Machine Learning Method for Day Classification to Understand Thermostatically Controlled Load Demand

Ying Guo
Data 61
The Commonwealth Scientific and Industrial Research Organisation (CSIRO)
Marsfield, NSW 2122, Australia
ying.guo@csiro.au

Nariman Mahdavi
Energy Flagship
The Commonwealth Scientific and Industrial Research Organisation (CSIRO)
Mayfield West, NSW 2304, Australia
Nariman.Mahdavimazdeh@csiro.au

***Abstract*—An accurate estimation of the thermostatically controlled loads (TCLs) demand is important for the load control to provide demand-side ancillary services, such as peak-demand reduction. CSIRO developed a mathematical model that provides the aggregate demand of a population of TCLs for a given set of model parameters and a given ambient temperature profile. The model parameters, however, need to be identified (estimated) from experimental data. By recognizing the importance of accurate and efficient identification of model parameters, in this paper, we developed a machine learning methodology using support vector machines (SVMs) to classify days into "hot", "cold" or "mild" days using very limited aggregated information. This classification is the most important part of the identification and is based on experimental evidence combined with ambient temperature readings and aggregate TCL demand. . The simulation results on real datasets across Australia indicate over 88% accuracy of successful classification.**

***Index Terms*—aggregated power demand, classification, machine learning, support vector machine, thermostatically controlled load**

I. INTRODUCTION

Global energy consumption has grown rapidly in the last decade and its increase is steadily accelerating. Power systems around the world all face the risk of outages. One of the recent examples was the blackout in Adelaide, South Australia on 8th February 2017, where more than 40,000 households and businesses were affected. The majority of these blackout events occurred during hot summer afternoons, even though a miniscule portion of the year falls under this category. Since the thermostatically controlled loads (TCLs) are the most significant temperature-dependant loads in Australia's electricity network, an estimation of the TCL demand in a load control scenario can reduce the peak-demand, and thus, contribute to a more resilient power system.

There are mainly two types of energy modelling approaches: bottom-up and top-down models, see [1] and [2] for example. While the former requires the energy data of individual households, the latter works on the aggregate demand of a population of households. Although some datasets record appliance-level measurements for research purpose, most of the readily available datasets are at the substation level, which includes the aggregate energy demand of all the customers within the supply area of a given zone substation. Therefore, a practical TCL demand estimation and prediction should be at the substation level. A number of elaborate mathematical models for the aggregate demand of such populations of loads have been developed in recent years, see [3] and [4] for example.

While most of the existing models typically characterize demand response to changes in thermostat setpoint, the model in [5] considers ambient temperature as the main excitation input. Thus, the aggregate demand of a heterogeneous population of TCLs has been modelled for a given set of physical parameters and a given ambient temperature profile. This new model only requires local weather and demand data to estimate the model parameters, with no need for additional metering or direct engagement of grid users.

In this paper we introduce support vector machines, a machine learning based method, to classify the days into "hot", "cold" or "mild" catalogues based on the statistical features from the aggregated substation level data. With the correctly classified days, the model can achieve a better performance by identifying proper parameters. Section II describes the SVM algorithm in details. Section III presents the experimental results in comparison with previous threshold based method. The paper concludes in Section IV with a discussion of extensions to the present work.

II. CLASSIFIER FOR TCL MODEL PARAMETER IDENTIFICATION

A. Problem Description

The parametric model developed by Mahdavi et al. can be used to predict the fraction of TCLs for a given population from aggregate load data when general ambient temperature is known [5]. In practice, these model parameters are not known and need to be identified from experimental data. Such parameter identification requires correlated data for ambient

978-1-5386-4950-3/17/$31.00 ©2017 Crown

temperature and aggregate TCL demand on a number of days where TCL are most likely in operation. In this paper, we refer to the days where TCLs are most likely to be in operation as "hot" days.

Additionally, the aggregate demand data available for electricity zone substations typically consists of a blend of all loads, of which the aggregate TCL demand makes a fraction that needs to be disaggregated. To approximate the aggregate TCL demand fraction in the total aggregate demand we use days classified as "mild", which are those where TCLs are most likely not in operation. In mild days TCL demand fraction is assumed to be negligible, which means that the total demand in those days is completely made up by non-TCL loads, which we refer to as "baseline load". By subtracting the total demand of a mild day from that of a hot day we obtain an estimation of the TCL demand for hot days. The same idea can be used for winter, using days classified as "cold", where TCLs are most likely in operation (as space heaters).

A simple method to classify days according to whether ACs are in operation or not can be based on applying a threshold on the maximum ambient temperature. We refer to this approach as the "threshold" method. The threshold is a constant across states, and can be determined by real data. A benefit of this method is its low computation cost, and it may be directly applied to data from any weather station without any other requirements. The classification would then be dependent on the threshold values chosen to make the decisions.

In this paper we develop an advanced classification method based on machine learning techniques, which use data to tune the decision process [6].

B. Support Vector Machines for Classification

To solve the classification problem, a support vector machine (SVM) has been developed. The SVM algorithm was originally derived from statistical learning theory. It is designed to maximize the predictive accuracy while automatically avoiding over-fit to the training dataset, hence achieve better empirical performance. The formulation uses the Structural Risk Minimization principle, which minimizes an upper bound on the expected risk, while most classification algorithms minimize the error on the training data. It is this difference which equips SVM with a greater ability to generalize, which is the goal in statistical learning [7].

Support Vector machines have a key feature: mapping training data series $\boldsymbol{X}$ nonlinearly into a higher dimensional kernel-introduced feature space, and computing the hyperplane. After the nonlinear mapping, a cost function that ignores errors within a certain distance of the true value is defined, aiming to find the hypothesis $f(\boldsymbol{X})$ with good generalization performance.

Two groups of key functions can affect the performance of SVM:

- Kernel function $k(x_i, x_j)=\langle\phi(x_i), \phi(x_j)\rangle$, which makes use of an implicit nonlinear map $\phi(x)$ from the input space into the feature space;
- Cost function $c(y, f(\boldsymbol{X}))$ where y is the real value and $f(\boldsymbol{X})$ the hypothesis.

There are many choices for kernel functions, such as polynomial kernels; radial basis functions or sigmoid kernels. For current study, we choose the radial basis function as the kernel function.

The training data $\boldsymbol{X}$ for SVM can be special features from original datasets. With $\boldsymbol{X}$ as the training input, the classification problem is to find the estimated class of any new testing dataset $\boldsymbol{X}_t$. The training process of SVM is a quadratic optimization problem, as explained in [8].

III. Experimental Results

The sample population used to develop and calibrate the method is the RBEES dataset [9], where the air conditioning (AC) demand and the total demand of around 200 residential houses are available. These houses are distributed in three states across Australia: Queensland (QLD), Victoria (VIC), and South Australia (SA).

According to the total aggregate demand of the population, ambient temperature, and humidity, we classify the days into most likely "on" days (ACs in operation), most likely "off" days (ACs not in operation), and others. For validation purposes, the most likely "on" days are considered to be the days with average aggregate AC demand (per-home) greater than 0.5 kWh per-home, and most likely "off" days are the days with average aggregate AC demand (per-home) less than 0.15 kWh per-home.

A. SVM Implementation

To train the SVM, we need to select a number of suitable "features" and their associated labels for the class of days. All these features are constrained to the aggregate power demand, external weather conditions, and the time of the day, rather than from household sub-metering. Two different combinations have been considered for the features:

- $\boldsymbol{X}_1$ = [peak demand; corresponding temperature; corresponding humidity; day of the week]; -- 4 features
- $\boldsymbol{X}_2$ = [peak demand; 2 hour temperature period; 2 hour humidity period; day of the week]; --1+4+4+1=10 features

where

- peak demand (kWh) – the daily peak demand (of the average demand at each time point across all households in a population);
- corresponding temperature (ºC) – the ambient temperature when the demand is the highest;
- corresponding humidity (%) – the external humidity when the demand is the highest;
- day of the week – [1-7] for Sunday to Saturday;
- 2 hour temperature period – 2 hours' ambient temperature before the demand is the highest;

- 2 hour humidity period – 2 hours' external humidity before the demand is the highest.

The associated labels for the class of the days are:

- 1 – cold day with peak of average TCL demand <0.15 kWh
- 2 – mild day with peak of average TCL demand between 0.15 kWh and 0.5 kWh
- 3 – hot day with peak of average TCL demand >=0.5 kWh

Experimental results show the effectiveness of the second feature X_2, which is the base of the results shown in this section. Without loss of generality, we only focus on the summer period of a year, where the results can be divided into two categories:

- Train the SVM with the first half of the summer days, and test it with the second half for the houses within the same state.
- Train the SVM with the whole summer days of a state, and test on the houses of other states. The states include QLD, VIC, and SA.

B. *SVM Experimental Results*

Based on RBEES dataset, we have three training sets and also three test sets, which are the populations in QLD, VIC, and SA. Therefore, we have nine different combinations of training and testing with a summary of the classification accuracy shown in Table I. Altogether the average of accuracy of SVM classifiers is 88.07%. Additionally, Figures 1 and 2 show the results of three cases, where the training is based on the QLD population and testing is done on all three populations. Notice that when the training set and the testing set are from the same population, we use the first 45 days' data for training, and the rest 45 days' data for testing.

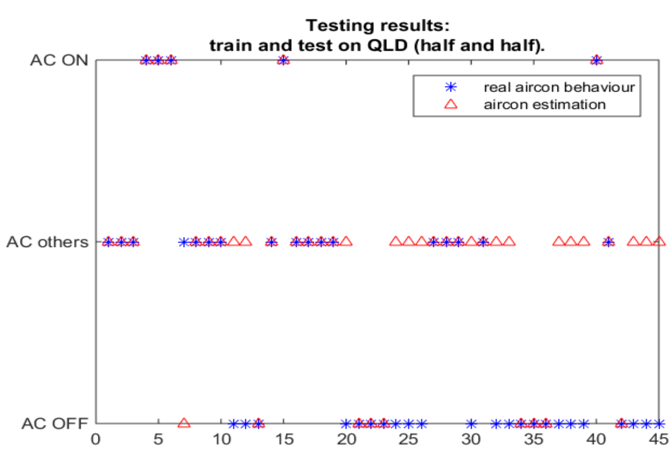

Figure 1. Classification of days, training and testing on 45 summer days in Queensland

Figure 2. Classification of the days, training with summer days in QLD Queensland, and test testing on 90 summer days in VIC Victoria

TABLE I. ACCURACY OF SVM CLASSIFIERS

Training state	**Testing state**		
	QLD	***VIC***	***SA***
QLD	79.26%	86.67%	88.15%
VIC	88.15%	94.07%	90.37%
SA	94.07%	87.41%	84.44%

QLD, Queensland; SA, South Australia; VIC, Victoria

We observe an inconsistency in the results. Namely, the testing result of Figure 2 appears better than those in Figure 1**Error! Reference source not found.**, which means that the results are better when the training and testing sets are different, contrary to what one would expect. This inconsistency only manifests itself when the QLD population is used as a training set. A factor that may be causing this inconsistency is the threshold levels used, which are postulated and the same to classify AC usage across three states with potentially very different behavioral patterns. This hypothesis could be tested by trying to estimate (rather than postulate) threshold values using data and find levels that maximize the number of predicted true positive and negatives, using for example a simple threshold classification method for testing. Another possible reason could be that the sample size is small – there are only 45 samples when the training and testing sets are from the same state. Larger sample size, such as data over a few summers' days, can be more representative of the feature of the classification problem.

C. *Comparison between SVM method and Threshold method*

In what follows we compare the TCL classification results of SVM method with the original threshold method. Firstly, we define comparison indexes as:

- True positive (TP): hit correctly – e.g. classified as "AC ON" for an "AC ON" day.

- True negative (TN): correct rejection – e.g. classified as “not AC ON” for a “not AC ON” day.
- False positive (FP): false alarm (aka Type I error) – e.g. classified as “AC ON” for a “not AC ON” day.
- False negative (FN): miss (aka Type II error) – e.g. classified as “not AC ON” for an “AC ON” day.
- Positive samples (Pos): Pos=TP+FN.
- Negative samples (Neg): Neg=FP+TN.

Table II shows a summary of the defined indexes. We also define an index for the overall accuracy of the model classification, which is a weighted average of true positive and true negative rates

$$\text{Accuracy} = (TN+TP) / (Pos+Neg) = (TN+TP)/(TP+FN+FP+TN).$$

TABLE II. COMPARISON INDEXES FOR THE SVM AND THRESHOLD METHODS

Ground truth	Predicted class	Index	Note
Yes	Yes	TP	Pos = TP + FN
Yes	No	FN	
No	Yes	FP	Neg = FP + TN
No	No	TN	

For each of the nine different combinations of the train and test set, we have 3 classes of “AC ON”, “AC OFF”, and “AC Others”. Thus, in each of the nine train/test cases, the days in the testing period are divided into three classes based on their actual AC consumption. We then compute and compare the accuracy index for each of these twenty-seven possibilities between the SVM method and the threshold method, shown in Table III. The higher accuracy of each case is shown in bold. The accuracy of the SVM method is higher in twenty-one cases, while the accuracy of the threshold method is higher in five cases. There is one case where both methods achieve the same accuracy. The average of accuracy of the threshold method is 79.26%, compared with an 88.07% from the SVM method. Overall, the SVM method achieved better performance.

A key factor that has not been directly considered is occupancy. While days with a maximum temperature between 30°C and 40°C would in thresholding be classified as “hot”, and thus indicate “AC ON”, if no one is at home (typically on weekdays for working residents), a false positive may arise. To somehow account for occupancy, days also divided into weekend and week days.

In this regard, the SVM approach may need further adjustments, as the data used for training applied to all days of the week indiscriminately. Yet, particularly on weekdays, the data may be overly noisy due to occupancy patterns. This issue may be addressed in the future by applying the classification to “similar” type days, such as weekends, which would have higher levels of occupancy.

D. The feature of misclassified days

For further comparison, we list the maximum ambient temperature of the FP and FN days produced by both classification methods. The histogram plots in Figure 3 and Figure 4, show the misclassified days’ maximum ambient temperature for “AC OFF” and “AC ON” classes for all 675 testing days in the 27 possible combinations (three same-state test cases and six cross-state test cases: 3*45 + 6*90 = 675). In each figure, the top two plots are the SVM results, number of FN and FP days, and the bottom two are the threshold method results.

For SVM, the similar numbers of FN days and FP days prove that SVM tends to achieve better empirical performance. Figure 4 shows similar performance for the “AC ON” situation. Overall, the SVM method demonstrated better accuracy over the whole RBEES dataset across 27 different combinations. The threshold method performs better than the SVM method in capturing the “AC OFF” class with only 2 cases of FP days, where the days that are incorrectly classified as OFF, see Figure 3. For the FN days, where the AC is classified as ON when it is in fact OFF, the threshold method misses many days when the maximum temperature is between 25°C to 30°C.

TABLE III. ACCURACY COMPARISON BETWEEN THE SVM METHOD AND THE THRESHOLD METHOD

	Testing state →	**QLD**			**VIC**			**SA**		
Training state ↓	***Class →*** / ***Method ↓***	***ON***	***OFF***	***Other***	***ON***	***OFF***	***Other***	***ON***	***OFF***	***Other***
QLD	SVM	**100%**	**68.9%**	**68.9%**	**94.4%**	85.6%	**80.0%**	**93.3%**	**88.9%**	**82.2%**
	threshold	97.8%	48.9%	46.7%	87.8%	**91.1%**	78.9%	86.7%	83.3%	72.2%
VIC	SVM	**95.6%**	**84.4%**	**84.4%**	92.2%	**98.9%**	**91.1%**	**91.1%**	**94.4%**	**85.6%**
	threshold	88.9%	**84.4%**	73.3%	**94.4%**	67.8%	62.2%	86.7%	83.3%	72.2%
SA	SVM	**93.3%**	**97.8%**	**91.1%**	93.3%	**87.8%**	**81.1%**	**92.2%**	84.4%	76.7%
	threshold	88.9%	86.7%	75.6%	**94.4%**	67.8%	62.2%	87.8%	**91.1%**	**78.9%**

been implemented to real dataset, RBEES, and achieves 88.09% accuracy. With the correctly classified days, the mathematical model can achieve better performance with more accurate model parameters.

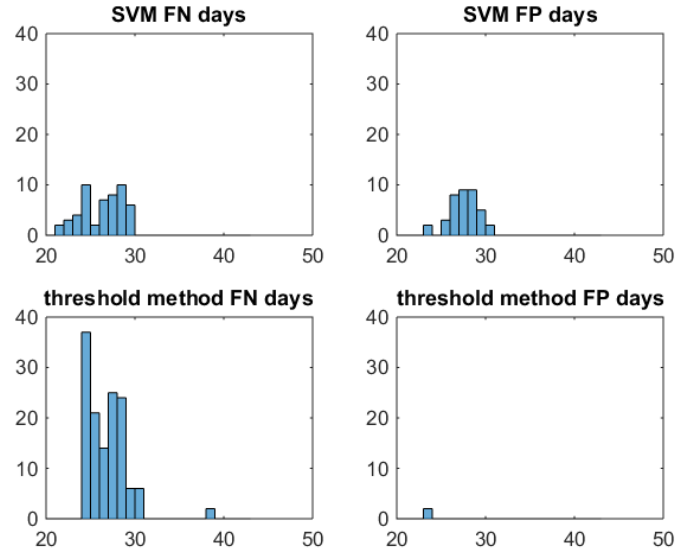

Figure 3. Histograms of maximum ambient temperature in the “AC OFF” class for misclassified days

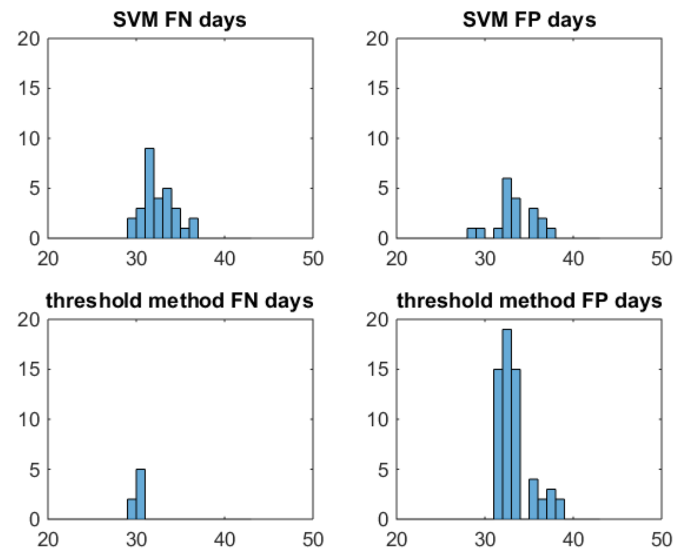

Figure 4. Histograms of maximum ambient temperature in the “AC ON” class for misclassified days

IV. CONCLUSION

An accurate estimation of the thermostatically controlled loads demand can be important for a resilient power system. This result contributes to the work of Mahdavi et. al. in developing a model that aggregates the demand of a population of TCLs within a given set of model parameters and a given ambient temperature profile. To improve the estimation of model parameters from experimental data, we developed a machine learning based method of support vector machines, which classify days into “hot”, “cold” or “mild” catalogues based on the statistical features from the aggregated substation level data. The presented approach has been implemented to real dataset, RBEES, and achieves 88.09% accuracy. With the correctly classified days, the mathematical model can achieve better performance with more accurate model parameters.

Methods for improvement of SVM accuracy are discussed in Section III. For instance, SVM can optimize the labelling of the ground truth data considering the geographical distribution of the datasets. Furthermore, the datasets could be catalogued by occupancy patterns to avoid noisy data. For instance, by applying the classification of “similar” type days, the higher occupancy levels on weekends would be recognized. With the possible availability of much richer datasets across Australia, the SVM has the potential to achieve better performance in future work.

ACKNOWLEDGMENT

This work was funded by the Department of Industry and Science under the Energy Use Data Model project. Thank you to Dr Adam Berry and Dr Julio Braslavsky for helpful comments and suggestions. We wish to acknowledge and thank the Department of Climate Change and Energy Efficiency (which was abolished in March 2013 and its energy efficiency functions transferred to the Department of Industry) for providing a proportion of the funding for the project under which the RBEES data was collected. We wish to acknowledge and thank the following organisations for their collaboration and substantial technical contributions to the project under which the RBEES data was collected: Energy Makeovers Pty Ltd, Energy Monitoring Solutions Pty Ltd, and Association of Building Sustainability Assessors.

REFERENCES

[1] L. G. Swan and V. I. Ugursal, “Modeling of end-use energy consumption in the residential sector: A review of modeling techniques,” *Renewable and Sustainable Energy Reviews*. 2009.

[2] M. Kavgic, A. Mavrogianni, D. Mumovic, A. Summerfield, Z. Stevanovic, and M. Djurovic-Petrovic, “A review of bottom-up building stock models for energy consumption in the residential sector,” *Build. Environ.,* 2010.

[3] C. Perfumo, E. Kofman, J. H. Braslavsky, and J. K. Ward, “Load management: Model-based control of aggregate power for populations of thermostatically controlled loads,” *Energy Convers. Manag.,* vol. 55, pp. 36–48, 2012.

[4] M. Kamgarpour, C. Ellen, S. E. Z. Soudjani, S. Gerwinn, J. L. Mathieu, N. Mullner, A. Abate, D. S. Callaway, M. Franzle, and J. Lygeros, “Modeling options for demand side participation of thermostatically controlled loads,” in Proc. IEEE Bulk Power Syst. Dyn. Control-IX Optim., Sec. Control Emerg. Power Grid (IREP), Greece, Aug. 2013.

[5] N. Mahdavi, J. H. Braslavsky, C. Perfumo, “Mapping the effect of ambient temperature on the power demand of populations of air conditioners,” *IEEE Trans. Smart Grid*, available online, 2016.

[6] Y. Guo, J. Wall, J. Li, S. West, “Real-Time HVAC sensor monitoring and automatic fault detection system,” In: Mukhopadhyay S., Postolache O., Jayasundera K., Swain A. (eds) Sensors for Everyday Life. *Smart Sensors, Measurement and Instrumentation*, vol. 23. Springer, Cham, 2017.

[7] V. N. Vapnik, Statistical Learning Theory. Wiley, N.Y., 1998.

[8] D. Dehestani, F. Eftekhari, Y. Guo, S. Ling, S. Su, H. Nguyen, “Online support vector machine application for model based fault detection and isolation of HVAC system,” *International Journal of Machine Learning and Computing*, vol. 1, pp. 66-72, 2011.

Ambrose MD, James M, Law A, Osman P, White S, “The Evaluation of the 5-Star Energy Efficiency Standard for Residential Buildings,” CSIRO, Australia, 2013.