Demand-side Management of Smart Appliance Based on Optimization Prediction Tools

Project proposal for University of Virginia SYS 6014 Decision Analysis Spring 2020

Richard Liu
tl7fha
University of Virginia
April 25 2020

***Abstract*—Demand-side management is one approach that can be applied to efficiently manage a site's energy consumption with the aim of cutting the costs of electrical energy supply. Based on this approach, we try to build an optimization tool for smart appliance to save energy cost for consumers. We collect historic prediction and real-time price data from NYISO to build models for our decision basis. To prove our model's reliability, we examine its performance and examine our assumptions, the test result shows that our optimization tool can save cost. In the future, we will make calibration to improve our model and integrate our model to make better prediction. We are hoping that we can approach the best optimal decision for the smart appliance to work step by step.**

***Index Terms*—Demand-side Management, Optimization Tool, Prediction Tool, Decision Analysis, Energy Saving.**

I. Introduction

Demand-side management is one approach that can be applied to efficiently manage a site's energy consumption with the aim of cutting the costs of electrical energy supply. Its general idea is to take group actions for grid and general system to save electricity charges for all people.

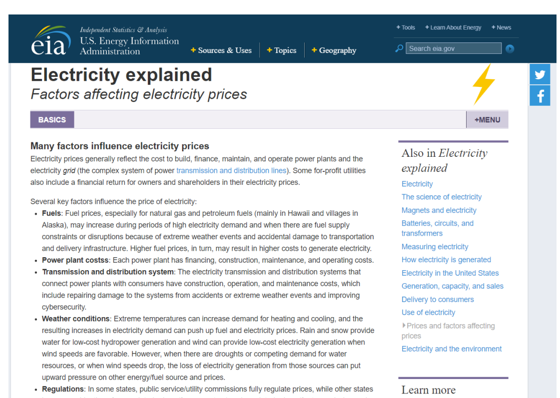
In real-world situation, the requirements of demand-side management is very straight-forward. We are hoping that consumers can choose to use less electricity during peak hours themselves so that we can reach a balance between the demand and the supply to reduce energy cost.

But the problem here is even if consumers acknowledge the benefit of demand-side management, is hard for themselves to implement this process, because of the limitation of technologies and knowledge. So there are several ways developed from energy companies' perspective to realize this goal. One way is to storage electricity when demand is relatively low and allocate them to consumers during peak hours. Another way is to manage energy generation process of grid operators so that demand and supply can be balanced.

These measures are widely applied and are making great benefits to the world energy saving. However, there are still lots of space for consumers themselves to take part in this decision process. As we mentioned before, it's hard for them to implement it directly. As a result, we will introduce smart appliance which allows working interactively and autonomously in this scenario. For example, your washing machine have a job to be done today. If a consumer don't need his washing job to be done as soon as possible, then smart appliance can hold the job and do it when electricity price is relatively low. Based on this idea, not only energy charges are saved but also the supply pressure is released.

In this scenario, we are trying to develop a simple way for consumers to deploy demand-side management. All consumers need to do is set up a working time range they need for smart appliance to complete the job. And smart appliance will decide its starting working time according to current flowing price. They don't need to know anything about the latent procedure of how smart appliance solves it out. Consumers only need to focus on their own needs and the flexibility of working time period they can provide smart appliance with so it's easy to perform this measure.

Now the problem switches to: How to get the value of flowing electricity price? The answer is: We need to build a prediction model for it. There are many factors that influence electricity prices, including fuel price, power plant costs, transmission and distribution system effect, weather condition and so on. As a result, building a model ourselves is too complicated and not reliable for us due to the large number of involved factors and lack of data access of these factors.


Independent Statistics & Analysis
U.S. Energy Information Administration

+ Sources & Uses | + Topics | + Geography | + Tools | + Learn About Energy | + News | Search eia.gov

Electricity explained

Factors affecting electricity prices

BASICS | +MENU

Many factors influence electricity prices

Electricity prices generally reflect the cost to build, finance, maintain, and operate power plants and the electricity *grid* (the complex system of power transmission and distribution lines). Some for-profit utilities also include a financial return for owners and shareholders in their electricity prices.

Several key factors influence the price of electricity:

- **Fuels**: Fuel prices, especially for natural gas and petroleum fuels (mainly in Hawaii and villages in Alaska), may increase during periods of high electricity demand and when there are fuel supply constraints or disruptions because of extreme weather events and accidental damage to transportation and delivery infrastructure. Higher fuel prices, in turn, may result in higher costs to generate electricity.
- **Power plant costss**: Each power plant has financing, construction, maintenance, and operating costs.
- **Transmission and distribution system**: The electricity transmission and distribution systems that connect power plants with consumers have construction, operation, and maintenance costs, which include repairing damage to the systems from accidents or extreme weather events and improving cybersecurity.
- **Weather conditions**: Extreme temperatures can increase demand for heating and cooling, and the resulting increases in electricity demand can push up fuel and electricity prices. Rain and snow provide water for low-cost hydropower generation and wind can provide low-cost electricity generation when wind speeds are favorable. However, when there are droughts or competing demand for water resources, or when wind speeds drop, the loss of electricity generation from those sources can put upward pressure on other energy/fuel source and prices.
- **Regulations**: In some states, public service/utility commissions fully regulate prices, while other states have a combination of unregulated prices (for generators) and regulated prices (for transmission and

Also in *Electricity explained*
Electricity
The science of electricity
Magnets and electricity
Batteries, circuits, and transformers
Measuring electricity
How electricity is generated
Electricity in the United States
Generation, capacity, and sales
Delivery to consumers
Use of electricity
Prices and factors affecting prices
Electricity and the environment

Learn more
Electric Power Monthly

Fig. 1: Factors that influence electricity price

Instead, we will be using the forecast electricity price data generated by New York independent system operator(NYISO) as our basis to predict the realized price of electricity. We will use historic real-time and forecast price data to generalize a model that can fit both data. Then we can predict the realized price one-day ahead according to the day-ahead predicted data.

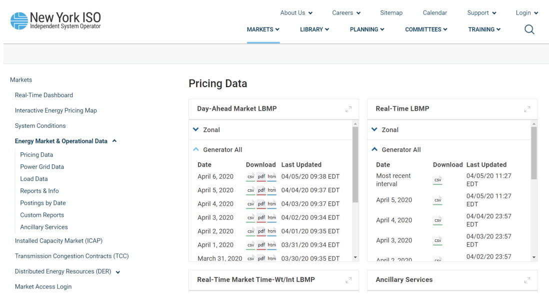


Fig. 2: NYISO Website

The result of the real-time updated data will be used to refine the model. Based on that predicted realized price value and the given working hour range, we can find the optimal working time for the smart appliance.

II. The decision problem and Formalism

A. Decision Problem

The decision-maker should be the smart appliance itself, to decide which time to start doing the job in order to get the lowest cost of electricity based on predicted realized LBMP in a typical day.

The analysis here uses statistical decision theory to generate a recommendation about which time to start working, in order to get a best potential price to save money.

B. Formalism

Parameters we have here:

H: length of the given working time range

L $\leq$ H :working hours needed to complete the work

Y=$< Y_H, Y_H-, ..., Y_1 >$: price on hours h= H,...,1

The decision maker's action set $\mathbb{H}$: choose a starting working time (certain hour h to start working) to complete the job. $h \in [1, H-L+1]$(range of left available starting working hours).

Uncertainty here is the arrangements to arrange working time for smart appliance according to the predicted price data one day ahead before the real price is known. Let Y denote the random electricity price. In this case, we represent the randomness of unobserved thing as the error distribution of the linear model we build (the fluctuation of price).

Payoffs: The loss function Let $L(h, y)$ denote the losses incurred by electricity usage by smart appliance.

$$L(h, y) = \sum_{h}^{h+L-1} Y_h \tag{1}$$

Decision criterion The decision criterion for smart appliance is: choose the start working hour that minimizes losses in expectation:

$$h^* = \operatorname*{argmin}_{h \in \mathbb{H}} E[L(h, Y)] \tag{2}$$

Where the expectation is taken over the distribution of Y. This distribution represents the information about uncertain electricity price we can predict using the predicted price data for smart appliance to make decision to work.

III. Data generating process

A. Raw Data Description

Instead of building a model ourselves, we will use the forecast electricity price data generated by NYISO (NewYork Independent System Operator).

- The source of data: https://www.nyiso.com/energy-market-operational-data - LBMP: locational based marginal pricing

1	Time Stamp	Name	PTID	LBMP ($/MWHr)	Marginal Cost Losses ($/MWHr)	Marginal Cost Congestion ($/MWHr)
2	11/01/2019 00:00	CAPITL	61757	20.18	0.38	-14.37
3	11/01/2019 00:00	CENTRL	61754	7.4	0.09	-1.88
4	11/01/2019 00:00	DUNWOD	61760	16.45	0.52	-10.51
5	11/01/2019 00:00	GENESE	61753	6.84	0.03	-1.39
6	11/01/2019 00:00	H Q	61844	5.16	-0.17	0.09
7	11/01/2019 00:00	HUD VL	61758	16.29	0.45	-10.42
8	11/01/2019 00:00	LONGIL	61762	16.59	0.66	-10.51
9	11/01/2019 00:00	MHK VL	61756	7.64	0.07	-2.14
10	11/01/2019 00:00	MILLWD	61759	16.49	0.48	-10.59
11	11/01/2019 00:00	N.Y.C.	61761	16.53	0.59	-10.51
12	11/01/2019 00:00	NORTH	61755	5.03	-0.4	0
13	11/01/2019 00:00	NPX	61845	17.9	0.4	-12.08
14	11/01/2019 00:00	O H	61846	7.16	0.02	-1.71
15	11/01/2019 00:00	PJM	61847	9.06	0.16	-3.47
16	11/01/2019 00:00	WEST	61752	7.34	0.11	-1.8
17	11/01/2019 01:00	CAPITL	61757	17.69	0.34	-12.28
18	11/01/2019 01:00	CENTRL	61754	6.76	0.08	-1.61
19	11/01/2019 01:00	DUNWOD	61760	14.52	0.48	-8.97
20	11/01/2019 01:00	GENESE	61753	6.27	0.02	-1.18
21	11/01/2019 01:00	H Q	61844	5	-0.16	-0.09
22	11/01/2019 01:00	HUD VL	61758	14.38	0.41	-8.9
23	11/01/2019 01:00	LONGIL	61762	14.74	0.65	-9.02
24	11/01/2019 01:00	MHK VL	61756	6.95	0.05	-1.83
25	11/01/2019 01:00	MILLWD	61759	14.55	0.44	-9.05
26	11/01/2019 01:00	N.Y.C.	61761	14.59	0.54	-8.98
27	11/01/2019 01:00	NORTH	61755	4.68	-0.39	0
28	11/01/2019 01:00	NPX	61845	15.94	0.38	-10.49
29	11/01/2019 01:00	O H	61846	6.74	0.04	-1.63
30	11/01/2019 01:00	PJM	61847	8.36	0.15	-3.14

Fig. 3: Forecast Electricity Price Data

	A	B	C	D	E	F	G	H
1	Time Stamp	Name	PTID	LBMP ($/MWHr)	Marginal Cost Losses ($/MWHr)	Marginal Cost Congestion ($/MWHr)		
2	11/01/2019 00:05:00	CAPITL	61757	19.46	1.14	0		
3	11/01/2019 00:05:00	CENTRL	61754	18.36	0.04	0		
4	11/01/2019 00:05:00	DUNWOD	61760	19.89	1.58	0		
5	11/01/2019 00:05:00	GENESE	61753	17.83	-0.49	0		
6	11/01/2019 00:05:00	H Q	61844	17.93	-0.39	0		
7	11/01/2019 00:05:00	HUD VL	61758	19.73	1.41	0		
8	11/01/2019 00:05:00	LONGIL	61762	20.28	1.96	0		
9	11/01/2019 00:05:00	MHK VL	61756	18.47	0.15	0		
10	11/01/2019 00:05:00	MILLWD	61759	19.82	1.5	0		
11	11/01/2019 00:05:00	N.Y.C.	61761	20.12	1.8	0		
12	11/01/2019 00:05:00	NORTH	61755	17.66	-0.66	0		
13	11/01/2019 00:05:00	NPX	61845	17.93	1.23	1.62		
14	11/01/2019 00:05:00	O H	61846	17.4	-0.92	0		
15	11/01/2019 00:05:00	PJM	61847	18.15	-0.17	0		
16	11/01/2019 00:05:00	WEST	61752	17.51	-0.81	0		
17	11/01/2019 00:10:00	CAPITL	61757	19.56	1.18	0		
18	11/01/2019 00:10:00	CENTRL	61754	18.38	0	0		
19	11/01/2019 00:10:00	DUNWOD	61760	19.96	1.58	0		
20	11/01/2019 00:10:00	GENESE	61753	17.85	-0.53	0		
21	11/01/2019 00:10:00	H Q	61844	17.99	-0.39	0		
22	11/01/2019 00:10:00	HUD VL	61758	19.79	1.42	0		
23	11/01/2019 00:10:00	LONGIL	61762	20.35	1.97	0		
24	11/01/2019 00:10:00	MHK VL	61756	18.54	0.17	0		
25	11/01/2019 00:10:00	MILLWD	61759	19.89	1.51	0		
26	11/01/2019 00:10:00	N.Y.C.	61761	20.18	1.8	0		
27	11/01/2019 00:10:00	NORTH	61755	17.7	-0.68	0		
28	11/01/2019 00:10:00	NPX	61845	18.01	1.25	1.62		
29	11/01/2019 00:10:00	O H	61846	17.4	-0.97	0		
30	11/01/2019 00:10:00	PJM	61847	18.18	-0.2	0		

Fig. 4: Real-time Electricity Price Data

Forecast electricity price data: every hour's LBMP based on forecast everyday(every hour's price, one-day ahead).

Historical real-time electricity price data: real-world recorded LBMP(every 5 minutes' price, continuously updated to today) we use the mean value of 12 every 5 minutes' price within each hour to represent its every hours' real-time electricity price.

B. The Predictive Model

$X1_h$: variable indicating realized value of historic predicted electricity price.

$Y1_h$: variable indicating realized value of historic real-time electricity price.

$X2_h$: variable indicating realized value of day-ahead predicted electricity price.

Y_h: random state variable indicating realized value of electricity price.

We will use the historic data (both predicted price and real-time price $X1_h$ and $Y1_h$) collected from the website to build a linear model to make prediction about the realized price(get distribution of Y_h) distribution given new day-ahead predicted price($X2_h$).

The process can be concluded as follows: First, we need to get historic predicted price($X1_h$) and historic real-time price($Y1_h$). Then we need to build a linear model and get its error distribution($Y1_h$= $\beta_0 + \beta_1 X1_h + \varepsilon_i$). Finally, given the day-ahead predicted price($X2_h$), we can predict the distribution of the realized price(Y_h).

That is, given a new predicted price data by the website one day ahead, we will have a prediction about the distribution of the realized price. We will use the mean value of every 5 minutes' real-time price within each hour to represent the every hour's real-time price($Y1_h$). And to match it with the predicted every hour's price($X1_h$) in the same location to build the model.

IV. Results

A. *Results and Analysis*

We are using historical real-time price data and prediction price data from February to March to build the linear model. The summary shows that the model is reliable which follows our assumption that these two variables(preprice and realprice) have a linear relationship.

```
Call:
lm(formula = realprice ~ preprice, data = df2)

Residuals:
    Min      1Q  Median      3Q     Max
-27.587  -3.428  -0.966   1.580 179.840

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.00368    0.18935   15.86   <2e-16 ***
preprice     0.82073    0.01088   75.46   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 7.892 on 21598 degrees of freedom
Multiple R-squared:  0.2086,	Adjusted R-squared:  0.2086
F-statistic:  5695 on 1 and 21598 DF,  p-value: < 2.2e-16

    Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
-27.5874  -3.4278  -0.9663   0.0000   1.5803 179.8402
```

Fig. 5: Linear Model Summary

When looking at the residual plot of the linear model we built, it seems like the error term is symmetrically distributed and there is no explicit pattern for the error. Our assumption of normal distribution for error term might be feasible.

Then based on our assumption that error term is normally distributed, we try to get the distribution of it. The plot shows that the distribution is matched with our assumption.

After examining the error distribution and model reliability, next step is using our prediction data to choose the optimal working time for consumers. We write a program in Python to simulate this process.

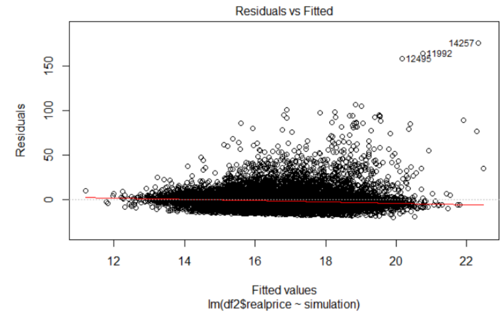


Fig. 6: Residual Plot

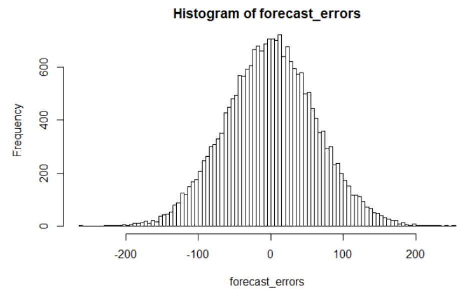


Fig. 7: Error Distribution

As we mentioned before, after inputting the date, the location work hours needed for the job and setting up working time range. The optimization tool will give its suggestion on which hour to start working and the cost it can save. We try several samples to test its performance.

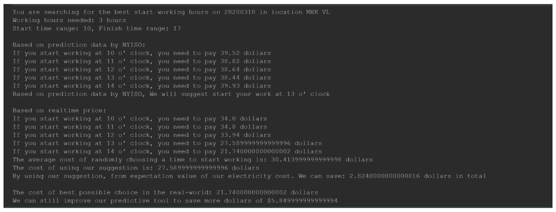
```
You are searching for the best start working hours on 20200314 in location HUD VL
Working hours needed: 1 hours
Start time range: 14, Finish time range: 20

Based on prediction data by NYISO:
If you start working at 14 o' clock, you need to pay 13.14 dollars
If you start working at 15 o' clock, you need to pay 13.31 dollars
If you start working at 16 o' clock, you need to pay 14.72 dollars
If you start working at 17 o' clock, you need to pay 16.74 dollars
If you start working at 18 o' clock, you need to pay 20.44 dollars
If you start working at 19 o' clock, you need to pay 21.73 dollars
Based on prediction data by NYISO, We will suggest start your work at 14 o' clock

Based on realtime price:
If you start working at 14 o' clock, you need to pay 11.95 dollars
If you start working at 15 o' clock, you need to pay 11.95 dollars
If you start working at 16 o' clock, you need to pay 11.66 dollars
If you start working at 17 o' clock, you need to pay 12.04 dollars
If you start working at 18 o' clock, you need to pay 12.92 dollars
If you start working at 19 o' clock, you need to pay 12.04 dollars
The average cost of randomly choosing a time to start working is: 12.093333333333334 dollars
The cost of using our suggestion is: 11.95 dollars
By using our suggestion, from expectation value of our electricity cost. We can save: 0.14333333333333442 dollars in total

The cost of best possible choice in the real-world: 11.66 dollars
We can still improve our predictive tool to save more dollars of $0.28999999999999915
```

Fig. 8: Model Performance For 1 Hour Work

```
You are searching for the best start working hours on 20200310 in location MHK VL
Working hours needed: 3 hours
Start time range: 10, Finish time range: 17

Based on prediction data by NYISO:
If you start working at 10 o' clock, you need to pay 39.52 dollars
If you start working at 11 o' clock, you need to pay 38.82 dollars
If you start working at 12 o' clock, you need to pay 38.64 dollars
If you start working at 13 o' clock, you need to pay 38.44 dollars
If you start working at 14 o' clock, you need to pay 39.93 dollars
Based on prediction data by NYISO, We will suggest start your work at 13 o' clock

Based on realtime price:
If you start working at 10 o' clock, you need to pay 34.0 dollars
If you start working at 11 o' clock, you need to pay 34.8 dollars
If you start working at 12 o' clock, you need to pay 33.94 dollars
If you start working at 13 o' clock, you need to pay 27.589999999999996 dollars
If you start working at 14 o' clock, you need to pay 21.740000000000002 dollars
The average cost of randomly choosing a time to start working is: 30.413999999999998 dollars
The cost of using our suggestion is: 27.589999999999996 dollars
By using our suggestion, from expectation value of our electricity cost. We can save: 2.8240000000000016 dollars in total

The cost of best possible choice in the real-world: 21.740000000000002 dollars
We can still improve our predictive tool to save more dollars of $5.849999999999994
```

Fig. 9: Model Performance For 2 Hours Work

As we can see, using the predictive tool's suggestion can help us save the cost compared with randomly doing the job. Even if it can't reach the best optimal cost at this point, in some cases the cost it saved can reach nearly 50 percent of the



```
You are searching for the best start working hours on 20200310 in location MHK VL
Working hours needed: 3 hours
Start time range: 10, Finish time range: 17

Based on prediction data by NYISO:
If you start working at 10 o' clock, you need to pay 39.52 dollars
If you start working at 11 o' clock, you need to pay 38.82 dollars
If you start working at 12 o' clock, you need to pay 38.64 dollars
If you start working at 13 o' clock, you need to pay 38.44 dollars
If you start working at 14 o' clock, you need to pay 39.93 dollars
Based on prediction data by NYISO, We will suggest start your work at 13 o' clock

Based on realtime price:
If you start working at 10 o' clock, you need to pay 34.0 dollars
If you start working at 11 o' clock, you need to pay 34.8 dollars
If you start working at 12 o' clock, you need to pay 33.94 dollars
If you start working at 13 o' clock, you need to pay 27.589999999999996 dollars
If you start working at 14 o' clock, you need to pay 21.740000000000002 dollars
The average cost of randomly choosing a time to start working is: 30.413999999999998 dollars
The cost of using our suggestion is: 27.589999999999996 dollars
By using our suggestion, from expectation value of our electricity cost. We can save: 2.8240000000000016 dollars in total

The cost of best possible choice in the real-world: 21.740000000000002 dollars
We can still improve our predictive tool to save more dollars of $5.849999999999994
```

Fig. 10: Model Performance For 3 Hours Work

overall optimal saved cost. However, the performance in some cases are not that good when price within the time range is similar or fluctuation is relatively low. Further improvements can be made to improve the model's performance.

B. Some assumptions we have for our model

First, it is supposed that there is a linear relationship between predicted price and real-time price. And the error term ε_i follows a normal distribution: $\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$.

Second, we are using the average of real-time 5 minutes' data to represent its each hours' price. Besides, when comparing the performance of choosing our suggestion's action and randomly choosing a starting working time, we use the average of all possible actions together to represent the performance of randomly starting working.

Finally, the process is assumed to be stationary. There is trend over time and months, years. So our prediction can be reliable.

V. Conclusions

A. Contributions

First, this project examines the linear relationship between real-time price and prediction price. The reliability of NYISO's prediction seems acceptable and we can make some adjustments to make it better.

To move further, this project examines the error distribution of our model, whether it follows our assumption of normal distribution. From our results, the error distribution of our model follows normal distribution and there seems no tendency that higher price brings higher fluctuation.

The biggest contribution is that we have built an optimization tool which can save cost for consumers. Even if it is a naive prototype, it kinds of realizing the thought of demand-side management. After inputting the date, the location, working time and working time range, it will suggest a working time for the user. And we are able to quantify the potential savings of our predictive tool.

An additional contribution is that in the process of using NYISO's data, we found that there are many errors in the raw data set. To implement our models, tools for aggregating data and data cleaning have been developed which can help detect the specific data file where its according error is located. It can be applied to NYISO's future data to make it follow the normalized format.

B. Future Improvements

There still needs to be process for getting better information. Although our model can save cost for consumers, it still can't reach the best optimal choice. After getting the real-time price data of the new day, We will fit it into the model again to get better results for the future.

Besides, we may also try to evaluate other electricity price prediction models in the same way to find a model with higher accuracy in predicting price. Moving forward, we may consider integrating the prediction model and introduce other variables to build a more concrete and solid model to predict electricity price rather than directly applying NYISO's prediction data.

Last, adding a working time preference variable to our loss function might be good to satisfy the consumers' need. For example, we may set a higher preference level value for consumers who would like their job to be done as soon as possible.

C. Conclusions

Demand-side management is really an effective idea for saving energy cost in today's society. Based on this idea, we try to build an optimization tool for smart appliance to save energy cost for consumers. We collect historic prediction and real-time price data from NYISO to build models for our decision basis. In addition to examining the reliability and error distribution of our model, the test result shows that our optimization tool can save cost. In the future, we will make calibration to improve our model and integrate our model to make better prediction. We are hoping that we can approach the best optimal decision for the smart appliance to work step by step.

VI. Reference

[1] Rui Tang, Shengwei Wang, Hangxin Li, Game theory based interactive demand side management responding to dynamic pricing in price-based demand response of smart grids, Applied Energy, 2019.

[2] Katia Gregio Di Santo, Silvio Giuseppe Di Santo, Renato Machado Monaro, Marco Antonio Saidel, Active demand side management for households in smart grids using optimization and artificial intelligence, Measurement, 2018.

[3] Nilotpal Chakraborty, Arijit Mondal, Samrat Mondal, Efficient Load Control Based Demand Side Management Schemes Towards a Smart Energy Grid System, Sustainable Cities and Society, 2020.