

Predicting Impact of Temperature Setpoint Adjustment on Household Energy Use Under Time Based Electricity Rate Plans

Mahesh Naidu and Henry Willem, Lawrence Berkeley National Laboratory

ABSTRACT

In summer months, households who have opted for time-based electricity plans are often faced with the challenge of choosing between limiting the use of air conditioners (ACs) to maximize the benefit of the plans and running the system continuously to meet their thermal comfort. One strategy that can be adopted to avoid higher bills would be to increase the setpoint temperature to reduce electricity consumption from AC operation when the electricity rate is high. However, AC users have little means to figure out, in advance, the impact of this adjustment or what the optimal setting should be. Past studies estimated annual savings of approximately ten percent if the thermostat could be set back 7-10 degrees F from its usual setting for eight hours daily. Climate factors, house/equipment types, occupants' behavior, and different rate plans influence the benefits of adjusting temperature setpoint. Therefore, a tool that determines savings is needed, so users can make decisions regarding adjusting their indoor temperature based on their price consciousness and thermal comfort needs. This paper presents a preliminary study of a data-driven model that predicts future energy use at various setpoint scenarios. A time-series prediction model is trained using the historical energy use data. Weather condition is included as an influencing factor, while indoor temperature is used in predicting impact of changing temperature setpoint. The model is then used to estimate monthly energy savings based on a time-of-use utility plan, to inform users in real-time about potential opportunities to adjust their temperature setpoint.

Introduction

Space heating and air-conditioning are the two major energy-consuming systems that account for over 51% of the total energy drawn by US households (EIA 2015). This is because they serve the basic needs of the occupants, i.e. thermal comfort, which is important for their health and well-being. Maintaining thermal comfort is also critical so the occupants can perform their daily activities in the indoor environment optimally. In the hot summer months, households rely heavily on air-conditioners, making it a major electrical expense. Air conditioning could alone lead to higher energy bills, especially during peak hours, if one opted for the time of use (TOU) plan.

Time of use (TOU) plans are implemented by utilities to control the heightened energy demand during peak hours. It incentivizes customers to shift some of their energy use activities from peak hours to off-peak hours. Through reduced rates, customers are encouraged to use high-energy appliances during off-peak hours and vice-versa. As much as it is important to save on electricity bills, it is also important to ensure thermal comfort. Therefore, switching-off air conditioning completely when it is needed the most is not an easy choice to make. In such situations, TOU customers tend to increase their setpoint temperature by some degrees. But with this approach, are the customers able to reduce their consumption substantially? The Department of Energy recommends turning thermostats back 7 to 10 degrees from the usual settings for 8

hours per day to achieve annual savings of up to 10% (DOE 2018). While setpoint adjustments could lead to a reduced electricity consumption, the impact is expected to vary depending on various aspects like building type, building material, climatic conditions, occupants' behavior, heating and cooling equipment types, etc. Therefore, it would be very useful and rewarding for the households to know the energy consumed by their ACs and how much they could save by increasing their set points while maintaining a comfortable indoor environment. With this perspective, this research is aimed at developing a real-time tool that can predict the AC energy consumption if setpoint temperature is increased.

In past studies, several methods and approaches were proposed to predict energy consumption of HVAC systems including air-conditioners in buildings. Some approaches involved development of mathematical (theoretical) models while others made use of data driven statistical models to predict energy use. Over the years, mathematical models representing physical and thermal processes involved in HVAC systems coupled with detailed computational techniques have been used to calculate equipment capacity and energy use requirements (McQuiston, Parker and Spitler 2004; ISO 13790:2008). Additionally, simulation tools such as EnergyPlus, BLAST, ESP-r, have also been designed to predict energy consumption of buildings or its subcomponents for their efficiency studies and analyses. Although these tools give relatively accurate estimations, their applications are limited or restricted in that they require high level details of the building and environment parameters. For example, Wang et al. (2013) presented an energy consumption calculation model to estimate the energy consumption of air conditioning systems in a data center at various temperature set-points. This mathematical model required a large number of measurements of many parameters specific to the cooling space, the building and its environment to compute the energy-use of the ACs at a given set-point. Such an approach may not be applied to all buildings and scenarios and is not practical because of the lack of detailed information and data; and the complex nature of its process. A data-driven energy use forecasting technique with a resulting model that utilizes minimal data inputs would be more practical for various field applications, including different building types and characteristics, as well as variations in indoor-outdoor conditions.

Statistical modeling approaches use simplified methods and fewer variables common to most buildings for making predictions, but they are known to be data intensive. Liao, Chang and Chang (2020) estimated monthly AC energy consumption of a building by using Machine Learning and Deep Learning models on selected factors such as air pressure, air temperature, floor area, and orientation. Their models produced expected results, however, the effectiveness of this tool to make real time prediction is subject to availability and processing time of a large number of influencing variables that have to be collected and made available for modeling in real-time making it a resource intensive and computation intensive process for individual households. The current research also uses historical data to train and develop predictive models but it aims to simplify the process by adopting a statistical time-series analysis directly on the target variables, i.e. energy use, to develop patterns that can be used to forecast the near-term future values. The current research used the actual time-based AC energy use-data of various households to develop time-series models unique to each household to predict the energy consumption patterns of the air-conditioner in real-time. Because past studies have shown strong dependence of AC energy use on weather conditions, this time-series analysis also considered external factors such as outdoor temperature. For example, Borowski and Zwolinska (2020) performed a comparative analysis of machine learning models that predicted energy consumption in buildings for cooling, using influencing factors such as weather variables, time,

and occupancy as inputs. Their analysis revealed that external temperature had an obvious influence on energy consumption, and thus determined the demand of cooling energy. Therefore, the outdoor air temperature data being a key factor determining energy usage, was also used in the model described in this study.

Using the predictive energy use models trained on real-time energy use data and influencing factors, future AC energy use can be forecasted or estimated if the set-point temperature is changed. Such an approach would allow households to select from a range of set points to suit their thermal comfort needs and budgetary constraints. Five households with existing data on their AC compressors energy use were selected for this study to understand the implications of this solution. Further, the analysis was extended to study the impact of TOU plans on adjusting the indoor temperature setpoints.

Methods

Air-conditioning operation and energy consumption is primarily driven by space heat gain, which is defined as the rate at which heat enters into and/or is generated within a space at a given instant. Heat can enter a house through various means, i.e. solar radiation through openings and windows, heat conduction through surfaces, heat generated within the space by occupants, lights, appliances, equipment and processes, and heat gained as a result of outdoor air exchanges and other miscellaneous sources. Solar radiation and heat conduction through walls and roofs are generally the dominant sources of heat gain in a home and may vary drastically within the day and from month to month. This means the outdoor air temperature variations and the resulting temperature differences with the indoor air temperatures are important determinants of heat gain and therefore, the cooling capacity of the AC systems. Other factors that influence AC energy use depend greatly on occupants' preferences and activities. Turning on lights, using heat generating devices or appliances, introducing warm outdoor air for ventilation, etc., all play a role in cooling load estimation for a home. However, their impact on heat gain is generally less than other sources, and given the lack of set patterns observed in these activities, it would be very challenging to effectively use them in the predictive tool. Therefore, the current approach assumes a minimal impact of the variations in heat generation by the occupants' activities and focuses primarily on measured time-series AC energy use data, as well as the indoor and outdoor temperatures of individual household units.

Four single-family homes (A,B,D,E) and one town home (C) situated in Texas were considered for this research. Data from the summer months of June, July, August in 2014 were used in the analysis for house A and house B, whereas data from the summer months of June, July, August in 2020 were used for houses C, D, and E. Summer months were selected because these are peak cooling months where consistent and prolonged use of the AC system is expected. Table 1 shows additional building and household characteristics of the selected sites. The information shows typical home sizes and occupancy and how closely correlated they are, i.e. house size increased with occupancy. All homes used a single AC system for their cooling. Single story homes used one thermostat, while two-story homes used two thermostatic controls.

Table 1. House characteristics of the selected sites

House	House age (years)	No. of floors	No. of rooms	House size (sq ft.)	No. of occupants	AC unit compressor	Indoor thermostat
A	13	2	7	1860	2	1	2
B	14	2	4	2800	4	1	2
C	13	2	9	1840	1	1	2
D	13	2	4	2600	2	1	2
E	53	1	9	2800	5	1	1

Figure 1 shows the typical measured weekly energy use pattern. While there are obvious day-to-day variations, the trend exhibits a repeatable pattern with energy use peaking around mid to late afternoon hours. The minor variations within each day are likely associated with occupancy rate and occupants' activities, while the more consistent daily patterns are closely associated with the daily weather variations, as shown by the other line in Figure 1. As discussed previously, heat conduction accounts for a substantial portion of overall heat gain in the home. As a result, outdoor air (ambient) temperature will have a strong effect on AC system operation time and thus, energy use. The compressor unit has to run more frequently to remove the excess heat that enters the indoor space as a result of temperature differences between indoors and outdoors to meet the preferred indoor temperature setpoint. It is therefore important to include the outdoor air temperature as a key determinant when analyzing and forecasting AC energy consumption. For outdoor air temperature data, the current research used historical weather data from weather stations in the region or city where the houses are located. To determine the temperature differences at any given time, the hourly outdoor air temperature data points were then paired with the measured indoor air temperature at each home.

A time series forecasting approach to predict future AC energy use based on indoor and outdoor air temperatures was developed. The method is discussed in the following section.

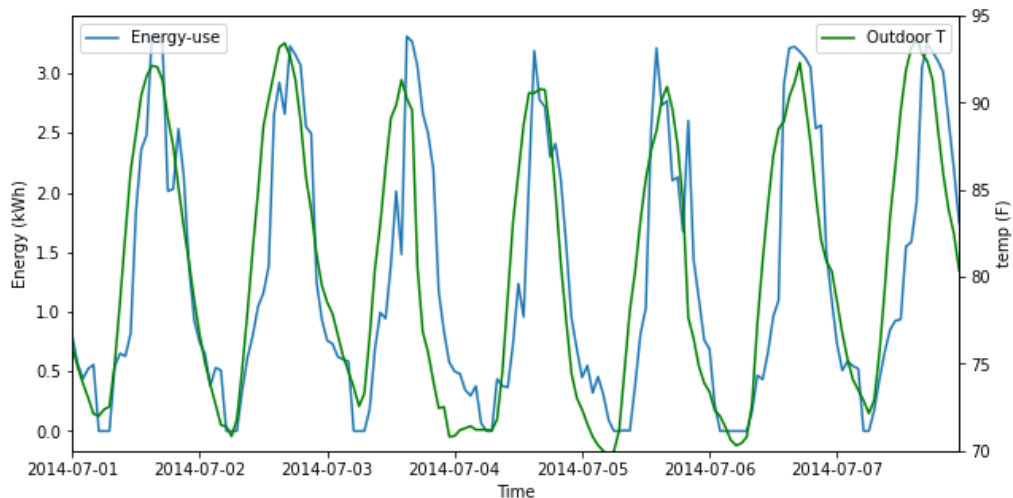


Figure 1. AC energy use and ambient (outdoor air) temperature profiles for a typical week

Development of time series models to predict energy use

The AC energy consumption data used in this research is time-dependent in nature and thus, analysis techniques exclusively used for time-series were chosen for in-depth analytics and forecasting future energy use. In this study, seasonal autoregressive integrated moving average (SARIMA) model was applied to the analysis. SARIMA belongs to the class of Autoregressive and Moving Average models, which are mathematically well-defined models generally favored in the domain of time series analysis. These comprehensive statistical models are able to efficiently predict a given time-series based on its own past values. SARIMA or seasonal ARIMA model was specifically chosen over its parent model ARIMA because it has the capabilities to detect seasonal patterns in a time-series (Oh and Morzuch, 2005). As noted above, the data that was analyzed in this study is also seasonal and repetitive in nature.

The ARIMA model in its various forms including SARIMA, has been extensively applied for forecasting and is well proved to be efficient. It is a combination of autoregressive (AR) model and the moving average (MA) model and is denoted by ARIMA (p,d,q) where p, d and q are the number of autoregressive terms, the degree of difference and the number of the moving average terms respectively. The model's forecasting equation is described below:

$$y_t = \mu + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} - \theta_1 e_{t-1} \dots - \theta_q e_{t-q} \quad (1)$$

where $y_{t-1}, y_{t-2}, \dots, y_{t-p}$ are past observations and $\phi_1, \phi_2, \dots, \phi_p$ are parameters of the AR model or coefficients of the AR terms; $e_{t-1}, e_{t-2}, \dots, e_{t-q}$ are past forecast errors and $\theta_1, \theta_2, \dots, \theta_p$ are parameters of the MA model or coefficients of the MA terms. μ is a constant.

The SARIMA model which incorporates seasonality in ARIMA, is denoted by SARIMA(p,d,q)(P,D,Q)m where P, D and Q are the number of seasonal autoregressive terms, seasonal degree of difference and seasonal moving average terms, respectively. 'm' is the number of time steps per season which is 24 for this data representing 24 hours in a day. The seasonal part of this model operates with the same structure as ARIMA, but across multiples of m (the number of periods in a season). The determination of the terms is essential to ensure high accuracy of the model.

To further enhance the accuracy of AC energy-use forecasting, ambient temperature was incorporated in the forecasting model. This was made possible by expanding the SARIMA model to the SARIMAX model, which allows additional exogenous variables in the analysis as predictors. SARIMA and ARIMA models are essentially regression models in which the dependent variable is the future value of the target variable and the independent variables are all lags of the dependent variable (Auto-regression) and/or lags of the errors in its prediction (Moving Average). Since ARIMA is a regression model, it is straightforward in principle to extend it by incorporating information from other leading indicators or exogenous variables. Such a model can be accomplished by simply adding one or more regressors to the previous forecasting equation, resulting in the equation shown below:

$$y_t - \phi_1 y_{t-1} \dots - \phi_p y_{t-p} = \mu - \theta_1 e_{t-1} \dots - \theta_q e_{t-q} + \beta(X_t - \phi_p X_{t-1}) \quad (2)$$

where X is the exogenous variable and β is the regression coefficient. It can be noted that this model is fitted to the errors of the regression of Y on X (i.e. the series Y minus beta X). In the

same manner, SARIMA with exogenous variables can be extended to generate the SARIMAX model. Past research has identified the improved performance of SARIMAX over SARIMA due to the inclusion of influencing factors as predictors. Chikobvu and Sigauke (2012) predicted daily peak electricity using SARIMA and SARIMAX models. Their analysis of the two models reported higher prediction accuracy for SARIMAX which included workdays and holidays as regressors. They also indicated that the time series model's performance can be further improved by including weather variables like temperature. In this research, the SARIMAX model is developed to forecast the energy consumption of ACs while accounting for the effect of external factors like temperature on the predicted variable.

Figure 2 depicts the modeling approach and steps used in this research. The time series analysis focused on the summer months of June, July and August. After checking the data for quality, the hourly time-series AC energy use data of the five households was divided into train and test datasets. The training dataset, used as input in the time-series model, consisted of hourly data points collected over a 30-day period. The trained model was then used to predict future energy use by the AC compressor on an hour-by-hour basis for the next 24 hours. For instance, hourly AC energy-use values collected from June 1 to June 30 were used to predict the hourly energy consumption for the following 24 hours on July 1. For the purpose of this paper, the process was repeated for 7 consecutive days so a one-week energy use estimation can be derived. This process is shown in Figure 3. All predicted values were then evaluated against the test dataset consisting of corresponding actual data values for the same hours.

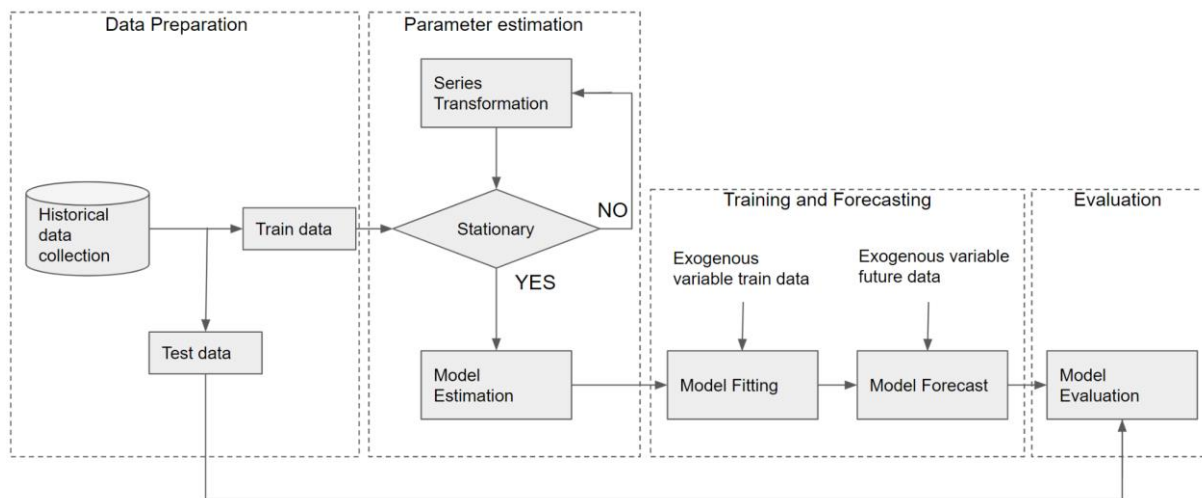


Figure 2. Flow chart of time series analysis methodology

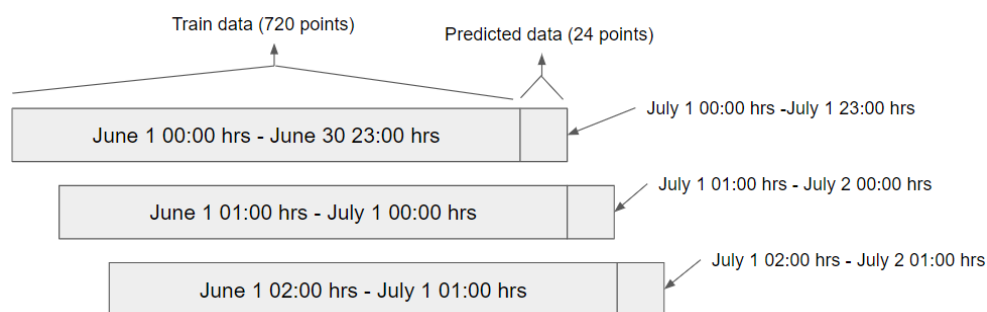


Figure 3. Data inputs and steps for predicting hourly energy use

This process was carried out for all households. Besides AC energy use data, ambient temperature data was also divided into training and testing data sets alongside AC energy use data for the same time intervals. The ambient temperature data was modeled with the AC energy use data in the SARIMAX model. The development of the SARIMAX model requires detailed specifications of the model composition. These specifications are the parameters of the model which are, number of autoregressive terms (p), degree of difference (d), number of moving average terms (q) as well as the terms P, D, Q which account for the seasonality in SARIMAX. Identifying the right model parameters is a crucial step in the modeling process for obtaining high prediction accuracy. Their values are estimated or learned from data and are initialized before the model is fitted. In the case of the SARIMAX model, prior to identifying the parameters, it is necessary to check whether the time series is stationary, that is it has constant mean, constant variance and there is presence of serial correlation. If not, then a data transformation procedure called ‘differencing’ is applied to the data once or twice in order to bring the data to stationarity. The one or two differencing is indicated in the model by the term ‘ d ’ (degree of difference). Stationarity of the data and the need for differencing is determined by performing the Augmented-Dickey fuller (ADF) test on the training dataset (Hyndman and Athanasopoulos, 2018). In this analysis, the ADF test was applied to the 1-month hourly AC energy-use datasets to check their stationarity. Seasonal differencing ($D = 1$), that is differencing after shifting 24 periods, precedes this step in order to remove any non-stationarity in data due to the seasonal component. If the ADF test determines that the training dataset is stationary, the parameter d is assigned the value 0. If non-stationary, then the series is differentiated once ($d = 1$) and the resultant series is tested again for stationarity. This process is repeated again if the series is not yet stationary. After attaining stationarity, autocorrelation function (ACF) and partial autocorrelation function (PACF) plots of the series were investigated to obtain the p, q and P, Q parameters (Box and Jenkins, 1970). The above techniques were applied to each training dataset before model fitting in order to obtain the optimal p, q, d and P, D, Q values which would be suitable for the datasets to provide better forecasting.

After determining the right model composition of the SARIMAX model, the models were fit using the time-series AC energy use data and the exogenous variable, i.e. ambient temperature. Model fitting and forecasting was followed by evaluation of the model’s prediction performance or accuracy. The predicted values were compared with the actual energy use values using two metrics: Coefficient of Variation of the Root Mean Square Error (CVRMSE) and Weighted Mean Absolute Percentage Error (WMAPE). For these statistical measures, lower values indicate better models.

Application of time series model to estimate energy savings at different set points

For given space heat gain, the ACs will respond to maintain the indoor temperature at the desired setpoint temperature by adjusting its runtime and cycling frequency. In the cooling season, when setpoint temperature is increased, the AC system will operate less because the cooling load that needs to be overcome is reduced. In the current research, the energy savings analysis related to changing setpoint temperatures focuses primarily on this effect and assumes that this process does not impact or is associated with other heat gain sources, whether constant, variable, or time-dependent. Based on this assumption, the SARIMAX modeling approach depicted in Figure 2 was re-applied using the energy use data from each home to forecast the hourly energy use when the setpoint is the same vs. when it is changed (increased) from the

existing temperature. In this case, instead of the ambient temperature, the difference between ambient air temperature and indoor temperature was used as the exogenous (predictor) variable.

With this adjusted model, households can estimate their future energy consumption at their current set-point as well as under a different scenario of increased set points. The energy use calculated at different set points can be estimated in real time at hourly intervals. With this information, the households would be able to make informed decisions regarding setting their AC setpoint temperature to their optimal and preferred conditions.

Households under TOU plans can significantly benefit from a prediction model that can inform them about their air-conditioner's energy use at different set points. By incorporating into the model electricity rates for peak and off-peak hours, households can estimate their savings if they want to increase their set point according to the rate schedule. To this end, estimates of savings when the set point is increased by 1 and 2 degrees is provided for each of the five households in the next section. Monthly savings were calculated based on a sample TOU rate schedule and the results from the respective households were analyzed. The rate schedule provided three electricity pricing tiers for use of electricity during different hours of the day. These hours are on-peak from 3 PM to 6 PM; mid-peak from 7 AM to 3 PM as well as 6 PM to 10 PM; and off-peak from 10 PM to 7 AM. No peak hours are scheduled for weekends. To calculate the energy and cost saved by increasing the set point from usual, the predicted energy consumed by the air conditioner set at a higher set point was subtracted from the predicted energy consumption at its normal setting, for a period of 24 hours. This produced the hourly AC energy use that could be possibly saved by the household if the set point is changed to a higher set point for 24 hours. Adding hourly AC energy saved during each hour gives the net energy saved during a typical day. The cost savings during the day were calculated by multiplying the energy saved during each of the on-peak, mid-peak and off-peak hours with their respective rates. If the household increases the temperature only during on-peak and mid-peak hours, then the energy-saved is obtained by deducting the predicted energy consumption at the higher set point from the predicted energy consumption at the normal setting for those hours only. The corresponding cost savings were calculated by multiplying the energy-saved with the respective rates. Thereafter, the net energy and percentage cost savings for all the days in a month were calculated and presented.

Results

Figure 4 shows the scatterplot of hourly AC energy use of house A as a function of ambient air temperature for one month. As expected, there is an apparent correlation between these variables, with subsequent regression analysis of energy use with air temperature as predictor showing a significant correlation (R-squared: 0.73, std. err. predictor: 0.002, std. err. const.: 0.181). To determine whether using the ambient air temperature as an exogenous variable would improve the time-series prediction accuracies, both SARIMA (without-) and SARIMAX (with-exogenous variable) models were developed. Figure 5 shows an example result of this analysis. The same hourly data over a period of 30 days were used as inputs to detect time series patterns and train the model for predicting energy use at each hour of the next 24 hours. Both models yielded reasonably good accuracies: CVRMSE values were 0.32 and 0.29 for SARIMA and SARIMAX models respectively while the WMAPE values were 0.24 and 0.21. The result shows that the SARIMAX generally predicts higher energy use than SARIMA models. Comparing the predicted models to the actual energy use for the same period reveals that

SARIMAX results are closer to the measured peak hourly energy use. The predicted models are also smoother than the actual measured data.

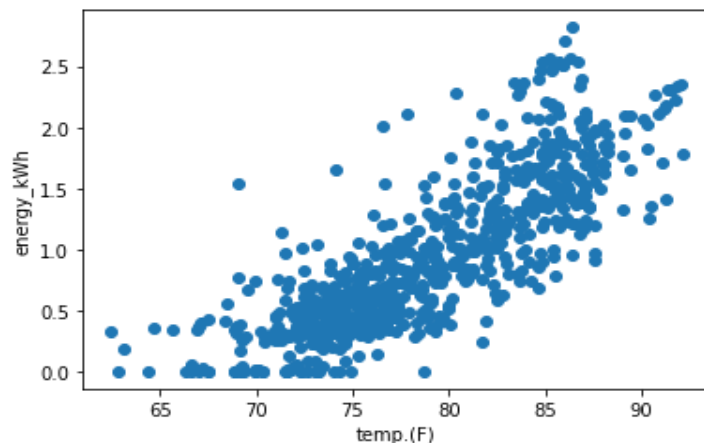


Figure 4. Hourly energy use as function of ambient air temperature

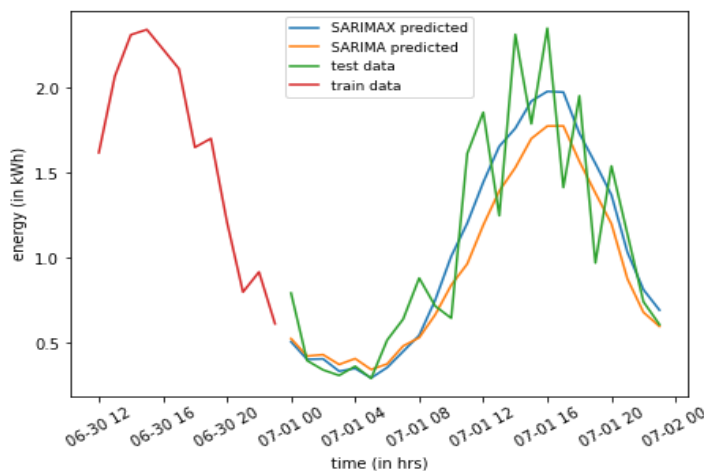


Figure 5. SARIMA and SARIMAX time series prediction model for 24 hours

In terms of model performance or accuracy metrics, i.e. CVRMSE and WMAPE, the analysis shows that SARIMAX offers slightly better accuracy than SARIMA. In four out of five homes, the CVRMSE and WMAPE of the SARIMAX models were equal or lower than the SARIMA models. However, the values are generally between 0.13-0.60 for both models, indicating relatively good predictions. Considering previous observations on the peak energy use prediction, the results presented in the subsequent section will be based on the SARIMAX models for all households.

For each of the households, the SARIMAX model forecasted the AC energy use for 7 consecutive days starting from July 31 up to August 6. The accuracy metrics of these models were in the range of 0.1-0.6 for CVRMSE and 0.1-0.4 for WMAPE, both suggesting a good accuracy level of the models. The predicted energy use for Houses A-E superimposed on the actual measured data is shown in Figure 6. Overall, there is a good agreement between the profiles or patterns of the predicted and the measured energy use. Looking in more detail, there were a few days in Houses B, C, and D specifically, where there was high variability in actual energy use and the predicted energy use could not mimic the variation accurately. It is possible

that during these days there were changes in occupancy, occupants' activities, and other heat gain sources that led to the rather unpredictable patterns of energy use. In Houses C and E, where the AC systems turned off occasionally in the early morning hours, the SARIMAX models were unable to match the turn off events, although the predicted patterns were reasonably similar. Further, there were sharp increases in energy use such as in Houses D and E that the SARIMAX models were unable to replicate. These variations are represented in Figure 7. This figure shows a summary box plots of hourly energy use differences between the predicted and measured data. The mean differences of all sites are close to zero and notable outliers can be observed for several houses (mean kWh $\pm$ st.dev): 0.001 $\pm$ 0.32 (House A), -0.086 $\pm$ 0.48 (House B), 0.042 $\pm$ 0.12 (House C), 0.12 $\pm$ 0.42 (House D), and -0.004 $\pm$ 0.46 (House E). This result confirms the ability of the time series model to accurately estimate energy use based on historical energy use data and the inclusion of an exogenous variable, i.e. ambient air temperature, even at hourly resolution. Current models were trained using 30-day data from a summer period. With more data as inputs over a longer period (multiyear), the model's prediction accuracy can be further improved.

Table 2 summarizes the percentage of energy savings for each household if they were to increase their temperature setpoint by 2 degrees F from the current daily setpoint for seven consecutive days. Assuming that heat gain is predominantly contributed by indoor and outdoor temperature difference, other heat gain factors remain the same, and impact of temperature variation is linear, the analysis result suggests a potential energy savings of 2.0-5.5% per degree increase of temperature over a 24-hour adjustment. The result shows that for some households increasing temperature may not be feasible given that their baseline setpoint temperature was already quite high, e.g. House B's daily setpoints were in the range of 77-78 degree F. The potential for energy savings varies among the five households. It can be observed that homes with lower setpoints (Houses A,D,E: 73-74F), tend to have greater energy saving potential than homes with higher setpoints (Houses B,C: 76-77F). Inter-day variability within each household demonstrates that energy savings can vary significantly from day to day depending on the weather conditions and other factors. The availability of a tool to predict the energy savings in real-time is therefore very useful to allow the household to make their preferred adjustments.

Table 3 presents the result of hourly analysis of energy use based on the prediction models at different temperature setpoints. Larger AC units (Houses A,B,D,E) saved about 3-4 kWh, 16-23 kWh, and 32-48 kWh per month, with one degree increase in setpoint for 3 hours (only weekdays), 15 hours (only weekdays), and 24 hours (weekdays and weekends) setpoint adjustment scenarios, respectively. Smaller AC unit (House C) saved about 1-7 kWh per month under the same scenarios. The increase in energy savings in kWh is almost linear with temperature change. The energy savings with two degree change in setpoint are approximately twice the savings from one degree change in setpoint. The trend in energy savings can be similarly observed in the electricity cost savings. Overall, the homeowner could save up to about 2.1-4.6% of their energy bill for cooling needs per month if they opt to increase setpoint by one degree F for all 24 hours of the day, or 0.4-1.0% per month if they opt to do the same for 3 peak hours (3-6 PM) during weekdays only.

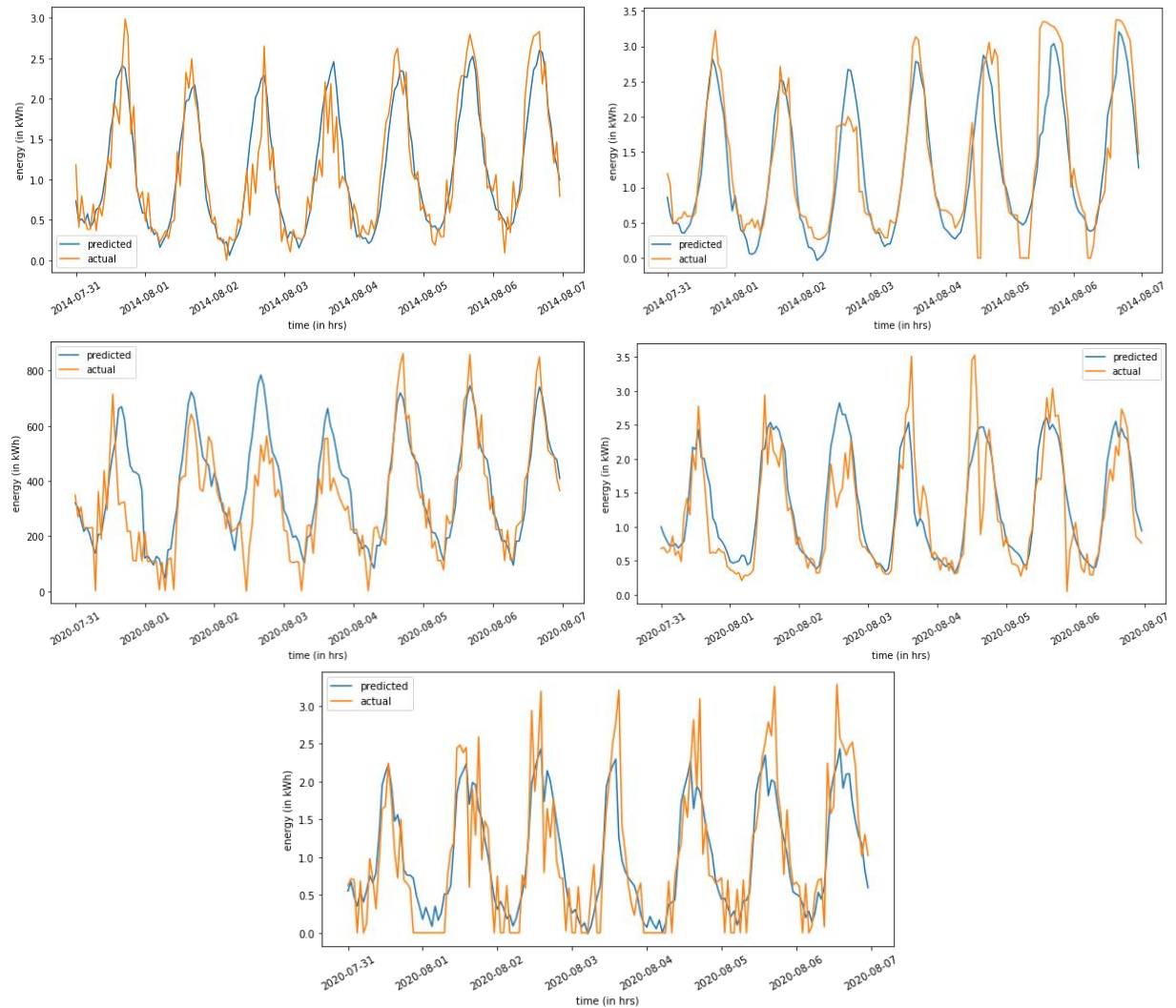


Figure 6. Hourly predicted vs. actual (measured) energy use over one week period for House A (top row-left), House B (top row-right), House C (middle row-left), House D (middle row-right), and House E (bottom row)
Note: x-axis tick marks indicate midnight (00:00)

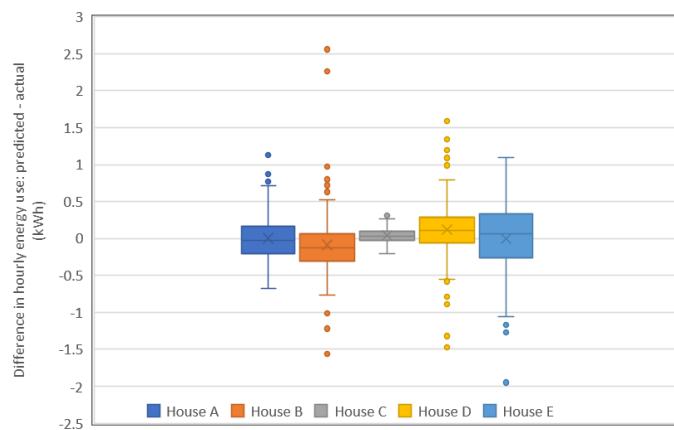


Figure 7. Boxplots of energy use difference (prediction error) between SARIMAX model and actual (measured) energy use for houses A-E

Table 2. Percentage of energy savings with a two degree increase in temperature setpoint

		Day1	Day2	Day3	Day4	Day5	Day6	Day7	Average
House A	Daily T setpoint (deg F)	73	73	73	74	73	74	73	73
	Energy saved by 2 deg. increase of T setpoint (%)	8.8	8.6	8.5	8.7	7.5	7.1	6.5	7.9
House B	Daily T setpoint (deg F)	78	77	78	78	77	77	77	77
	Energy saved by 2 deg. increase of T setpoint (%)	7.2	6.4	5.8	6.5	5.7	5.9	6.5	6.3
House C	Daily T setpoint (deg F)	75	76	76	75	76	76	76	76
	Energy saved by 2 deg. increase of T setpoint (%)	4.6	3.5	3.7	4.2	4.1	4.3	4.2	4.08
House D	Daily T setpoint (deg F)	73	74	73	73	74	74	74	74
	Energy saved by 2 deg. increase of T setpoint (%)	9.5	8.8	11.5	8.7	8	8.5	9.1	9.1
House E	Daily T setpoint (deg F)	74	73	73	74	74	74	74	74
	Energy saved by 2 deg. increase of T setpoint (%)	10.5	9.8	16.4	11.5	10	9.8	9.5	11.07

Note: Houses A,C,D,E setpoints based on data from Aug 1-7, House B based on Aug 31-Sep 6

Table 3. Estimated monthly energy use and electricity cost savings from changing setpoint

		Increase T setpoint by 1 deg. F			Increase T setpoint by 2 deg. F		
		On peak hours only (3-6 PM)	On and mid-peak hours (7 AM to 10 PM)	Full day (24 hours)	On peak hours only (3-6 PM)	On and mid-peak hours (7 AM to 10 PM)	Full day (24 hours)
Monthly savings							
House A	Energy use (kWh)	3.2	15.9	31.8	6.3	31.8	76.2
	Electricity cost (%)	0.6	1.8	3.1	1.2	3.5	6.2
House B	Energy use (kWh)	3.6	18.0	42.7	7.2	36.0	85.4
	Electricity cost (%)	0.5	1.8	3.0	1.2	3.5	6.2
House C	Energy use (kWh)	0.6	3.3	7.0	1.3	6.6	14.1
	Electricity cost (%)	0.4	1.2	2.1	0.8	2.4	4.0
House D	Energy use (kWh)	3.8	18.9	40.3	7.5	37.7	80.6
	Electricity cost (%)	0.7	2.1	3.4	1.4	4.2	6.9
House E	Energy use (kWh)	4.5	22.5	48.3	9.0	45.1	96.2
	Electricity cost (%)	1.0	2.8	4.6	1.9	5.7	9.2

Note: Monthly estimate for full day (24 hours) scenario is based on all days in the month including weekdays and weekends. Other scenarios are based on weekdays only.

Discussion

In this study, statistical time-series modeling techniques were applied to predict the air conditioning energy consumption of households at various set points. They are seen as relatively less complex than theoretical models, which generally involve a large number of calculations and computations modeling various processes and variables. On the other hand, statistical techniques

require comparatively fewer inputs/variables, making them more suitable for practical applications, particularly in the field, where input data are limited. In this research, it was shown that to predict AC energy use, a time-series model such as SARIMAX can be trained using only inputs such as past energy consumption and weather variables that can be easily obtained concurrently from local weather stations. The results have shown that these models can provide accurate predictions as long as the model composition and parameters are accurately determined. On the other hand, in theoretical models, high complexity and lack of precise input information can lead to errors in estimation.

Autoregressive moving average models use time as a factor in forecasting analysis. For the houses considered in this research, space cooling operation and energy consumption was most often driven by the heat transfer between the surroundings and the internal conditioned space. Since the surrounding conditions are highly variable and typically vary based on the time of the day, time was considered as an important factor in determining future AC energy consumption. In this research, the time varying patterns observed in AC energy use data were captured by SARIMA and SARIMAX models. With the help of the SARIMAX model, ambient temperature, a parameter that largely defines surrounding conditions, was used to further enhance the accuracy of the forecasting models.

Borowski and Zwolinska (2020) used Artificial Neural Networks (ANNs) to build a cooling energy use prediction model trained with weather variables and other influencing factors as predictors. Besides using ambient temperature as a strong predictor of AC energy-use, the variables, hour and day of the week representing the time factor were also included in the prediction modeling. The ANN models provided average CVRMSE and WMAPE values of 0.29 and 0.21 respectively, which was also obtained by the current study's SARIMAX statistical model. Although neural networks are good at handling linear as well as non-linear relationships in data, they are known for being data-intensive, requiring a lot of historical data to train on. Insufficient historical time-series data can result in a poorly fitted model. On the other hand, fitting a comprehensive model as neural networks with a substantial amount of data requires high data processing capabilities, making them computationally intensive too. Thus, statistical time series models have a high potential in this area since they give an accuracy near to that of neural networks by using simple strategies and limited past data.

Temperature difference between indoors and outdoors, seen as a major cause of heat gain, was given as the exogenous variable to the SARIMAX model to forecast AC energy use for increased setpoints. It gave the same accuracy as the model that used only outdoor temperature as the predictor variable. This is because variance observed in indoor temperature was considerably lower than the variance in ambient temperature making the temperature difference almost identical to the ambient temperature. To use this model for forecasting the energy use at increased setpoints, corresponding temperature differences were provided to make the predictions.

Many studies have demonstrated the positive impacts of adjusting temperature setpoint through various approaches. Aarish et al. (2015) studied the application of programmable and smart thermostatic controls. The report shows that both approaches could reduce cooling electric consumption by 13-14%. The difference between the average indoor temperatures after new thermostats were installed and initial setpoints before the installation, was observed to be in the range of 2-3 degree F, which indicated potential savings in the range 4-7% for each degree adjustment. Another similar study by Parker et al. (2016) evaluated the use of smart thermostats in 22 homes and reported average cooling energy savings of 9.6% and median savings of 6.3%

per year, but with a very high degree of variation. Although the study did not discuss the temperature settings in the homes, the finding was relatively consistent with the Aarish et al.'s study. The results of the current study, based on real-time data inputs and statistical modeling, yield slightly lower energy savings potential, i.e. 2.0-5.5% per degree increase, assuming a 24-hour adjustment to the temperature setpoint. The results from earlier studies accounted for within a day and day-to-day variations in the thermal controls by collecting data from direct interventions in the homes. Such a study approach is needed to understand the field application and realize the potential of the current time-series predictive tool. It is important to learn how occupants' may use the information feedback to adjust their setpoint and how the TOU considerations are taken into account in their decision making.

To study the influence of changing setpoint on energy consumption, Moon and Han (2017) simulated a typical U.S. single family home and developed theoretical models based on various parameters associated with heat gain and climate conditions. They tested the model in two climate zones, representing cold and hot-humid conditions. Their result from the hot-humid climate analysis shows that one degree F increase in temperature setpoint resulted in an average of 120 kWh or 5.7% energy savings, which is in good agreement with the upper limit of the percentage energy savings in the current results. Blonsky, Maguire, and Vincent (2020) conducted a study to quantify the energy cost savings from varying setpoints by 1, 3, 5, 7 degree C. They simulated one year temperature variations in five cities in different climate zones. They reported energy cost savings in the range of 2.1-4.0% per degree F increase in temperature setpoints. This result agrees very well with the findings of the current study, which shows savings in the range of 3.0-4.7% cooling energy cost savings for single family homes.

Comparing different TOU scenarios, i.e. on-peak, on-peak+mid-peak, vs. all time (24-hour) adjustments of temperature setpoints, it is apparent that energy cost savings are highest for the latter scenario. However, this does not mean it is the most optimal solution. If they opt to adjust setpoint only during the peak hours, they would only save between 0.5-1.0% cost for one degree F increase in setpoint but they will be under this condition for only about 8% of their time. If they extended the adjustment to include mid-peak hours, 1.8-2.8% of energy cost can be saved but they would be increasing their setpoint for about 42% of their time. The decision should be made by the occupants in terms of compromising their comfort, the amount of time they would be willing to live in slightly warmer temperatures, or the costs they are willing to pay for cooler temperatures. Expanding this research to field applications by including real-time energy rates in the prediction model along with predicted energy costs at various set points should be the next target goal to evaluate the feasibility of the technique and determine its influence on occupants' decision making and behavior.

Conclusion

This study shows that advanced techniques such as SARIMA/SARIMAX time-series modeling has significant potential for use in forecasting energy use when historical data is available to train the energy use model. In households where the occupants' generated heat remains relatively constant, outdoor air conditions and indoor temperatures can be considered the main factors contributing to the variance in the energy use data. Given that these factors are time-dependent, the time-series model could accurately predict the outputs at the next time stamps/scales as long as sufficient data is used. The current technique uses historical data from past 30 days at hourly intervals to predict the next hours up to one full day. This data resolution

appears sufficient for the current model needs, as observed from the accuracy metrics, i.e. CVRMSE and WMAPE.

The results of this study are in good agreement with past studies. A one degree increase in setpoint corresponds with about 2.0-5.5% energy saving for a 24-hour adjusted setpoint. In a single family home setting, this saving could lead to cost savings in the range of 3.0-4.7%. If occupants choose to only adjust their setpoint during peak hours (8% of time), their savings would be between 0.5-1.0% per one degree F increase. If they extend the adjustment to include mid-peak hours (42% of time), 1.8-2.8% of energy cost can be saved.

For future research, it would be important to expand the study to include field deployment of the calculation algorithms in home's thermostatic control and factored in real-time energy rates in the prediction model along with predicted energy costs at various set points and time.

References

U.S. EIA (Energy Information Administration). 2015. "Energy Used in Homes." www.eia.gov/energyexplained/use-of-energy/homes.php

U.S. DOE (Department of Energy). 2018. "Programmable Thermostats." www.energy.gov/energysaver/programmable-thermostats

McQuiston, F., J. Parker, and J. Spitler. 2004. *Heating, Ventilating and Air Conditioning: Analysis and Design*. Hoboken, NJ: Wiley.

ISO (International Organization for Standardization). 2018. *Energy performance of buildings — Energy needs for heating and cooling, internal temperatures and sensible and latent heat loads — Part 1: Calculation procedures*. ISO Standard No. 52016-1:2017.

Wang, N., J. Zhang, and X. Xia. 2013. "Energy consumption of air conditioners at different temperature set points." *Energy and Buildings* 65: 412-418.

Liao, J., M. Chang, and L. Chang. 2020. "Prediction of Air-Conditioning and Energy Consumption in R&D building using multiple machine learning Techniques." *Energies MDPI* 13:1847-1869.

Borowski, M., and K. Zwolinska. 2020. "Prediction of Cooling Energy Consumption in Hotel Building using Machine Learning Techniques." *Energies MDPI* 13:6226-6245.

Oh, C., and B. Morzuch. 2005. "Evaluating Time-series Models to Forecast the Demand for Tourism in Singapore." *Journal of Travel Research* 43: 404-413.

Chikobvu, D., & Sigauke, C. 2013. "Modeling influence of temperature on daily peak electricity demand in South Africa." *Journal of Energy in Southern Africa* 24(4):63-70.

Hyndman, R.J., and G. Athanasopoulos. 2018. *Forecasting: principles and practice*. Melbourne, Australia: OTexts.

Box, G., G. Jenkins, and G. Reinsel. 2008. *Time Series Analysis: Forecasting and Control*. Hoboken, NJ: Wiley.

Aarish, C., M. Pesussi, A. Rietz, D. Korn. 2015. *Evaluation of the 2013-2014 Programmable and Smart Thermostat Program*. Evansville, IN: Cadmus Group, Inc.

Parker, D., K. Sutherland, D. Chasar. 2016. "Evaluation the Space Heating and Cooling Energy Savings of Smart Thermostats in a Hot-Humid Climate using Long-term Data". *In proceedings of the 2016 ACEEE Summer Study on Energy Efficiency in Industry* 8:1-15. Pacific Grove, CA: ACEEE

Moon, J., S. Han. 2010. "Thermostat strategies impact on energy consumption in residential buildings." *Energy and Buildings* 43 (2011): 338-346

Blonsky, M., J. Maguire, Vincent T. 2020. "The Cost of Comfort in Your Home: Energy and Cost Savings from Setpoint Adjustment of Space Heating and Cooling." *In Proceedings of the 2020 Mines Graduate Research and Discovery Symposium*. Golden, CO.