for peak load prediction.

We evaluate the contribution of individual features by incrementally building the feature set starting with x_t^1, and measuring the cross-validation RMSE at each step. Shown in Figure 6 is the variation in RMSE for SVR averaged over all 24 homes' datasets. Similar results are obtained for LS-SVR; we omit them due to space limitations. The decrease in RMSE on adding x_t^2 (day of week) and x_t^3 (ambient temperature) is significantly smaller than on adding x_t^4 (variance) and x_t^5 (last peak load) to the feature set. Hence for hourly peak load prediction, the contribution of the variance and peak load observed during one hour is more significant for predicting the next hour's peak load, than the temperature, time and day.

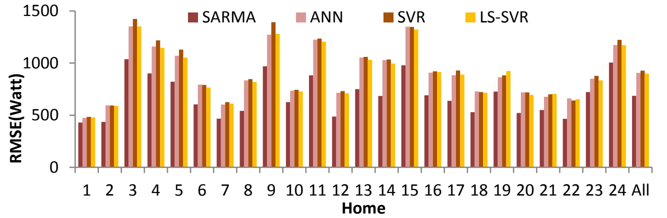

Fig. 2. RMSE for peak load prediction using SARMA, ANN, SVR, LS-SVR

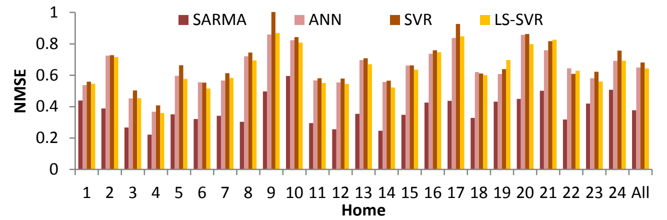

Fig. 3. NMSE for peak load prediction using SARMA, ANN, SVR, LS-SVR

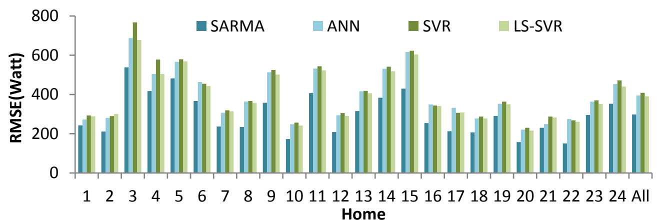

Fig. 4. RMSE for mean load prediction using SARMA, ANN, SVR, LS-SVR

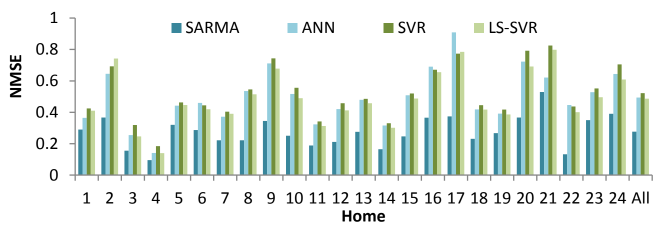

Fig. 5. NMSE for mean load prediction using SARMA, ANN, SVR, LS-SVR

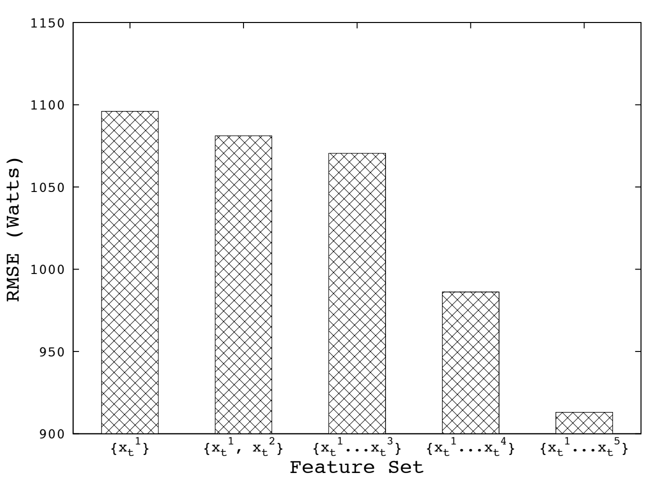

Fig. 6. RMSE for peak load prediction using LS-SVR averaged over all 24 homes

B. SARMA

To find the optimal values of parameters p and q in our ARMA model, we study the Lagged Autocorrelation of the hourly peak load values (shown in Figure 7). We observe that values corresponding to the last 5 hours and the 22nd to 25th hours have more significant correlation to peak load than others. Further, we study the partial correlation of the hourly peak load (shown in Figure 8). This allows us to observe the correlation between y_t and y_{t-k} with the linear dependence of $y_{t-1} \ldots y_{t-k-1}$ eliminated. Optimal values of p and q obtained using least square error minimization are 5 and 30 respectively.

Since for time series based approaches such as SARMA cross-validation cannot be performed, we evaluate it by sequential prediction over the dataset using the optimal parameter values for p and q. Shown in Figures 2 and 3 are the RMSE and NMSE values for SARMA and their comparison with SVR and LS-SVR. We observe that SARMA outperforms all other three approaches-SVR, LS-SVR and ANN. Similar results are observed for mean load prediction (shown in Figures 4 and 5). By comparing Figures 2 and Figure 4, and, Figures 3 and 5, we observe the prediction accuracy improvement of SARMA is greater for peak load prediction than for mean load prediction. Hourly mean load bears less stochasticity than peak load, due to latter's higher dependence on consumer activity. SARMA targets stochastic consumer activities and is more suitable for predicting peak load. Finally, SARMA provides a peak load prediction RMSE of 685 W (averaged over all homes), and NMSE improvement of at least 10% in each of the homes' datasets, and 30% when averaged over all homes, as compared to LS-SVR.

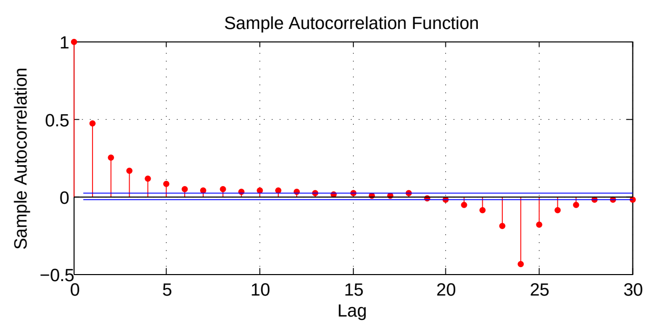

Fig. 7. Auto-Correlation Function of the time series with seasonality removed

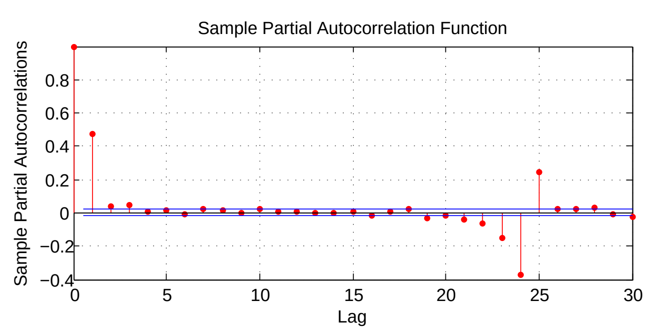

Fig. 8. Partial Auto-Correlation Function of the time series with seasonality removed

VI. Discussion

A. Consumer Activity

From our experiences with home load prediction, we conclude that predicting consumer activity is most significant for home load prediction. This is evident from our comparative evaluation of the different features. Due to the high temporal resolution (6 seconds) of our dataset, we were able to use the variance of measured load values as a coarse indicator of consumer activity, but better indicators should be investigated. For example, load disaggregation techniques [26] can be used to identify operation of specific appliances using the home load which in turn reflect activity. However, this requires prior knowledge (or $learning$) of the various home appliances. We defer the investigation of these possibilities to future work.

B. Temperature as a SARMA Input

One might expect that a combination of physical and history-based features might yield even better predictive performance. Therefore, we investigate the use of ambient air temperature as an input to the SARMA model. We apply the Auto-Regressive Moving Average with *Exogenous Inputs (ARMAX)* model to do so. $\mathrm{ARMAX}(p, q, r)$ is defined as:

$$y_t = \underbrace{\delta + \sum_{i=1}^{p} \phi_i y_{t-i}}_{Auto-Regressive} + \underbrace{\epsilon_t - \sum_{i=1}^{q} \theta_i \epsilon_{t-i}}_{Moving\ Average} + \underbrace{\sum_{i=1}^{r} \beta_i T_{t-i}}_{Exogenous\ Input} \quad (10)$$

where $\{T_t\}$ is the ambient air temperature at hour t, $\beta_i, i \in \{1, \ldots, r\}$ are model parameters, and $\{p, q, r\}$ are the degrees of the model. Similar to SARMA, the parameters in ARMAX are obtained by minimizing squared error.

However, we observed that the improvement in prediction accuracy due to ARMAX is not significant and negative in certain cases (detailed results omitted due to space limitations). This is due to the fact that the effect of temperature is not greatly observable at an hourly or smaller timescale. Intuitively, a higher temperature is expected to increase the load due to consumers' use of air-conditioning and similar heavy load appliances. Thus the effect of temperature will be observable on a larger (e.g. daily) timescale because of an increase in the *base* load at that timescale. Since the SARMA model only uses only up to the last 27 hours' observations, it fails to capture the effect of temperature on a daily or longer timescale.

VII. Conclusion

We find that existing machine learning methods for hourly mean load prediction do not perform well for peak load prediction. Further, when using regression, physical features like temperature, day and time, which are helpful for daily or weekly load prediction, are not as effective for hourly prediction as history-based features. On the other hand, SARMA, a time series based model, is 30% more accurate than SVR, LS-SVR, and ANN, both for peak and mean load prediction. This is because SARMA can capture consumers' routine and stochastic activities, which are the most significant factors influencing home load prediction. We plan to use our approach to study peak load reduction in using energy storage.

References

[1] Current Cost Envi. http://www.currentcost.com/product-envi.html.

[2] Global adjustment and the IESO market. http://www.ieso.ca/imoweb/pubs/ga/GA_Presentation.pdf.

[3] Ontario power authority. ipsp stakeholder consultation: Conservation. http://www.powerauthority.on.ca/sites/default/files/page/IPSPn

[4] Ontario time-of-use pricing. http://www.ontario-hydro.com.

[5] Transitional demand response program. http://www.ieso.ca/imoweb/pubs/settlements/se_TDRP_Program.pdf.

[6] University of Waterloo Weather Station. http://www.weather.uwaterloo.ca/.

[7] I. Bayram, G. Michailidis, M. Devetsikiotis, S. Bhattacharya, A. Chakrabortty, and F. Granelli. Local energy storage sizing in plug-in hybrid electric vehicle charging stations under blocking probability constraints. In *Proc. of SmartGridComm, 2011*.

[8] C. Bishop. *Neural networks for pattern recognition*. Oxford University Press, USA, 1995.

[9] T. Carpenter, S. Singla, P. Azimzadeh, and S. Keshav. The impact of electricity pricing schemes on storage adoption in ontario. In *Proc. of e-Energy*, 2012.

[10] C.-C. Chang and C.-J. Lin. Libsvm: A library for support vector machines. *ACM Trans. Intell. Syst. Technol.*, 2(3):27:1–27:27, May 2011.

[11] 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*, 19(4):1821 – 1830, November 2004.

[12] C. Chen and D. J. Cook. Energy outlier detection in smart environments. In *Proc. of Artificial Intelligence and Smarter Living*, 2011.

[13] R. E. Edwards, J. New, and L. E. Parker. Predicting future hourly residential electrical consumption: A machine learning case study. *Energy and Buildings*, March 2012.

[14] S. Fan and L. Chen. Short-term load forecasting based on an adaptive hybrid method. *IEEE Transactions on Power Systems*, 21(1):392 – 401, feb. 2006.

[15] H. Hahn, S. Meyer-Nieberg, and S. Pickl. Electric load forecasting methods: Tools for decision making. *European Journal of Operational Research*, 199(3):902 – 907, 2009.

[16] H. Huang, R. Hwang, and J. Hsieh. A new artificial intelligent peak power load forecaster based on non-fixed neural networks. *International Journal of Electrical Power and Energy Systems*, 24(3):245 – 250, 2002.

[17] P. Mandal, T. Senjyu, N. Urasaki, and T. Funabashi. A neural network based several-hour-ahead electric load forecasting using similar days approach. *International Journal of Electrical Power and Energy Systems*, 28(6):367 – 373, 2006.

[18] T. C. Mills. *Time Series Techniques for Economists*. Cambridge University Press, 1991.

[19] A. Pardo, V. Meneu, and E. Valor. Temperature and seasonality influences on spanish electricity load. *Energy Economics*, 24(1):55 – 70, 2002.

[20] S. Ramesh, B. Natarajan, and G. Bhagat. Peak load prediction using weather variables. *Energy*, 13(8):671 – 679, 1988.

[21] C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning*, chapter 2 Regression. MIT Press, 2006.

[22] V. N. Vapnik. *The nature of statistical learning theory*. Springer-Verlag New York, Inc., New York, NY, USA, 1995.

[23] P. Vytelingum, T. D. Voice, S. D. Ramchurn, A. Rogers, and N. R. Jennings. Agent-based micro-storage management for the smart grid. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems*, Richland, SC, 2010.

[24] Y. Yao, Z. Lian, S. Liu, and Z. Hou. Hourly cooling load prediction by a combined forecasting model based on analytic hierarchy process. *International Journal of Thermal Sciences*, 43(11):1107 – 1118, 2004.

[25] H. Zareipour, C. Caizares, and K. Bhattacharya. The operation of ontario's competitive electricity market: Overview, experiences, and lessons. *IEEE Transactions on Power Systems*, 22(4):1782 –1793, 2007.

[26] M. Zeifman and K. Roth. Nonintrusive appliance load monitoring: Review and outlook. In *IEEE Transactions on Consumer Electronics*, volume 57, pages 76–84. IEEE Computer Society, 2011.

[27] B. Zhang and Z. Dong. An adaptive neural-wavelet model for short term load forecasting. *Electric Power Systems Research*, 2001.