Effects of ADHCR on Normalized Electricity Consumption

Lingqi Huang(hlingqi@umich.edu)

December 2023

Abstract

This paper investigates the Discrepancy in Heating and Cooling Ratios (DHCR) across various census regions and climate zones, utilizing data from the 2018 Commercial Buildings Energy Consumption Survey (CBECS). Our analysis delves into the impact of DHCR variations on overall electricity consumption and seeks to identify an appropriate model through the application of different statistical methods, leveraging key variables of interest.

Github: https://github.com/KingOliver666/HW-.git

1 Introduction

In contemporary business operations, the utilization of electricity and its associated costs represent a pivotal factor influencing a company's annual expenditure. It is imperative to discern the impact of various potential factors on the total electricity consumption. In our analysis of the 2018 Commercial Buildings Energy Consumption Survey (CBECS) data, we gathered information from 6,357 diverse buildings with distinct purposes, examining a comprehensive set of 1,250 variables. Our primary emphasis was on investigating the influence of a specifically constructed variable, known as DHCR, on electricity consumption. DHCR is defined as follows:

$$DHCR = \frac{\text{Electricity Consumption on Heating} - \text{ Electricity Consumption on Cooling}}{\sqrt{\max\{\text{ Electricity Consumption on Heating, Electricity Consumption on Cooling}\}}}$$

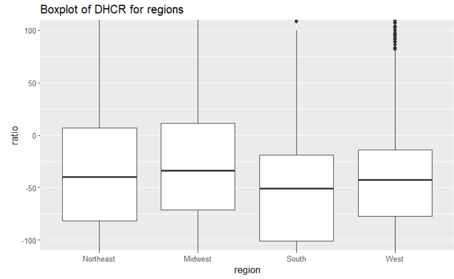
The rationale behind introducing this constructed variable lies in its ability to discern the predominant usage of electricity within a building. A positive sign for DHCR indicates a higher annual electricity consumption for heating, while a negative sign suggests a greater consumption for cooling. In cases where the disparity between heating and cooling consumption approaches zero, signifying nearly equal usage, the denominator of DHCR plays a crucial role. It accurately incorporates the magnitude of electricity consumption for both heating and cooling, preventing instances where a building may exhibit a small total electricity consumption for heating and cooling but possess a disproportionately large DHCR.

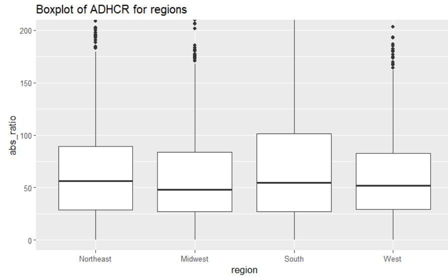
Additionally, we introduce a new variable, the absolute value of DHCR, denoted as ADHCR ($|DHCR|$). This variable is inherently positive and is deemed a plausible factor influencing the total annual electricity cost. High consumption in either cooling or heating, as indicated by ADHCR, can impact the overall annual electricity cost. While DHCR provides insight into both the magnitude and direction (heating or cooling), ADHCR specifically highlights the absolute magnitude, serving as a potential predictor for the annual electricity cost.

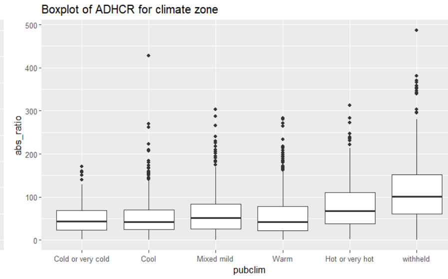
Prior to developing a prediction model, we conducted descriptive statistics and ANOVA on DHCR and ADHCR across various census and climate regions.

2 Data

To initiate our analysis of DHCR, we excluded 606 observations with zero costs for both heating and cooling to ensure the meaningfulness of DHCR. Subsequently, for numerical stability in our data and predictions, we normalized electricity consumption. Notably, across the four census regions—Northeast, Midwest, South, and West—we observed mean DHCR values of -37.48, -28.24, -61.9, and -40.8, respectively. The corresponding mean ADHCR values were 65.4, 59.7, 71.3, and 60.4, respectively. It is noteworthy that the sign of DHCR for all four regions is negative, indicating that, on average, these regions incur higher costs for cooling than for heating. Additionally, the magnitude of DHCR for the South region is notably larger.We also did a ANOVA of the mean of ADHCR for these 4 regions with $\alpha = 0.05$ and we reject our null hypothesis that four regions have the same average ADHCR.(The ANOVA can be used because it satisfys conditions of ANOVA, which are Independence, Normality, and Homoscedasticity)







We observed noteworthy findings in the analysis of DHCR for five distinct climate zones—Cold or very cold, Cold, Mixed mild, Warm, Hot or Very hot—along with one category labeled as "withheld" due to confidential reasons. Notably, the mean DHCR for the 'withheld' category is extremely high at -87.5, as is its ADHCR at 112.3. Additionally, the normalized mean electricity consumption for this "withheld" is 1.06, surpassing that of any other climate zone(since all others have negative average z-score). We then did a ANOVA of mean of ADHCR for these 6 climate zone with $\alpha = 0.05$ and we reject our null hypothesis that 6 climate zone have the same average ADHCR.

Recognizing a potential link between a higher Assessed Design Heating and Cooling Requirement (ADHCR) and increased normalized electricity consumption, we are motivated to explore the relationship between ADHCR and normalized electricity costs. This investigation leads us to the process of model selection.

3 Model Selection

We initially conducted a simple linear regression, examining the relationship between normalized electricity consumption and the Assessed Design Heating and Cooling Requirement (ADHCR). The resulting linear model is represented as

$$\widehat{Norm}_{Elec} = -0.718 + 0.011 * \widehat{ADHCR}$$

In this simple linear regression model, the t-value for ADHCR is approximately 51.07, and the associated P-value is close to zero. Consequently, we can reasonably conclude that ADHCR is statistically significant, indicating a positive relationship between ADHCR and normalized electricity consumption.

Next, we extended our analysis to include several variables of interest, namely sqft/1000 (square footage), renelc (Electrical upgrade), wkhrs (Total hours open per week), pctermn (Number of desktop computers), renlgt (Lighting upgrade), monuse (Month in use), and laptpn (Number of laptops). The goal was to assess the significance of ADHCR in predicting normalized electricity consumption in the presence of these additional variables.

We adopted three methods to selecting linear model, which are backward elimination, R^2-adjusted based criterion, and stepwise regression on both direction, and the three models are:

Backward Elimination:

$$\widehat{Norm}_{Elec} = -0.86 + 0.0039\widehat{ADHCR} + 0.0017I\left(\frac{\widehat{sqft}}{1000}\right) - 0.122\widehat{renelNo} \tag{1}$$

$$+ 0.0032\widehat{wkhrs} + 0.0006\widehat{pctermn} + 0.072\widehat{renlgtNo} \tag{2}$$

Stepwise Regression:

$$\widehat{Norm}_{Elec} = -0.91 + 0.0039\widehat{ADHCR} + 0.0017I\left(\frac{\widehat{sqft}}{1000}\right) - 0.122\widehat{renelNo} \tag{3}$$

$$+ 0.0036\widehat{wkhrs} + 0.0006\widehat{pctermn} + 0.072\widehat{renlgtNo} \tag{4}$$

Adjusted R^2:

$$\widehat{Norm}_{Elec} = -1.1 + 0.0039\widehat{ADHCR} + 0.0016I\left(\frac{\widehat{sqft}}{1000}\right) - 0.124\widehat{renelNo} \tag{5}$$

$$+ 0.0032\widehat{wkhrs} + 0.0006\widehat{pctermn} + 0.073\widehat{renlgtNo} + 0.018\widehat{monuse} \tag{6}$$

In all three selected models, the adjusted R^2 values are approximately 0.6. Notably, our targeted variable, ADHCR, consistently exhibits a positive and statistically significant relationship with normalized electricity consumption across these models. This consistency in the positive relationship provides evidence suggesting no obvious collinearity between ADHCR and the other variables.

4 Conclusion

Upon comparing various census regions and climate zones, a noteworthy observation emerged in the southern area and the climate zone denoted as 'withheld' due to confidentiality concerns. In these regions, there is a conspicuous trend of exceptionally high DHCR and ADHCR , accompanied by higher electricity consumption. Subsequent analysis revealed a positive relationship between ADHCR and normalized electricity consumption. Notably, three distinct modeling approaches—backward elimination, stepwise regression, and the criterion of Adjusted-R^2 —all indicate that ADHCR stands out as a strong predictor for normalized electricity consumption.

5 Discussion

Although we identified a positive relationship between ADHCR and normalized electricity consumption, it's important to note that we excluded observations with zero electricity consumption on both heating and cooling. By doing so, we acknowledge that we may have omitted valuable information from these eliminated observations, and, consequently, ADHCR may not be a suitable predictor of normalized electricity consumption for this specific subset.

In addition, our models were constructed using only a small subset of predictors, potentially introducing bias. We also did not explore the independence between ADHCR and other variables not considered in our analysis. Furthermore, for numerical stability, we opted for normalized electricity consumption instead of the raw values. While this choice enhances stability, it complicates the interpretation of the effect of variables on electricity consumption.

Lastly, we omitted numerous observations during model selection due to the presence of many missing values in our data and we manually removed them. This exclusion may introduce imprecision in estimating the effects of our predictors and result in larger errors when predicting normalized electricity consumption.