A Naive Bayesian Classification Model for Determining Peak Energy Demand in Ontario

Bon Ryu(✉), Tokunbo Makanju, Agnieszka Lasek, Xiangdong An, and Nick Cercone

Department of Electrical Engineering and Computer Science, Lassonde School of Engineering, York University, Toronto, ON, Canada
bonryu2@yorku.ca

Abstract. In the Canadian Province of Ontario, electricity consumers pay a surcharge for electricity called the Global Adjustment (GA). For large consumers, having the ability to predict the top 5 daily energy demand hours of the year, called 5 Coincident Peaks (5CPs), can save millions of dollars in GA costs, and help decrease peak energy usage. This paper presents a Naive Bayesian classification model for predicting the 5CPs. The model classifies hourly energy demand as being a 5CP hour or not. The model was tested using hourly energy demand for the province of Ontario over a 21 year period (1995–2015). Classifying a day as a 5CP hour containing day yielded a mean precision and recall of 0.49 (0.18) and 0.88 (0.23) (Standard deviation is in brackets), respectively. Targeting the 5CP hours to within three candidate hours of potential 5CP containing days yielded a mean precision and recall of 0.47 (0.19) and 0.83 (0.22), respectively.

Keywords: Global adjustment · Energy · Demand · Peak · Prediction · Naive Bayesian · Classification

1 Introduction

In the Canadian province of Ontario, electricity consumers pay above the market price for energy (price has units of $/energy unit, and energy has units of watt-hr, e.g. kilowatt-hr or KWHr). The additional price of energy, referred to as the *Global Adjustment* (GA) rate, is the difference between the market price and the guaranteed prices paid to regulated and contracted generators [2,6]. The GA rate applied to consumers depends on their classification as a *Class A* or a *Class B* customer. Class A customers are the largest energy consumers with top average hourly energy demands of 5 megawatts (MW) or higher, and they pay GA costs through the *Industrial Conservation Initiative* (ICI), established by the *Independent Electricity System Operator* (IESO[1]) [7].

[1] The IESO is a crown corporation whose mandate is to oversee the health and efficiency of the electrical grid. They are tasked with several responsibilities, such as ensuring adequate supply of electricity, and promoting the decrease of peak energy usage.

A. Leon-Garcia et al. (Eds.): Smart City 2015, LNICST 166, pp. 517–529, 2016.
DOI: 10.1007/978-3-319-33681-7_43

The ICI program encourages Class A consumers to shift energy use away from the **5 Coincident Peak** (**5CP**) hours of the year. Henceforth we define 5CP as the top 5 daily maximum energy demands of a fiscal year and **peak** as one of the 5CPs. Demands are reported hourly and have units of Watts. A daily maximum is simply the maximum province wide demand for a given day. The term "Coincident" is in reference to the multiple sources of demand in Ontario during a single hour. Each of the 5CPs are daily maximum demands and must happen on different days of the fiscal year [7].

Each fiscal year (from May 1st to April 30th), the IESO maintains a table sorted by the top 10 daily maximum energy demands of the fiscal year, aggregated for all of Ontario [7]. For each row in the table, the hour for which the energy demand occurred is also recorded. Furthermore, each row in the table has an associated total Ontario wide GA cost for Class A customers [6]. At the end of the fiscal year, a Class A customer pays a percentage of this cost for the first five rows in the list (i.e. the 5CP hours), equal to the percentage of the total Ontario demand they were responsible for during those five hours.

For a given year, the total GA cost incurred by a Class A customer can be worth millions of dollars. Thus, a Class A customer can have very large savings if they can ramp down their energy usage during the 5CP hours.

In this study, we present a *Naive Bayesian Classification Algorithm* to classify hourly energy demands as a peak (*peak*) or non-peak ($peak^c$) based on the definitions outlined in the ICI program. In this context, one is actually predicting which hour will be one of the 5CP hours of the year. This is very different from traditional peak demand forecasting algorithms which attempt to predict a numerical value of demand during predetermined hours of the day.

1.1 Related Work

Currently, we are aware of only one other study in literature that aims to solve the exact same Ontario peak prediction problem [8]. In the paper, Jiang et al., modify and test a few different algorithms from literature that solved similar problems [5]. They refer to these algorithms as the following: "Californias Critical Peak Pricing, "Stopping", and "Optimization". They compared these adapted algorithms with their own novel method, which they called "Probabilistic" and report as the best performer.

Their novel algorithm utilized 14-day ahead daily maximum demand predictions from the IESO (let us call this dataset the **14-DayAhead-Dataset**). For each day of the fiscal year, their algorithm takes the maximum predicted demand for the next day (from the *14-DayAhead-Dataset*), and calculates the probability of that demand being within the top 5 of all daily maximums since the start of the fiscal year until 14 days ahead. If this probability is above some static threshold τ_p, they classify the next days maximum demand to be a one of the 5CPs for that year.

The algorithm calculates probabilities based on some probability theories (*e.g.* order statistics, distribution of differences, *etc.*), as well as employing simple `IF` statements to increase or decrease a demand threshold. Their logic however, requires setting a static value for "extreme temperature". At the same time, their static threshold for probability, τ_p, should be adjusted to acquire acceptable *precision* and *recall*. (See Subsect. 2.4 for definitions of Precision and Recall).

Since the 14-DayAhead-Dataset only exits for 2006 and on-wards, they used 2006 as a training set for their initial demand threshold, and tested on years 2007 to 2013. These test years only had summer peaks; a fact that will be important in comparing our own algorithm to this previous work.

1.2 Motivation

In this study, we attempt to solve this problem by classifying hourly energy demands with a Naive Bayesian Classification model. The intention was to create an algorithm that does not heavily depend on heuristic thresholds. In addition, from a machine learning perspective, we preferred an approach that can easily accommodate many variables as inputs, and which has the potential for testing the best combination of inputs.

The results of the initial model were reported in a previous work [3], which had the limitations of discretizing continuous variables, and not predicting winter coincident peaks. The current version of the model solves these issues.

Furthermore, we wished to train and test our model on more years, including those that have winter peaks. The occurrence of winter peaks is a very real possibility and we wanted a model that could easily accommodate such occurrences.

2 Methods

2.1 A Naive Bayesisan Classifier to Classify Energy Demand as Peaks or Non-peaks

In a Naive Bayesian classification model [11], variables used as inputs create a vector,

$$\boldsymbol{x} = (a, b, c, d, e, f) \tag{1}$$

$\boldsymbol{x}$, is also called a *tuple*, and the elements, a to f are called *attributes*.

Using Bayes theorem, the model calculates the following two conditional probabilities that an hourly demand is a *peak* or $peak^c$, given a tuple $\boldsymbol{x}$.

$$P(peak|\boldsymbol{x}) = \frac{P(\boldsymbol{x}|peak)P(peak)}{P(\boldsymbol{x}|peak)P(peak) + P(\boldsymbol{x}|peak^c)P(peak^c)} \tag{2}$$

$$P(peak^c|\boldsymbol{x}) = \frac{P(\boldsymbol{x}|peak^c)P(peak^c)}{P(\boldsymbol{x}|peak)P(peak) + P(\boldsymbol{x}|peak^c)P(peak^c)} \tag{3}$$

The left hand side of the equations are called posterior probabilities, and there are only two because we only have two classifications, *peak* and $peak^c$. Their sum equals 1, and the final classification of an hourly demand corresponds to whichever posterior probability is largest.

$P(peak)$ is the probability of a peak. Let N_{train} be the number of training years. Since there are only five top-5 peak demands in a year and demand is reported hourly,

$$P(peak) = \frac{\#\,\text{of Peaks}}{\#\,\text{of training hours}}, \tag{4}$$

$$\#\,\text{of Peaks} = 5N_{train}, \tag{5}$$

$$\#\,\text{of training hours} = \sum_{year_i \,\in\, training\ year} \#\ \text{tuples in}\, year_i \tag{6}$$

The probability of a non-peak is simply

$$P(peak^c) = 1 - P(peak) \tag{7}$$

The following conditional probabilities are called *likelihoods*:

$$P(\boldsymbol{x}|peak) = \prod_{A=a..f} P(A|peak) \tag{8}$$

$$P(\boldsymbol{x}|peak^c) = \prod_{A=a..f} P(A|peak^c) \tag{9}$$

where A is a variable for a specific attribute from a to f.

To train the classification model, one first partitions the training tuples into a set of peaks ($\{peak\}$) and non-peaks ($\{peak^c\}$) based on prior knowledge of how to classify them (i.e. $\boldsymbol{x} \in \{peak\}$ `IF` **demand is top-5,** `else` $\boldsymbol{x} \in \{peak^c\}$.

For each discrete attribute that has disjoint bins, one calculates the conditional probability of a tuple belonging to each bin by simply counting. *E.g.* If A is *discrete*,

$$P(A = A_{bin1}|peak) = \frac{\#(\{peak\} \cap \{\boldsymbol{x}|\boldsymbol{x}_A \in A_{bin1}\})}{\#\{peak\}} \tag{10}$$

$$P(A = A_{bin1}|peak^c) = \frac{\#(\{peak\} \cap \{\boldsymbol{x}|\boldsymbol{x}_A \in A_{bin1}\})}{\#\{peak^c\}} \tag{11}$$

In Sect. 2.3, we describe how the likelihoods for the continuous variables were calculated.

2.2 Data Sources and Software Tools

The attributes used to train the Naive Bayesian classification model prototype were Ontario energy demand, hour of the day, day type (*e.g.* holiday or workday), temperature, humidex, and windchill. Historical demand data from May 1^{st}, 1994 April 30^{th}, 2015 is publicly available and were obtained from the IESO [1].

Unfortunately, we were not able to acquire the historical 14-DayAhead-Dataset from the IESO, which Jiang et al. says exists for 2006 onward. Currently, on a real-time hourly basis, the IESO provides hourly predictions up to 7 days

Table 1. Attributes for the Naive Bayesian Classification Model. For continuous attributes, the bandwidths used for Gaussian kernel density estimations of PDFs are shown. The estimation algorithm was developed by Kristan *et al.* [9]

Discrete attribute	#Bins	Bins
Hour of day	24	1, 2, 3, ... , 24
Day type	6	Weekend or holiday, workday Monday, workday Tuesday, workday Wednesday, workday Thursday, workday Friday
Continuous attribute	**PDF**	**Bandwidth**
Temperature	$P(Temperature\|peak)$	Automatic
	$P(Temperature\|peak^c)$	Automatic
Humidex	$P(Humidex\|peak)$	2
	$P(Humidex\|peak^c)$	0.35
Windchill	$P(Windchill\|peak)$	Automatic
	$P(Windchill\|peak^c)$	0.35
Normalized Demand (NormDem)	$P(NormDem\|peak)$	Automatic
	$P(NormDem\|peak^c)$	Automatic

in advance but we could not to the best of our abilities locate such information for past days and years.

Hourly weather data was obtained for the Toronto Buttonville Airport from the climate website of the government of Canada [4]. Raw data was originally imported into a MYSQL database with php scripting. The rest of the work such as querying, model creation, training, and testing were all done in MATLAB [10]. All work was performed on a machine with an Intel Core i5 M 540 CPU, with 8 Gb of ram.

2.3 Attributes and Training

Discrete Attributes. Since peaks also never occurred during the weekends or holidays, the attribute *day type* was designed to differentiate between weekends/holidays, and workday weekdays. *Hour of day* is also an important attribute for the model since high energy demands usually occur during specific hours. Both these attributes had naturally occurring discrete bins.

Continuous Attributes. To avoid the issue of binning continuous attributes, Gaussian Kernel Density Estimation (KDE) was used to estimate probability density functions (PDF). A MATLAB implementation was borrowed from

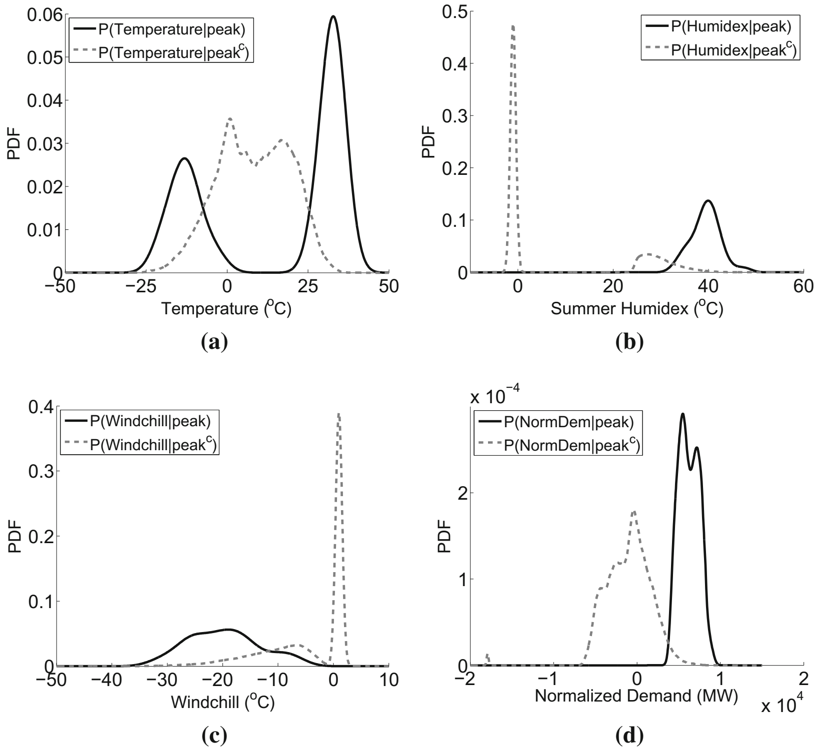

Fig. 1. Example of Gaussian kernel density estimations of continuous attribute PDFs. The PDFs above were derived by training on all years except for 1997. (a) Temperature, (b) Humidex, (c) Windchill, (d) Normalized Demand (NormDem).

Kristan *et al.* [9]. Gaussian KDE involves treating each count of a histogram as a normal Gaussian (i.e. a Gaussian kernel), adding all the kernels, and dividing the resulting function by the number of kernels added in order to re-normalize. Kristan *et al.*'s implementation automatically estimates the bandwidth of the Gaussian kernels to be added (**bandwidth** is a smoothing parameter that modifies the widths of the kernels). This automatic estimation had to be manually adjusted for a few of the distributions, to achieve relatively smooth PDFs. For the sake of helping readers reproduce our results while using Kristen *et al.*'s algorithm, the set bandwidths are listed in Table 1. The PDFs of the continuous variables used for test year 1997 is shown as an example in Fig. 1.

Temperature is an important continuous attribute. High and low temperatures are correlated with summer and winter peaks, respectively. As an additional way of helping differentiate between peaks and non-peaks, we also used humidex and windchill, which are indices that combines the effects of high temperatures with humidity and wind speed, respectively.

There were no humidex values below 25 °C, and no winchill values above 0 °C, and empty humidex and windchill values only existed for $peak^c$ tuples. During pre-processing, empty humidex and windchill values were assigned a value of −1 and 1, respectively. This simplified our code by allowing the KDE algorithm to take care of maintaining the relative size of the empty and non-empty domains of the $peak^c$ PDFs. The exact values assigned to the empty values are not important as long as the ratio $P(A|peak)/P(A|peak^c)$ approaches zero somewhat monotonically, as the humidex and windchill approaches the assigned values from their non-empty domains.

Normalizing Demand. To decrease the variance in the demand *probability distribution functions* (PDFs), the demand is normalized by subtracting from each fiscal year's demand data, the average maximum daily demands of the first 15 working days of the fiscal year. Thus the demand data from all the years are shifted to a common starting point (i.e. to zero), to remove the effects of different baseline demands of different years. Luckily no peaks have ever occurred during the first 15 days of May. Thus in real-time use of the algorithm, this process should not negatively affect the detection of summer peaks.

2.4 Testing and Evaluation

Since we could not obtain the historical 14-DayAhead-Dataset from the IESO (as mentioned previously in Subsect. 2.2), during testing we used the actual historical hourly demand for the testing years. This is a limitation of this current study, which we hope to address in the future by at the very least downloading prediction data on a daily basis and estimating the predicted demand with a random error term.

We did however, test on many years of historical data (*i.e.* for fiscal years ending in 1995 to 2015), whereas in literature [8], tests were only done for 2005 on-wards. Since there were 21 full years of demand data, the model was trained 21 times, each time excluding a desired testing year, and testing on the excluded year.

To help compare the best performing peak prediction algorithms in literature [8] with ours, we compute similar metrics such as precision and recall.

Evaluation Metrics. Precision is the number of true positives (TP) divided by the total number of positive predictions, which is the sum of true positives and false positives (FP). Precision is also know as positive predictive value.

$$Precision = \frac{\#\text{True Positives}}{\#\text{True Positives} + \#\text{False Positives}} = \frac{TP}{TP + FP} \quad (12)$$

Recall is the number of true positives (TP) divided by the number of actual positives (AP), which is the sum of true positives and false negatives (FN). Recall is also known as *sensitivity*.

$$Recall = \frac{\#\text{True Positives}}{\#\text{Actual Positives}} = \frac{TP}{TP + FN} \tag{13}$$

Note that the set of tuples counted as TP or FN make up the set of peaks, $\{peak\}$, and $TP + FN = \#$Actual Positives = 5. FN and True Negative counts (TN) are not reported, but they are simply the complements of TP and FP, respectively.

2.5 Prediction Time-Frame: 24 Hrs, 3 Hrs, and 1 Hr

In the work by previous authors, predictions of peaks are performed to classify an entire day as a peak day or not, and precision and recall is calculated only on daily peak predictions [8].

In the current study, we go a step further and attempt to predict whether or not a peak occurs among the Top 3 hours of the day, or during a single hour. Let these three additional prediction types be called *3 Hr*, and *1-Hr* predictions. The following describes what constitutes a TP or FP for these different prediction types. Let the prediction of a *peak* and $peak^c$ be labeled as `true` and `false`, respectively.

Daily Peak Prediction. A Daily Peak is simply an entire day that is labeled as `true` because it contains at least a single hourly tuple whose value of $P(peak|\boldsymbol{x})$ is greater or equal to 0.5. If an actual *peak* occurs on this day, then increase TP by 1. If an actual *peak* does not occur on this day, then increase FP by 1.

3-Hr Peak Prediction. Once a Daily Peak is predicted, look at the *three hours with the highest values* of $P(peak|\boldsymbol{x})$, and label these three hours as `true`. If an actual *peak* occurs during one of the three hours, then increase TP by 1. If an actual *peak* does not occur during one of the three hours, then increase FP by 1.

1-Hr Peak Prediction. Once a Daily Peak is predicted, look at the *hour* with the highest value of $P(peak|\boldsymbol{x})$, and label this hour as `true`. If an actual *peak* occurs during this hour, then increase TP by 1. If an actual *peak* does not occur on this hour, then increase FP by 1.

3 Results

True positive count (TP), false positive count (FP) count, precision (Eq. 12) and recall (Eq. 13) were computed for all 21 test years, and are displayed in Table 2. This table compares the precision and recall for the three different prediction types mentioned above.

Precision for daily predictions ranged from 0.25 and 1.00, and was below 0.4 for five out of the 21 test years (1996, 1998, 2003, 2006). The low precision for these years mean that the FP count was high compared to the TP count.

Table 2. TP, FP, Precision, and Recall for all 21 testing years. Fiscal year ends are from 1995 to 2015. Color scales were generated in Microsoft Excel. The blue scale is between 0 and 5, red scale is between 0 and 12, and green scale is between 0.0 and 1.0. *#S Pks* and *#W Pks* stand for number of actual summer and winter peaks, respectively.

			TP			FP			Precision			Recall		
Test Year	#S Pks	#W Pks	Day	3-Hr	1-Hr	Day	3-Hr	1-Hr	Day	3-Hr	1-Hr	Day	3-Hr	1-Hr
1995	0	5	5	5	3	4	4	6	0.56	0.56	0.33	1.0	1.0	0.6
1996	0	5	3	3	1	9	9	11	0.25	0.25	0.08	0.6	0.6	0.2
1997	0	5	2	2	1	2	2	3	0.50	0.50	0.25	0.4	0.4	0.2
1998	1	4	1	1	1	3	3	3	0.25	0.25	0.25	0.2	0.2	0.2
1999	1	4	5	5	5	6	6	6	0.45	0.45	0.45	1.0	1.0	1.0
2000	3	2	5	5	5	7	7	7	0.42	0.42	0.42	1.0	1.0	1.0
2001	2	3	3	3	2	0	0	1	1.00	1.00	0.67	0.6	0.6	0.4
2002	5	0	5	3	2	5	7	8	0.50	0.30	0.20	1.0	0.6	0.4
2003	5	0	5	4	4	11	12	12	0.31	0.25	0.25	1.0	0.8	0.8
2004	2	3	5	5	5	6	6	6	0.45	0.45	0.45	1.0	1.0	1.0
2005	1	4	5	5	4	5	5	6	0.50	0.50	0.40	1.0	1.0	0.8
2006	5	0	5	5	3	10	10	12	0.33	0.33	0.20	1.0	1.0	0.6
2007	5	0	5	5	3	8	8	10	0.38	0.38	0.23	1.0	1.0	0.6
2008	5	0	5	5	4	7	7	8	0.42	0.42	0.33	1.0	1.0	0.8
2009	5	0	5	5	2	2	2	5	0.71	0.71	0.29	1.0	1.0	0.4
2010	5	0	5	5	4	2	2	3	0.71	0.71	0.57	1.0	1.0	0.8
2011	5	0	5	4	2	6	7	9	0.45	0.36	0.18	1.0	0.8	0.4
2012	5	0	4	4	1	6	6	9	0.40	0.40	0.10	0.8	0.8	0.2
2013	5	0	5	5	4	7	7	8	0.42	0.42	0.33	1.0	1.0	0.8
2014	5	0	5	4	1	4	5	8	0.56	0.44	0.11	1.0	0.8	0.2
2015	2	3	4	4	2	1	1	3	0.80	0.80	0.40	0.8	0.8	0.4
min	0	0	1	1	1	0	0	1	0.25	0.25	0.08	0.20	0.20	0.20
max	5	5	5	5	5	11	12	12	0.19	0.20	0.16	0.29	0.29	0.30
mean	3.19	1.81	4.38	4.14	2.81	5.29	5.52	6.86	0.49	0.47	0.31	0.88	0.83	0.56
σ	2.01	2.01	1.13	1.12	1.40	2.88	3.00	3.04	0.18	0.19	0.15	0.23	0.22	0.28

Recall for daily predictions was very high for all 21 years except for three: 1996, 1997, and 1998. Recall was 1.0 for all but 6 of our 21 test years. Of these six test years, two had recalls of $0.8 = 4/5$ (2012 and 2015), two had recalls of $0.6 = 3/5$ (1996 and 2001), one had recall of $0.4 = 2/5$ (1997) and finally one had recall of just $0.2 = 1/5$ (1998).

As expected, making daily predictions yielded the highest precision and recall, compared to targeting a smaller time-frame of prediction (*e.g.* 3 h or 1 h). Performance was not as good for the 3-Hr and 1-Hr peak predictions methods, which were preliminary attempts to hone in on smaller prediction time-frames. The authors have yet to investigate all possible methods of targeting smaller prediction time-frames, and are hopeful that further study will be fruitful.

Table 3. Precision and Recall averaged across 7 years from **2007 to 2013**, inclusively, for the three different prediction types. σ is standard deviation. The results of the daily predictions of the "Probabilistic" algorithm from literature [8] are displayed for comparison. The green color scale was generated in Microsoft Excel and is between 0.0 and 1.0

	Precision				Recall			
	Daily	**Probab-ilistic**	3-Hr	1-Hr	**Daily**	**Probab-ilistic**	3-Hr	1-Hr
min	**0.38**	**0.40**	0.36	0.10	**0.8**	**0.8**	0.8	0.2
max	**0.71**	**0.71**	0.71	0.57	**1.0**	**1.0**	1.0	0.8
mean	**0.50**	**0.55**	0.49	0.29	**0.97**	**0.94**	0.94	0.57
σ	**0.14**	**0.11**	0.14	0.14	**0.07**	**0.11**	0.09	0.22

Additionally, precision and recall of the Daily predictions were averaged for years 2007 to 2013 to help compare with the results in literature [8], and these results are show in Table 3. As previously mentioned, the "Probabilistic" algorithm in literature made daily predictions and was tested on data for fiscal years 2007 to 2013 [8]. The average precision, recall, and their corresponding standard deviations of the "Probabilistic" algorithm are shown in Table 3 for comparison with our Daily peak predictions. In Table 3, the minimum and maximum precision values were estimated from the plots given in literature, and the fact that their minimum and maximum precisions as fractions must have numerators (TP in Eq. 12) of 4 or 5.

The current model, had slightly higher recall and slightly lower precision in comparison to the model in literature. Furthermore, both the Daily and 3-Hr prediction methods had high recall for many years for which winter peaks existed. Years with winter peaks were not tested in literature.

4 Discussion

In making daily peak predictions, the Naive Bayesian Classification algorithm had a precision and recall that is comparable to work by previous authors [8]. The model was tested on many more years, however, and was trained and tested on years with winter peaks, which was not previously done. While the results are good, there is still potential for even better performance.

4.1 Low Recall for 1996, 1997, 1998

Low recall (0.6 or less) for the test years 1996, 1997, and 1998 likely cannot be attributed to whether the low recall test years had more winter peaks since both summer and winter peaks occurred commonly during many years for which the model performed with perfect recall. Most likely the issue is related to the normalization procedure described in Sect. 2.3. The procedure was carried out to eliminate the effects of large differences in baseline demand from year to year, and help line up the demand thresholds from one year to the next.

The point was to decrease the variance in the demand and increase the separation of the peak and non-peak distributions. For the testing years in question, this procedure may have been insufficient, or it may have inadvertently shifted the testing years' demand thresholds away from the average demand threshold of the training years.

4.2 Demand Threshold Prediction

The current model does not predict a demand threshold, and the authors believe that incorporating an adaptive estimation of the demand threshold (i.e. the midpoint between the 5th and 6th top demand) may solve the low precision and recall problems during a few of the years.

Most of the algorithms tested in literature are demand threshold prediction algorithms [8], and any such algorithms can be incorporated to the current one in the following way. Every time a demand is being tested, one could shift the test demand value by the same amount one would shift the current demand threshold to line it up with the intersection of the *peak* and $peak^c$ PDFs in Fig. 1d. The more accurate the estimation, the better the model would perform. Even if the estimation is not perfect, the other attributes may be just extreme enough to push the classification to be a peak (or mild enough to push the classification to be a non-peak). Even without a threshold demand prediction incorporated into the model, the model works fairly well. Thus, any demand threshold predictions that works better than no prediction, would likely improve the current model.

If there exists an algorithm that can predict the threshold perfectly, then the problem is solved completely. However, this is very unrealistic and the best one can do is predict the threshold as well as possible. Then, in the case of the Naive Bayesian algorithm, use other attributes to take the final step in deciding if a peak will occur.

4.3 Testing on Years with Winter Peaks

The current model benefited from experimenting and training on years with winter peaks such as the years previous to 2006 as well this past fiscal year ending in April 2015. It has been tested and shown to work well for many years with summer and/or winter peaks. The previous authors did not train nor test on years with winter peaks because predicted day ahead Ontario demands were only available from 2006 to 2013 when winter peaks did not occur [8].

4.4 Testing on Actual Versus Predicted Demand Data

A limitation of the current study is that the current model was tested on actual demand data, due to the authors being unable to acquire daily historical *predicted* demand (*i.e.* the *14-DayAhead-Dataset*). Our Daily Prediction method (see Subsect. 2.5) may perform poorer using such demand data. In that regard, we commend the work of the previous authors who managed to test their model on day ahead predicted demand [8].

Due to using actual demand data for testing, the 3-Hr and 1-Hr peak predictions were designed to not utilize the knowledge of the highest hourly demand of the day. If the 14-DayAhead-Dataset can be used for testing, then knowledge of the predicted maximum demand hour can be used. This would likely increase the precision and recall of the 3-Hr and 1-Hr prediction methods closer to that of the Daily prediction method.

5 Conclusion and Future Work

The Naive Bayesian Classification Model was successfully trained and tested 21 times. The mean precision and recall of Daily classifications were 0.49 (0.18) and 0.88 (0.23), respectively[2]. Predicting whether a peak will occur in a smaller time-frame is more difficult. When the model was used to choose three hours of a potential peak day as candidates for the top 5 yearly peak hours, the mean precision and recall were 0.47 (0.19) and 0.83 (0.22), respectively. The results are promising and with some additional work, the authors are confident that the model's statistical evaluation metrics will improve significantly.

6 Future Work

Besides incorporating a demand threshold prediction algorithm to the current model, further testing of the model will involve acquiring and testing on predicted demand data. At the very least, testing will use estimated predicted demand, which can be derived by considering the standard deviation of predicted demand with respect to actual demand. Past weekly and current daily predicted demand can be downloaded easily from the IESO. Once such testing is possible, we will also optimize methods of predicting peaks within the smaller time-frames of one, two, and three hours.

This study is part of a much larger project that will involve integrating many different algorithms to a common cloud based big data server and visual analytics system. One of goals of the project will be to help different types of Ontario energy consumers reduce peak energy usage as well as decrease overall energy consumption.

Acknowledgments. This work was supported by a Collaborative Research and Development (CRD) grant from the Natural Sciences and Engineering Research Council of Canada (NSERC), grant number 461882-2013. The authors of this work thank their collaborative development partner Fuseforward Solutions Group. The authors also thank Ricky Fok for suggesting the use of Kristan *et al.*'s kernel density estimation algorithm.

[2] Standard deviation is in brackets.

References

1. IESO Data Directory. http://www.ieso.ca/Pages/Power-Data/Data-Directory.aspx. Accessed 27 July 2015
2. Office of the Auditor General of Ontario - Report on Renewable Energy initiatives. http://www.auditor.on.ca/en/reports_en/en11/303en11.pdf. Accessed 30 June 2015
3. An, X., Cercone, N.: A statistical model for predicting power demand peaks in power systems. In: Proceedings of the 12th International Conference on Fuzzy Systems and Knowledge Discovery, pp. 1022–1026 (2015)
4. Environment Canada: Climate Data - Environment Canada, October 2011. http://climate.weather.gc.ca/. Accessed 01 July 2015
5. Gallagher, S.H.: Electric Demand Response Tariffs Clean-Up. http://www.pge.com/nots/rates/tariffs/tm2/pdf/ELEC_3221-E.pdf. Accessed 06 July 2015
6. IESO: IESO Global Adjustment. http://www.ieso.ca/Pages/Ontario%27s-Power-System/Electricity-Pricing-in-Ontario/Global-Adjustment.aspx. Accessed 30 June 2015
7. IESO: IESO Global Adjustment, for Class A customers. http://www.ieso.ca/Pages/Participate/Settlements/Global-Adjustment-for-Class-A.aspx. Accessed 30 June 2015
8. Jiang, Y.H., Levman, R., Golab, L., Nathwani, J.: Predicting peak-demand days in the Ontario peak reduction program for large consumers. In: Proceedings of the 5th International Conference on Future Energy Systems, e-Energy 2014, pp. 221–222. ACM, New York (2014). http://doi.acm.org/10.1145/2602044.2602076, Accessed 30 June 2015
9. Kristan, M., Leonardis, A., Skoaj, D.: Multivariate online kernel density estimation with Gaussian kernels. Pattern Recognit. **44**(1011), 2630–2642 (2011). http://www.sciencedirect.com/science/article/pii/S0031320311001233, Accessed 09 July 2015
10. MATLAB: version 7.8 (R2009a). The MathWorks Inc., Natick (2009)
11. Russell, S., Norvig, P.: Artificial Intelligence: A Modern Approach. Prentice-Hall, Englewood Cliffs, New Jersey (1995)