ENERGY SYSTEM IMPACTS OF HEATING SYSTEM ELECTRIFICATION IN NEW YORK STATE

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Master of Science

by

Rashika Mittal

May 2022

Abstract

New York State (NYS) has taken a leadership role in deep decarbonization by passing the Climate Leadership and Community Protection Act (CLCPA), which commits New York to reach net-zero greenhouse gas emissions and requires 40 percent emissions reductions from 1990 levels by 2030 and 85 percent emissions reductions by 2050. As is in other cold-climate regions, heating in the residential and commercial sectors is the single largest end-use category in New York State. And over 80 percent of the heating-related energy consumption comes from the onsite combustion of fossil fuels dominated by natural gas and fuel oil.

Electricity-driven heating technologies such as air-source and ground-source heat pumps provide a viable pathway to decarbonize the heating sector as they can integrate with renewable electricity generation such as wind and solar. However, the main challenges to electricity-driven heating arise from its impact on the power system, notably from an increase in winter peak electricity demand in the distribution networks.

By using optimization models and statistical techniques, this thesis estimates that the total heating electricity demand arising from converting fossil fuel heating systems to electric heating systems in the NYS residential and commercial sectors is about 48 percent more than the current requirement. The results show a clear shift from summer peak to winter peak due to heating system electrification in NYS. Additionally, we show some heating electrification scenarios using different heat pump models. Scenarios using various heat pump adoption rates based on studies that support heating system electrification in the state and surrounding areas were also developed. The findings of the study are of importance to state and local policymakers and energy market participants. This quantification will help the policymakers in taking more defined steps towards developing resilient energy systems, building energy-efficient buildings and heating systems, and empowering greater adoption of sustainable energy systems.

Keywords: energy demand, heating electrification, heat-pump, county, residential and commercial sector, optimization, statistics

Biographical Sketch

Rashika was born and brought up in the industrial city of Kanpur, Uttar Pradesh in India. In Kanpur, she experienced days of power cuts and on the bus ride to school, she could smell the stench in the air because of improper waste disposal by the tanneries in the area. These experiences made her interested in the energy industry, and she understood that managing power and pollution is the only way of improving the lifestyle of the people around her. She did her schooling at Seth Anandram Jaipuria School where she felt inspired when she saw her teachers planting trees on the school land before every monsoon.

Rashika is an ardent supporter of the fight against climate change. She believes that every small step counts and tries to incorporate energy-conserving habits in her everyday life. She is amazed by the various ways in which people are using technology to reduce the use of fossil fuels. To get hands-on experience with such technologies, she joined Manipal Institute of Technology to do her bachelor's in chemical engineering. During her bachelor's, she participated in a variety of research projects ranging from photocatalysis for wastewater treatment to synthesizing lithium-sulfur batteries for hybrid electric vehicles. Rashika graduated in 2018 with a university gold medal for graduating summa-cum-laude in Chemical Engineering.

After a year-long stint at Deloitte in Bangalore, India, Rashika came to Cornell in August of 2019 to pursue her master's in Chemical Engineering. Since then, she has been working with Professor Max Zhang and Professor Jefferson Tester to determine the electricity demand upon the electrification of heating systems in the State of New York to aid the State in moving further with its 2030 Energy Goals.

To my parents and brother

Acknowledgments

No one can whistle a symphony; it takes a whole orchestra to play it. The symphony of this thesis was played by several musicians—my advisors, mentors, the very knowledgeable people at and outside Cornell, friends, and parents.

First and foremost, I would like to thank my advisors and committee members—Professor Max Zhang and Professor Jeff Tester—without their constant support, guidance, and encouragement, this thesis would not have materialized. Professor Tester has been very encouraging towards my work, always providing suggestions that push the results to be as close to practical as possible. He was very supportive in getting me in touch with the right people when I needed expertise in a domain to which I was new.

Professor Zhang was the prime mover of this thesis. He not only supported but actively encouraged exploring novel ideas and methods employed in this work. Most importantly, he helped me walk the way when things seemed out-of-control and unachievable. The most important thing that I got to learn from him was his unwavering focus on becoming excellent and not losing sight of the real-world application of the research. He also understood my long-term goals and supported my endeavors outside academics like my summer internship at Environmental Defense Fund (EDF). Thank you so much, Professor Zhang and Professor Tester.

I would like to express my deepest gratitude to my mentors—Elizabeth Stein and Pamela McDougal who were great guides and supporters during my internship at EDF and helped me understand many aspects of the energy sector in the real world, David L. Kay who was extremely helpful in data analysis and critical decision making at a very important time in my thesis work, and Professor Snow whose communication classes were godsent.

I would also like to thank my group members— Zach for his expert guidance in the department of building heating systems and coding in Python; Jeff, who has guided me through the ups and downs in my thesis journey; Ali and Bo for their valuable inputs and willingness to help whenever possible; friends—Saumya, for helping me with good coding practices and being the fresh eyes to look at the code bugs be it weekday or weekend; Niharika, for helping me improve the editorial quality of the thesis (and being my best friend); Shikhar, who tirelessly helped me during the most uncertain times in graduate school; my seniors and batchmates for keeping me sane in times of crisis—Venktesh,

Shrikant, Akshay, Anmol, Sruti, Ritika, Amit, Karthik, Sahil, Murti, and Sushrut. My heartfelt thanks to all of you!

I would also like to thank all other people who have been an integral part of my thesis—Koenraad who helped with critical questions regarding heat pumps; Patrick who helped with information on heat pump models; Polly, who helped in meeting organization; Wendy, for helping me with all the IT and computer issues; and all the staff from CBE and NYS government.

Finally, I would like to express my deepest gratitude to my parents and my brother. It is only due to their empowering wishes and unwavering support, I am where I am today.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

HP – Heat Pump(s)

COP – Coefficient of Performance

CES – Clean Energy Standards

EIA – Energy Information Administration

NYS – New York State

GSHP – Ground Source Heat Pump

HSPF – Heating Seasonal Performance Factor

MSHP – Multi-split Heat Pump

NRDC – Natural Resources Defense Council

NREL – National Renewable Energy Laboratory

ORPTS – Office of Real Property Tax Services

PLUTO – Primary Land Use Tax Lot Output

RECS – Residential Energy Consumption Survey

VEIC – Vermont Energy Investment Corporation

ccASHP – cold climate Air Source Heat Pump

CLCPA – Climate Leadership and Community Protection Act

NYISO – New York Independent System Operator

NYSERDA – New York State Energy Research and Development Authority

1. Introduction

Over 60% of households in the United States (and 80% in the Northeastern U.S.) utilize on-site fossil-fuel-based space heating, accounting for 10% of the country's greenhouse gas emissions [1]. As the world moves toward a carbon-free energy grid, a major shift to electric space-heating systems is needed to mitigate carbon emissions to arrest climate change [2]. This can be expected to increase the electricity demand. In the areas where electricity demand currently peaks in winter, space heating electrification can exacerbate an already strained grid infrastructure [3], [4]. With about 40% of the U.S. population living in cold climate states [5], predicting electric heating demand is critical to the planning, development, and management of more efficient and robust energy infrastructure.

Only through a granular estimate of the increase in the demand for electricity resulting from the electrification of current fossil-fuel-based heating systems, can we estimate both the extent of increase in electricity generation capacity needed, as well as when and where this need will arise. This quantification will also help the policymakers in taking more defined steps towards maximizing the adoption of clean resilient energy systems. Thus, analyzing the impact of electrification on the transmission grid is necessary for long-term planning for power system upgrades [6].

Given the importance of heating electrification towards reducing greenhouse gas emissions and the sheer impact it can have on the electricity infrastructure, research interest in this area is growing exponentially. There are studies that calculate the increase in the demand for heating electricity that will result from the electrification of the heating systems. White et al. [7] estimate heating electrification demand for the residential sector in Texas after it experienced a major blackout. The study reports a 36% increase in the peak electricity demand as compared to the current residential electricity consumption with the existing mix of heating systems. Waite et al. [8] model the electrification demand for residential and commercial sectors for the entire United States, at the census tract resolution, using country-wide data. The study reports that space heating electrification could require a 70% increase in the nation’s electricity system capacity. There has been research on determining the space heating demand upon building envelope refurbishment [9] or climate change [10] in the residential sector. Some studies analyze the building energy consumption to understand the energy consumption trend in residential buildings [11]. As the world moves toward electrification, studies that calculate the heating electrification demand for both the residential and commercial building sectors become increasingly important.

Most research work on estimating building energy consumption use a constant setpoint temperature [9], an ambient reference temperature [8], or simulation-based data [12] instead of real-world data to replicate

the temperature settings preferred in general households. Furthermore, it is very common that the studies that predict or calculate heating demand or building energy demand use simulation data from software packages like EnergyPlus, eQuest [13], etc. to validate their methodology or have no way to verify their work [8]. The analyses that use simulation tools usually require a lot of details related to the building for which the calculations are performed. This can be a tedious process and hard to scale for larger geographies. The ones that do not verify their work, run the danger of deriving conclusions based on a potentially flawed methodology and are not indicative of the actual scenario at hand.

Our study showcases a method to calculate the heating electrification demand for residential and commercial sectors taking New York State (NYS) as an example. It holds great importance for NYS because it is a cold climate state and will experience major electrification in the near future. As per the Strategic Outlook for 2021-2024 by the New York State Energy Research and Development Authority (NYSERDA) [14], NYS aims to make the building sector carbon neutral with the help of electrification. These steps are inspired by the goals set by the Climate Leadership and Community Protection Act (CLCPA) towards reducing 22 million tons of carbon through energy efficiency and electrification [5].

Our work is different from the studies done before because of two main reasons: 1) We use two datasets—the Ecobee smart thermostat dataset and the NYS tax parcel dataset—which haven't been used before to perform such a calculation. 2) Another noteworthy aspect is that this model is validated against actual operational system data.

To elaborate on the points above, the hourly setpoint temperature data for residential buildings was taken from the newly available Ecobee Donate your Data smart thermostat dataset [16]. This dataset was analyzed to find that over 90% of smart thermostat users employ unique setpoint schedules, and the cumulative effect of these programmed setpoints can have a considerable impact on energy demand profiles. Thus, instead of using a constant setpoint temperature for the various homes, through this dataset, we bring this study closer to a real-life scenario by taking into account the actual variable temperature settings. Furthermore, we use the NYS Tax Parcel dataset to find the values of building floor area, building heating fuel type, etc. Using a state-owned dataset makes it more relevant for calculating values for NYS specifically.

Additionally, to ensure that the methodology that the study follows is reliable, we compare the demand values generated by the model with publicly available 2019 New York Independent System Operator (NYISO) data [17]. NYISO data is metered load data recorded by NYISO at a five-minute resolution and covers 11 New York Control Area load zones. It is measured in MWh, quality checked for consistency and completeness, and updated every day. We compare the NYISO data values and trend with the model-

calculated demand values, to find peaks in our model at the same points as the NYISO data values. We see that our model closely follows the trend shown by the NYISO data with only a 7% average difference between the hourly values. Since method validation is an important part of comprehensive quality management to ensure the scientific integrity of analytical data, using such a comprehensive dataset to verify our methodology reinforces the dependability of this study.

For NYS specifically, our study computes approximately a 48% increase in heating electricity demand for residential and commercial sectors. We have assumed the installation of a multi-split heat pump (MSHP) to electrify fossil-fuel-based heating systems. Our result is an estimated value because it will keep reducing with advancement in heat pump (HP) technology resulting in more efficient heating systems that will require less energy to provide the same amount of thermal comfort. To elaborate on this, our work develops various electrification scenarios using highly efficient HP models. We show that technologically advanced HPs that are currently available can reduce electrification demand by as much as 20% as compared to the standard HPs. In addition, we also develop scenarios of different HP penetration values based on a number of studies on heating electrification conducted in NYS. We use these values to help us visualize the potential electricity demand profiles under various rates of HP adoption.

Thus, this study facilitates in providing insights into the grid implications of heating electrification initiatives and regulations. It is valuable to grid planners and policymakers, particularly those in regions largely driven by summer cooling demands to prepare for the shift in seasonal peaks. Our methodology can be used to find demand at a county level; thus, it can assist in the development of specialized electrification plans with numerous localized benefits. The approach used in this study might be applied to other geographies to investigate the effects of space heating electrification with various other heat pump types like ground source heat pumps (GSHPs) or other more technologically advanced heat pumps which could have the potential to bring the demand further down from the current levels.

The results of this work can be used by future studies to determine the impact of such an increase in electricity demand. They can also be used to determine the amount of energy that can be provided by renewable resources, along with helping stakeholders evaluate if the current renewable infrastructure is enough to support the demand. If the current installations are not enough, and new installations are needed, studies building on our research can also determine the most effective composition of renewable energy sources—such as wind, hydropower, solar, etc.—that can be used to power the grid in New York State.

The remainder of the thesis is organized into three more sections: section 2 explains the datasets that were used and outlines the methodology followed by the project, section 3 discusses the results, section 4 concludes the thesis, and section 5 talks about future work.

2. Methods and Materials

2.1. Data Collection and Description

2.1.1. NYS Tax Parcel Dataset

This study uses the NYS Tax Parcel centroid point data which is available for all 62 counties. It has several attributes including building type, building area, heating fuel type, etc., which are populated by the GIS Program Office using Assessment Roll tabular data from the NYS Department of Tax and Finance's Office of Real Property Tax Services (ORPTS). A property assessment roll is a public record collected for each property on an assessing unit (municipality or county) to calculate the taxable value of the property. It contains information on the market value of the property, the date when the valuation was done, and personal homeowner information like phone number and address.

OBJECTID	COUNTY_NAME	PROP_CLASS	BLDG_STYLE	BLDG_STYLE_DESC	HEAT_TYPE	HEAT_TYPE_DESC	FUEL_TYPE	FUEL_TYPE_DESC	SQFT_LIVING	GFA	USED_AS_CODE	USED_AS_DESC
1	Jefferson	322	0	0	0	0	0	0	0	0	0	0
2	Jefferson	322	0	0	0	0	0	0	0	0	0	0
3	Jefferson	321	0	0	0	0	0	0	0	0	0	0
4	Jefferson	322	0	0	0	0	0	0	0	0	0	0
5	Jefferson	105	0	0	0	0	0	0	0	0	0	0
6	Jefferson	112	8	Old style	2	Hot air	1	Oil	1937	0	0	0
7	Jefferson	105	0	0	0	0	0	0	0	0	0	0
8	Jefferson	210	6	Contemporary	3	Hot wtr/stm	1	Oil	1792	0	0	0
9	Jefferson	120	0	0	0	0	0	0	0	0	0	0
10	Jefferson	210	17	Manufactured Housing	2	Hot air	1	Unknown	1404	0	0	0

Figure 1 This snapshot of the 2019 NYS Tax Parcel data for Jefferson County shows the columns like 'PROP_CLASS', 'SQFT_LIVING', 'GFA', etc. that we used in our study.

In this study, the 2019 NYS Tax Parcel dataset was used to obtain the residential and commercial building floor area values; $p_{elec,Heat,b}$, which is the fraction of residential or commercial building class floor area that use electric heating (described in Section 2.2.1.1.) ; and $p_{FF,Heat,res}$, which is the fraction of residential building class floor area that use fossil-fuel-based heating (described in Section 2.2.1.2.). We used the 'SQFT_LIVING' column to get the residential area for the buildings with residential property codes ('PROP_CLASS') between 200 and 300, as defined in the data dictionary. Similarly, we used the 'GFA' column to get the commercial area for the buildings with commercial property codes ('PROP_CLASS') between 400 and 500, as defined in the data dictionary.

Even though the dataset contains information for all NYS counties, the Nassau County data is not completely populated and contains no values for 'SQFT_LVING' and 'GFA'. Since these values are needed for our work, we exclude Nassau County data completely in our calculations.

Similarly, the data for New York City (NYC) counties in this dataset was also incomplete because NYC uses a different system for data collection. Therefore, to fill this gap, we used the 2021 PLUTO dataset (explained below) to get the residential and commercial building floor areas for this region.

2.1.2. PLUTO Dataset

To get the residential and commercial building floor areas for NYC counties, we used the 2021 Primary Land Use Tax Lot Output (PLUTO) dataset [18]. PLUTO contains land use and geographic data at the tax lot resolution. It combines data from various city departments like the Department of City Planning (DCP), Department of Finance (DOF), Department of Citywide Administrative Services (DCAS), and Landmarks Preservation Commission (LPC). These departments collect building data, e.g., types of use, location, etc., in real-time. As the city acquires new or disposes of old properties, the datasets are updated. PLUTO contains three main data categories: 1) Tax Lot Characteristics, 2) Building Characteristics, and 3) Geographic/Political/Administrative Districts.

2.1.3. EIA Data

The U.S. Energy Information Administration (EIA) [19] is the Department of Energy's statistics and analytical arm. EIA gathers, analyzes, and disseminates independent and unbiased energy data to support smart policymaking, efficient markets, and to create public awareness of the role of energy in the economy and environment.

The actual value of 2019 sector-wise monthly electricity and fossil fuel [23],[21] consumption in NYS was taken from EIA. EIA obtained the sector-wise energy consumption data by conducting surveys of both utility and non-utility companies that sell/deliver gas or electricity to customers to obtain retail sales and revenue data for all end-use sectors, viz, residential, commercial, industrial, and transportation.

In terms of units of measure, the sector-wise electricity retail sales data is available in million kWh or thousand MWh. Fossil fuel consumption data is available in million cubic feet or gallons which we converted to Wh for the sake of unit consistency.

2.1.4. Weather Data

We used the 2019 National Solar Radiation Database (NSRDB) [22] data to get the hourly temperature for all the counties. NSRDB contains meteorological data at an hourly and 30-minute resolution for the

whole of the United States at different geographical scales. Multi-channel measurements from geostationary satellites are used to model the present NSRDB.

NSRDB data is not separated by county. To remedy this, we used the latitudes and longitudes of the most populated areas in each county to estimate county-wide temperatures for the year 2019 at an hourly resolution.

2.1.5. Ecobee Thermostat Data

Smart-home devices with built-in communication technology have become increasingly common in recent years, allowing for wider implementation of smart heat pump control. Smart thermostats, for example, have gained significant acceptance in recent years and are expected to be in 40% of US households by 2021. Smart thermostats provide a variety of energy-saving features such as unique setpoint schedules, different comfort configurations like "sleep" and "away", and occupancy sensors that cut energy use while the house is unoccupied.

Hourly setpoint temperature data for residential buildings was taken from the newly available Ecobee Donate your Data smart thermostat dataset [16]. This dataset was analyzed to find that over 90% of smart thermostat users employ unique setpoint schedules, and the cumulative effect of these programmed setpoints can have a considerable impact on energy demand profiles [2]. The dataset contains temperature data of various homes located in different counties in NYS at a five-minute resolution. We averaged the temperature of all the NYS homes for the year 2019 and filtered it by the hour.

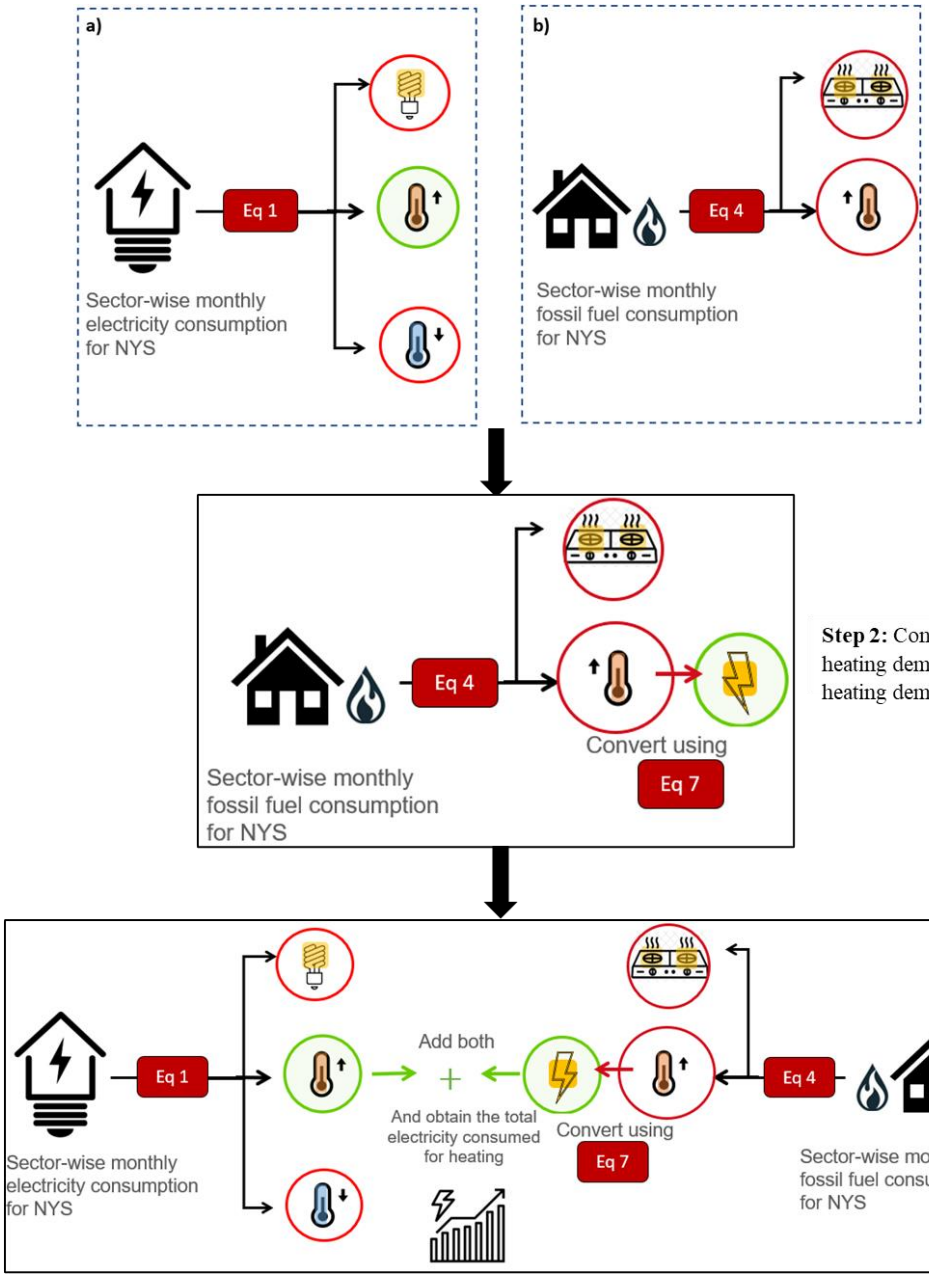
2.1.6. Data for Constants – $p_{AC,b,c}$, $p_{FF,Heat,com,c}$

The table below describes the data sources for some of the constants used in the heating demand calculations. These constants are: the fraction of floor area of each building class, b ($com\ or\ res$), in each county, c, that uses electric cooling, AC, is denote by $p_{AC,b,c}$; and the fraction of floor area of each building class, b ($com\ or\ res$), in each county, c, that uses fossil-fuels for heating, $Heat$, is denoted by $p_{FF,Heat,b,c}$.

Table 1 Data Sources for Calculating the Constants Used in the Model Equations

Constants	**Data Source**
The fraction of the floor area in **commercial** buildings that uses electric cooling ($\boldsymbol{p_{AC,com,c}}$)	NYSERDA Commercial Baseline Study [24] is a study conducted by NYSERDA in collaboration with the New York State Department of Public Service (DPS). The study assesses the existing commercial building stock and energy-consuming equipment (e.g., HVAC), including equipment count and characteristics. The data is based on telephone/web survey responses of 3,882 commercial customers and on-site audits at 826 businesses.
The fraction of the floor area in **commercial** buildings that uses fossil fuel for heating ($\boldsymbol{p_{FF,Heat,com,c}}$)	
The fraction of the floor area in **residential** buildings that use electricity for cooling in NYC ($\boldsymbol{p_{AC,res,c}}$)	The U.S. Census American Community Survey (ACS) is a tool that helps local governments, community leaders, and companies understand how their areas are changing. The source of the data is ACS 'SELECTED HOUSING CHARACTERISTICS' TableID: DP04 Product: 2019: ACS 1-Year Estimates Data Profiles [25].
The fraction of the floor area in **residential** buildings that use electricity for cooling in all other counties except NYC counties ($\boldsymbol{p_{AC,res,c}}$)	Residential Energy Consumption Survey (RECS) 2015 is a multi-year project that includes a Residential Survey phase, data collecting from household energy suppliers, and calculation of end-use consumption and expenditures [1].

2.2. Methodology



Step 1: Split total energy use into component parts:

a) Lighting, heating, and cooling for electricity

b) Lighting and heating for fossil fuels

Step 2: Convert fossil-fuel based heating demand into electric heating demand

Step 3: Add existing electric heating demand and converted electric heating demand to get total electric heating demand

Figure 2 illustrates the methodology followed in this study and shows that it is broken into three steps: 1) split the energy usage into its respective components, 2) convert fossil-fuel-based heating energy demand into heating electricity demand, 3) add up the existing heating electricity usage and converted heating demand to obtain the total heating electricity demand.

Figure 2 is a step-by-step explanation of how the model equations will be used to evaluate the heating electricity demand.

According to step 1, Equation 1 takes sector-wise monthly electricity consumption data pulled from the Energy Information Administration as input and splits it into three broad usage categories—lighting (and other non-temperature-dependent usages), heating, and cooling. Similarly, Equation 4 takes the sector-wise monthly fossil fuel consumption data as input and splits it into two broad usage categories—cooking (and other non-temperature-dependent usages) and heating.

In step 2, we convert the fossil-fuel-based heating usage into electric heating usage using Equation 7. While doing this conversion, we use the temperature-dependent coefficient of performance (COP) developed using Equation 9 to account for electric heating system efficiency. These equations are elaborated in great detail in the following section.

Finally, in step 3, we add up the existing electric heating demand and the converted (fossil fuel to electric) heating demand to get the approximated total electricity demand for when all the heating systems in NYS become electric.

To get a holistic view of the potential increase in the demand for electricity, we evaluate the electrification scenario developed by using HPs to replace the fossil fuel-based space heating systems in the residential and commercial building sectors. Additionally, we perform calculations at varying levels of HP penetration and find the peak hourly demand at the state level to understand the hours when peaker plants—the power plants that are only used when there is a significant demand for electricity— could potentially be required to satisfy the demand.

2.2.1. Model Equations

The model equations in this thesis were inspired by the methodology followed in a study at Columbia University by Waite et al. [8]. Our work makes improvements to that study by bringing about three major changes: 1) the use of NYS Tax Parcel data to get building-related information as compared to the National data used by Waite et al, 2) the introduction of internal setpoint temperature variability in

residential buildings using Ecobee data instead of using a single constant temperature for all the buildings for which calculations were made, and 3) by evaluating the baseline prediction using actual NYISO data, which is exclusive to our study.

2.2.1.1. Temperature-Dependent Electricity Usage Model

The estimated value of temperature-dependent electricity usage [8], i.e., for cooling and heating calculated for each building class, b, in each county, c, at each time step, t as shown in Equation 1, is found out by splitting the total monthly electricity usage of a county into three parts: 1) temperature-independent electricity usage (e_b^{const}), i.e., electricity used for lighting, refrigeration, etc.; 2) increasing-temperature-dependent electricity usage (e_b^{+}), i.e., electricity used for cooling; and 3) decreasing-temperature-dependent electricity usage (e_b^{-}), i.e., electricity used for heating. The building floor area for each building class (residential and commercial), b, and county, c, $A_{b,c}$, is determined using the NYS Tax Parcel dataset. $p_{AC,b,c}$ is the fraction of the floor area of each building class in each county that uses electric cooling (AC). $p_{elec,Heat,b,c}$ is the fraction of floor area of each building class in each county that uses electricity for heating ($Heat$).

Thermostat setpoint for residential buildings is chosen such that it enhances human comfort and the Ecobee smart-thermostat data [16] captures the variability in setpoint temperatures in different homes which greatly improves the accuracy of the calculations performed. The hourly reference temperatures are taken as statewide average Ecobee thermostat hourly setpoint temperatures. Ecobee thermostats are only employed in residential buildings and therefore these temperatures were not used for commercial buildings. The setpoint temperature for actual commercial buildings varies significantly depending on the type of building and many other factors, but a value of $T_{ref,com}$ = 68°F was taken as a constant setpoint temperature for all commercial buildings for enhanced energy savings [26]. $T_{c,t}$ is the hourly outdoor temperature for each county in NYS taken from NSRDB.

$$\hat{E}_{b,c,t} = A_{b,c}[e_b^{const} + p_{AC,b,c}e_b^{+}\left(T_{c,t} - T_{ref,b}\right)^{+} + p_{elec,Heat,b,c}e_b^{-}\left(T_{ref,b} - T_{c,t}\right)^{+}] \quad (1)$$

The '+' outside the parentheses indicates that the value of the term is zero when negative. The increasing and decreasing temperature-dependent terms are weighted by the fraction of building floor area using

electricity for cooling and heating, respectively. We assume that the primary end-use of increasing and decreasing temperature-dependent electricity usage is cooling and heating only.

The fraction of floor area in each building class, in each county that uses electricity for heating, $p_{elec,Heat,b,c}$, was determined in the same way for both residential and commercial building classes by dividing the building class area that uses electricity by the total building class floor area found using NYS Tax Parcel data. However, $p_{elec,Heat,b,c}$ was evaluated differently for the NYC counties. For $p_{elec,Heat,res,c}$ for NYC, we assumed that this fraction is equal to the number of households having heating equipment divided by the total number of occupied household units as per the American Community Survey data [25]. For $p_{elec,Heat,com,c}$ for NYC, we used the NYSERDA Commercial Baseline Study. The commercial baseline study was conducted by NYSERDA in collaboration with the New York State Department of Public Service (DPS). The study assesses the existing commercial building stock and energy-consuming equipment (e.g., HVAC), including equipment count and characteristics.

The county-wise fraction of building class floor area with air conditioning, $p_{AC,b,c}$, was determined differently for each of the commercial and residential building classes. For $p_{AC,com}$, we use the fraction of businesses having electric cooling in place as per the NYSERDA Commercial Baseline Study [24]. For the residential building class, $p_{AC,res,c}$ was found from the Residential Energy Community Survey (RECS) 2015 (latest available) [1]. We assume that this fraction is equal to the number of households having electric cooling equipment divided by the total number of occupied household units.

2.2.1.2. Model Fitting and Error Calculation for Electricity Usage Model

$$RSS_{elec,b} = \sum_{m=1}^{12} \left[E_{b,m}^{(act,2015)} - \sum_{c \in NYS} \sum_{t \in m} \hat{E}_{b,c,t} \right]^2 \tag{2}$$

The decision variables e_b^{const}, e_b^{+}, and e_b^{-} are selected for the residential and commercial building classes so that the residual sum of squares error $RSS_{elec,b}$ is minimized [8]. The minimization is performed against the actual statewide monthly electricity usage for each building class in 2019, as obtained from publicly available retail sales data of the U.S. Energy Information Administration (EIA) [19].

The estimated peak electricity demand is calculated for each building class in each county by taking the maximum of the estimated hourly electricity demand:

$$P_c = max_t \sum_b \hat{E}_{b,c,t} \tag{3}$$

2.2.1.3. Temperature-Dependent Fossil Fuel Usage Model

The estimated value of temperature-dependent fossil fuel usage [8], i.e., for heating, calculated for each building class, b, in each county, c, at each time step, t as shown in Equation 4, is found by splitting the total monthly fossil fuel usage of a county into two parts: temperature-independent fossil fuel usage (f_b^{const}), i.e., fossil fuel used for cooking; and decreasing-temperature-dependent fossil fuel usage (f_b^-), i.e., fossil fuel used for heating. There is no increasing-temperature-dependent fossil fuel usage because the underlying data shows that not much fossil fuel is used for cooling purposes in either building class. The building floor area for each building class (residential and commercial), b, and county, c, $A_{b,c}$, is determined using the NYS Tax Parcel dataset. Lastly, $p_{FF,Heat,b,c}$ is the fraction of floor area of each building class in each county that uses fossil fuel for heating ($Heat$).

Similar to the temperature-dependent electricity usage model, one important change introduced is the use of variable internal setpoint temperatures. The thermostat setpoint for residential buildings is based on the optimum seasonal temperature for human comfort as shown by Ecobee smart-thermostat data. The hourly reference temperatures are taken as statewide average Ecobee thermostat hourly setpoint temperatures. Ecobee thermostats are only employed in residential buildings and therefore these temperatures were not used for commercial buildings. The reference temperature for actual commercial buildings varies significantly depending on the type of building and many other factors, but a value of $T_{ref,com}$ = 68°F was taken as a constant setpoint temperature for all commercial buildings for enhanced energy savings [26]. $T_{c,t}$ is the hourly outdoor temperature for each county in NYS was taken from NSRDB.

$$\hat{F}_{b,c,t} = A_{b,c}[f_b^{const} + p_{FF,Heat,b,c} f_b^- \left(T_{ref,b} - T_{c,t}\right)^+] \tag{4}$$

The '+' outside the parentheses indicates that the value of the term is zero when negative. The decreasing temperature-dependent terms are weighted by the fraction of building floor area using fossil fuel for heating. We assume that the primary end-use of decreasing temperature-dependent fossil fuel usage is heating only.

The fraction of floor area of the residential building class in each county that uses electricity for heating, $p_{FF,Heat,res,c}$ was calculated by dividing the building class floor area using electricity for heating with the

total building class floor area of all residential buildings. However, for the NYC counties, $p_{FF,Heat,res,c}$ was determined using data from the 2019 American Community Survey Data [25]. It was assumed that the county-wise commercial floor area using fossil fuel for heating was scaled linearly with the fraction of households using air-conditioning in each county, i.e., $p_{FF,Heat,com,c}$, was calculated using the fraction of businesses having fossil-fuel-based heating in place as per NYSERDA Commercial Baseline Study [24]. The commercial baseline study was conducted by NYSERDA in collaboration with the New York State Department of Public Service (DPS). The study assesses the existing commercial building stock and energy-consuming equipment (e.g., HVAC), including equipment count and characteristics.

2.2.1.4. Model Fitting and Error Calculation for Fossil fuel Usage Model

$$RSS_{FF,b} = \sum_{m=1}^{12} \left[F_{b,m}^{(act,2019)} - \sum_{c \in NYS} \sum_{t \in m} \hat{F}_{b,c,t} \right]^2 \tag{5}$$

The decision variables f_b^{const} and f_b^{-} in Equation 4 are selected for the residential and commercial building classes in the state of New York to minimize the residual sum of squares error, $RSS_{FF,b}$ [8]. The minimization is performed with respect to the actual statewide monthly electricity usage for each building class in 2019, as obtained from publicly available retail sales data of the U.S. Energy Information Administration (EIA) [19].

Fuel oil and propane are delivered in bulk and stored on-site for a season or many years, whereas natural gas use is measured monthly. Therefore, the total fossil fuel consumption was scaled with natural gas consumption on a monthly basis. The unit of measurement for these fossil fuels was in million cubic feet or gallons which was then converted to Watt-hours before plugging these values into the minimization equation. Thus, the monthly, m, fossil fuel usage, $F_{b,m}^{(act,2019)}$, for each building class, b, is computed using monthly natural gas usage, $F_{NG,b,m}^{(act,2019)}$, annual distillate fuel oil usage, $F_{FOK,b}^{(act,2019)}$, and annual propane usage, $F_{prop,b}^{(act,2019)}$] data from EIA[19]—[21]:

$$F_{b,m}^{(act,2019)} = \frac{F_{NG,b,m}^{(act,2019)}}{\sum_{m=1}^{12} F_{NG,b,m}^{(act,2019)}} \left[\sum_{m=1}^{12} F_{NG,b,m}^{(act,2019)} + F_{FOK,b}^{(act,2019)} + F_{prop,b}^{(act,2019)} \right] \tag{6}$$

2.2.2. Heating Electrification Model

To develop a heating electrification model, we converted the amount of fossil fuels used into electricity used with the help of the model equations explained in section 2.1.2.1. below. The electrification equations were inspired by the study performed at Columbia University [8]. This study introduces a new model equation to evaluate temperature-dependent COP in section 2.1.2.2. which is used calculate to heating electrification demand.

2.2.2.1. Conversion of Fossil Fuel Heating to Electric Heating

To compute the projected new county-level temperature-dependent electricity demand for each building class, we create different scenarios of heat pump penetration levels to indicate the fraction of existing fossil fuel heating being replaced.

The electricity demand upon converting the fossil-fuel-based heating systems to HP-based heating systems for each building class in each county at each time step, $E_{b,c,t}$, for a given HP penetration, $x_{b,c}$ = {0: 1}, is given by:

$$E_{b,c,t}(x_{b,c}) = x_{b,c} \times A_{b,c} \times \frac{\left[f_b^{-}(T_{ref,b} - T_{c,t}) \right] \eta_{FF}}{COP\ (T_{c,t})} \tag{7}$$

We generate various other scenarios at different levels of electrification using different values of $x_{b,c}$. The electricity demand, $E_{b,c,t}$, for a given building class and HP penetration is computed by:

$$E_{b,c,t}(x_{b,c}) = A_{b,c}[e_b^{const} + p_{AC,b,c} e_b^{+} (T_{c,t} - T_{ref,b})^{+} + p_{elec,Heat,b,c} e_b^{-} (T_{ref,b} - T_{c,t})^{+} + x_{b,c} \times \frac{\left[f_b^{-}(T_{ref,b} - T_{c,t}) \right] \eta_{FF}}{COP\ (T_{c,t})} \tag{8}$$

2.2.2.2. Coefficient of Performance Modelling Equation

To create a temperature-dependent COP model that is representative of real-life HPs, we used the field measurement data of ductless HPs (or MSHPs) from a study conducted in the Northwest US by the National Renewable Energy Laboratory (NREL) [27]. The purpose of the NREL study was to evaluate two tools: 1) EnergyPlus and 2) the Simple Energy and Enthalpy Model (SEEM) to compare their energy consumption projections and make model improvements where there were substantial variances. The purpose was to minimize the variability in energy savings estimates and have a better understanding of simulation engine limits for more reliable savings estimates of efficiency measures and regional potential.

While modeling the temperature-dependent COP equation, variable speed HPs were not considered because the calculations for such HPs are more complex than a COP versus outdoor temperature equation. The COP is also a function of compressor speed (and as a result, the heating load of the space and the heat pump size), outdoor humidity, and indoor temperature as well. There are also non-linear effects like defrosting and backup heat. All these effects can cause real-world COP to vary +/- 30%.

Using Ref [27], we have derived outdoor temperature-dependent COP model equations that represent the range of real-world COP values seen in the figure.

$$COP_{HP}(T) = \begin{cases} -0.0044 * T^2 + 0.096 * T + 3.7406 & , \quad 4.5 < T < 27 \\ 0.09 * T + 3.6 & , \quad -18 < T < 4.5 \\ 1 & , \quad T \leq -18 \end{cases} \tag{9}$$

Here, temperature values—represented by T—are in degrees Celsius.

Several simplifying assumptions were made while developing the heating electrification model. We assumed that the entire electrification is done with only cold climate air source heat pumps. We have not considered GSHPs because these kinds of HPs are not very widely used yet in residential or commercial buildings. We also assumed a single existing fossil fuel heating system efficiency, η_{FF} = 0.78, based on Ref [8].

Since we did not use any building energy models, we base the heating electrification calculation on these derived equations. We make some assumptions to match the actual performance of a heat pump as it varies with ambient temperature. We assume that once the predicted COP goes below 1 at minimum heat pump operating temperatures, the system will function as a complete electric resistance system with an

efficiency of 1. As a result, the overall HP system COP also becomes 1. For a few higher temperature points (T > 27) where the field measurement data was not available, we have not considered any values for COP.

The graph in **Figure 3** shows the field measurement data values from the NREL study and shows the fitted equations that would be used to represent temperature-dependent COP. For temperatures below ~4 degrees, COP follows the trend shown by the blue line and for temperatures above ~4 degrees, COP follows the trend shown by the orange dotted line.

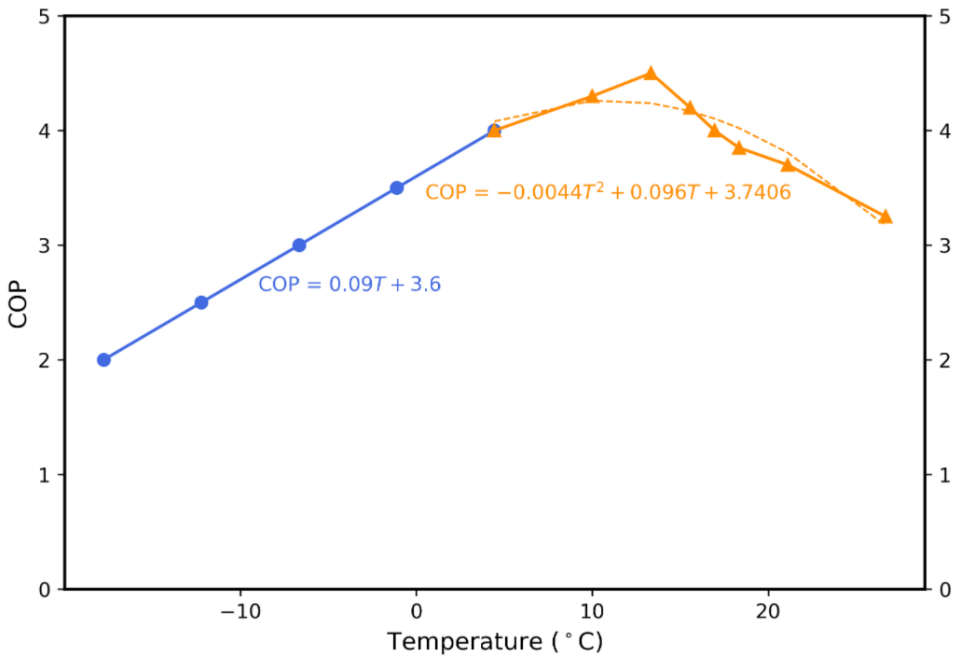


Figure 3 shows the model temperature-dependent coefficient of performance. For temperatures below ~4 degrees, COP follows the trend shown by the blue line and for temperatures above ~4 degrees, COP follows the trend shown by the orange dotted line.

3. Results and Discussion

In this section, we first discuss the qualitative and quantitative model validation techniques used by this study, followed by the results obtained upon the electrification of the heating systems in NYS.

3.1. Model Validation

Figure 4a validates the methodology followed in this study by qualitatively comparing the metered NYISO data from 2019 with the demand values calculated using the model equations. The metered NYISO data is a combination of all the sectors, viz, residential, commercial, industrial, and transportation, and is highlighted using a black line. The Estimated R+C curve marked using a blue dotted line shows the demand values calculated by model for the residential (R) and commercial (C) sectors that are of primary focus in this study. The Estimated R+C demand curve was generated using Ecobee data for residential buildings and constant setpoint temperature for commercial buildings. The orange dotted line represents the Estimated Total which is the sum of residential and commercial demand values calculated by the model plus the EIA industrial and transportation [24] data to get the total of all sectors. The demand values calculated by the model and the NYISO data are at a daily resolution. However, EIA provides sector-wise monthly data, so we have assumed that the industrial and transportation values were divided equally among all days of the year to generate the daily curve.

We can see that the Total NYISO and the Estimated Total values are a close match with an average difference of 7% between them. There is an average difference of about 25% between the Estimated R+C values and the Total NYISO data. The model generated peaks match the peaks in the NYISO data showing that the model accurately evaluates the rise and fall in demand. Hence, we can say that the difference between the Estimated R+C curve is attributed to the absence of industrial and transportation sector demand. This proves that the estimated residential and commercial values are closely replicated by the model, and this model can be further used to calculate the total heating electrification demand—as elaborated in the later sections.

Figure 4b quantifies the above comparison. To do this, we calculate the two relative differences: i) the relative difference between the daily Total NYISO demand data and the Estimated Total demand data, and ii) the relative difference between monthly Total NYISO demand data and the Total EIA data (which was used to model the calculated data). Each of the NYISO, the EIA, and the Estimated Total demand data are a combination of the electricity consumed by the residential, commercial, industrial, and transportation sectors. i) is represented by box and whisker plots, and the colored dots represent ii). We

can see that the two relative differences are comparable with an average relative difference of 0.07. Therefore, we can use the calculated data to justify our model.

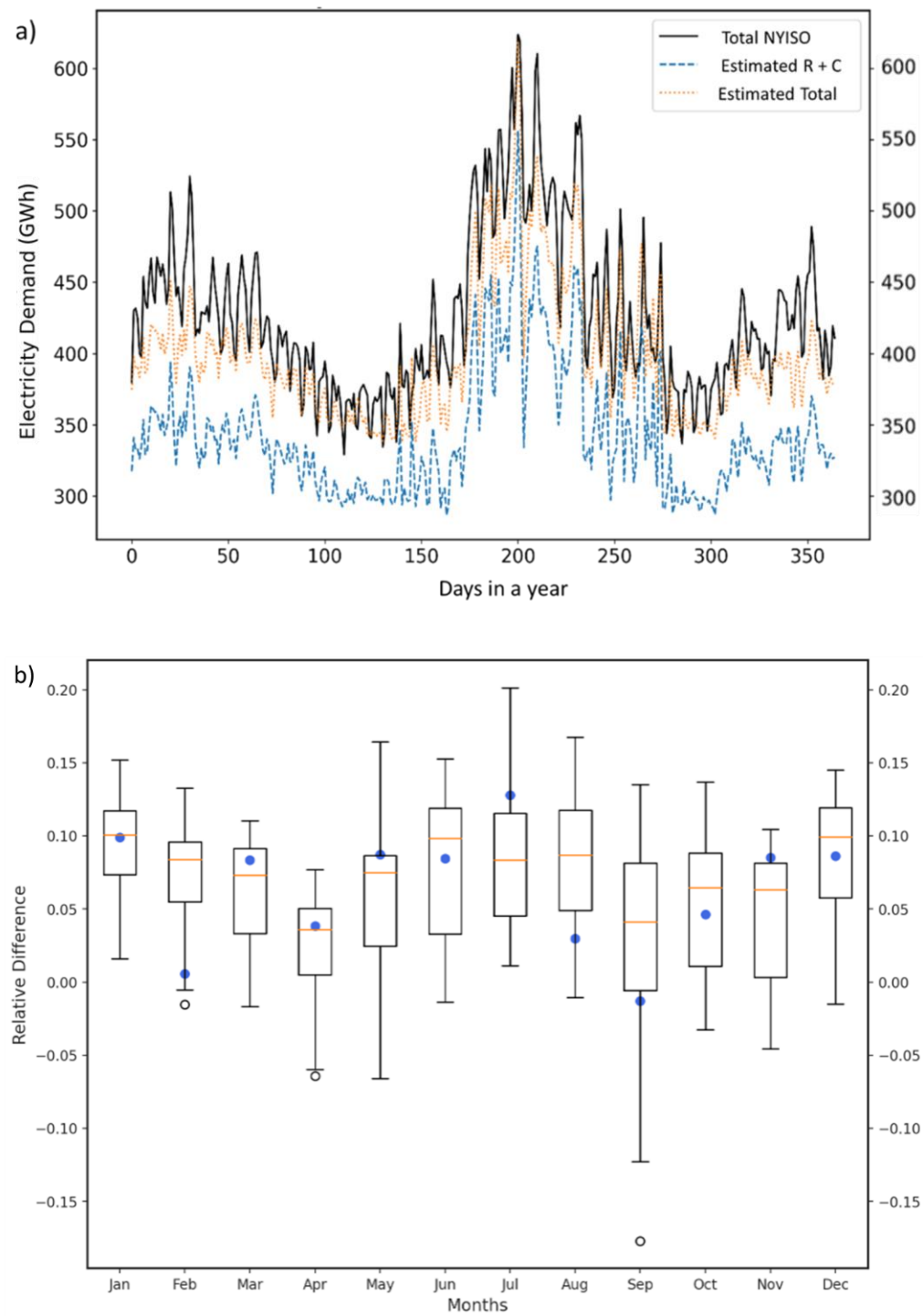


Figure 4 *a) validates the methodology followed in this study by qualitatively comparing the metered NYISO data from 2019 with the demand values calculated using the model equations. Total NYISO and Estimated Total values*

match closely, and the model-generated peaks match the peaks shown by the NYISO data indicating model accuracy. Hence, the difference between Estimated R+C and NYISO curves is attributed to the absence of industrial and transportation sector values. This proves that the estimated residential and commercial (Estimated R+C) values are closely replicated by the model, and this model can be further used to calculate the total heating electrification demand. b) comparison of the two relative differences between (NYISO and Estimated Total Demand) and (NYISO and Total EIA Demand) data using box plots; box plots show (NYISO-Estimated Total Demand)/NYISO, marked points show (NYISO-Total EIA Demand)/NYISO.

To further reinforce the reliability of the model, this relative difference was also statistically justified. The **statistical** method used for this was the two-sample t-test with which we checked whether the mean of the difference between i) and ii) was in an acceptable range or not.

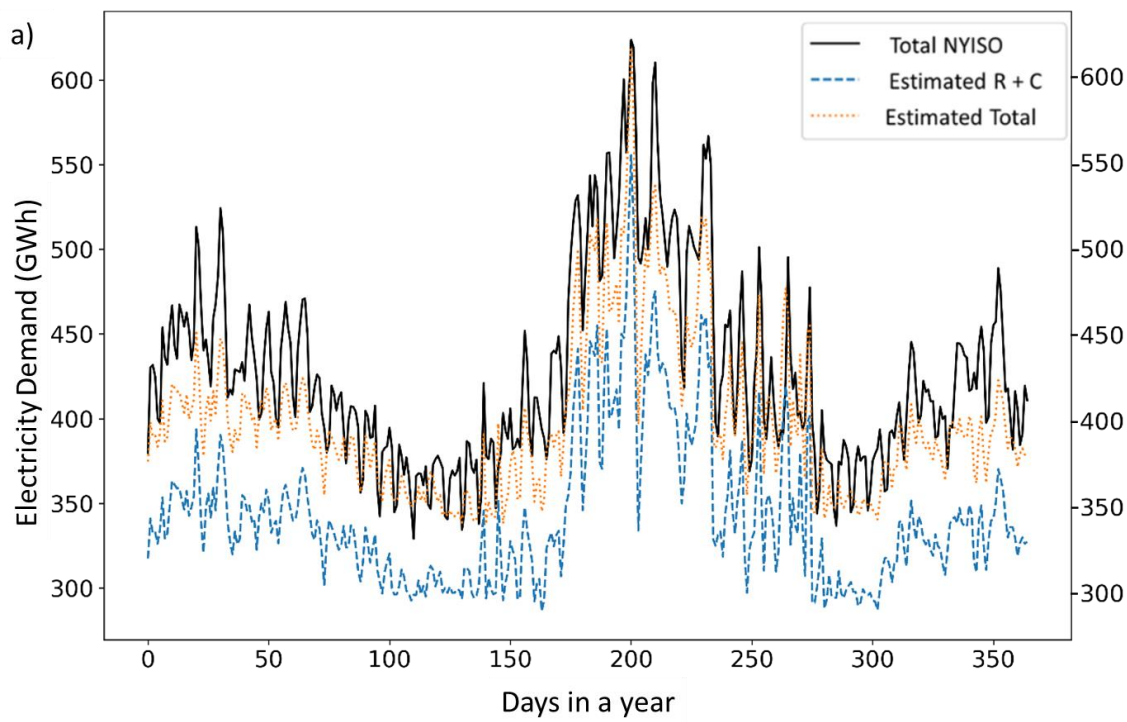
The two hypotheses for this particular two-sample t-test were: 1) the null hypothesis H0 which assumes that the two population means are equal ($\mu 1 = \mu 2$) and 2) the alternative hypothesis HA which assumes that the two population means are not equal ($\mu 1 \neq \mu 2$).

The t-statistic value obtained for the test was -0.1 and the p-value obtained was 0.92. This is a positive indicator because, by definition, a t-value of 0 signifies that the sample results exactly equal the null hypothesis. As the difference between the sample data and the null hypothesis increases, so does the absolute value of the t-value. Since we have a significantly low t-value, this means that our null hypothesis is true.

Because the p-value of our test (0.92) is greater than alpha = 0.05, we fail to reject the null hypothesis of the test. We do not have sufficient evidence to say that the mean difference in electricity demand between the two populations is different.

Figures 5a and 5b compare the hourly electricity demand trend of Total NYISO values with the Estimated Total. 5a shows the demand comparison when calculations were made using variable internal setpoint temperature values taken from Ecobee data. On the other hand, 5b shows the demand comparison when calculations were made using a single constant internal setpoint temperature value of 20 degrees Celsius [28]. On studying the data, we observe that the calculated energy demand follows the trend better in 5a than it does in 5b. The areas of comparison are highlighted by the black boxes in 5b. We can see that the Estimated Total demand values exceed the Total NYISO values as reflected in the black box drawn at the top of figure 5b. This can be attributed to the larger temperature delta values obtained with constant setpoint than would be obtained with variable temperature values. The black box on the lower half of the

figure depicts that lesser variability is introduced in the demand values when a constant setpoint temperature is used, as a result of this, the values follow a flatter trend.



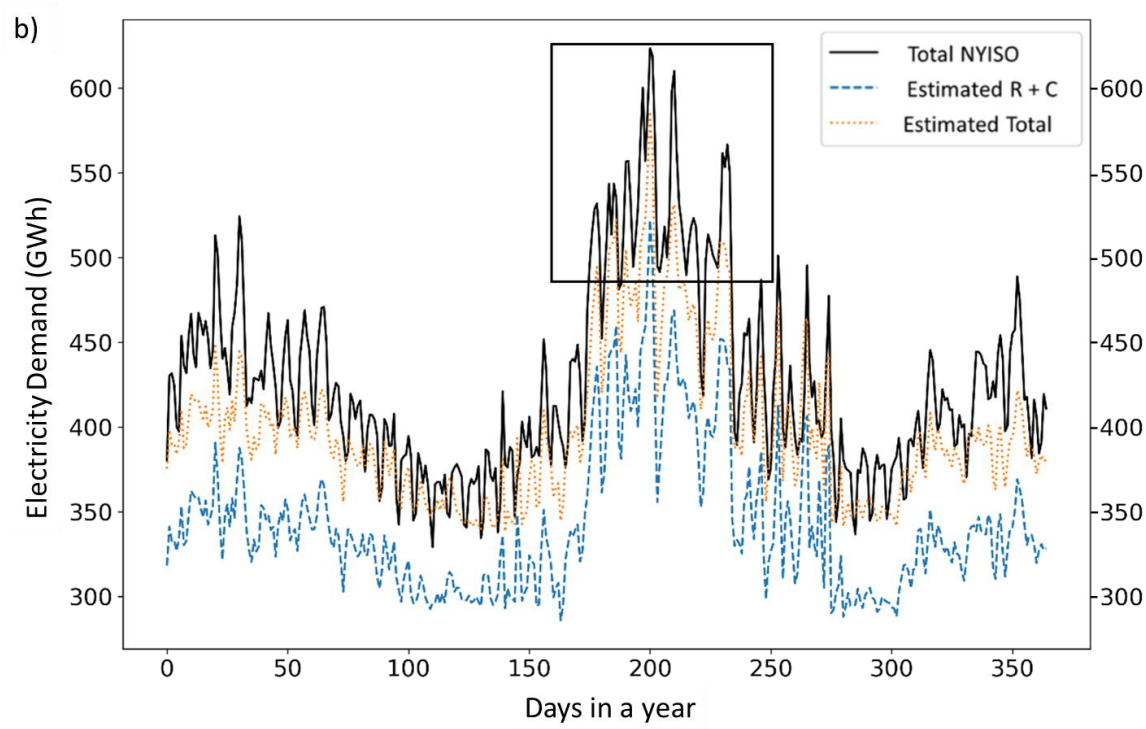


Figure 5 Comparison between the trend of NYISO vs. estimated electricity demand in case a) with variable internal setpoint temperature values taken from Ecobee data and case b) with a single constant internal setpoint temperature value. We can see that the calculated energy demand follows the trend better in case a) when compared to case b) with the area of comparison highlighted by the black boxe on b). We can see that the Estimated Total demand values at the peaks are lesser than the Total NYISO values in the box at the top in case b) and are matched better in case a). This can be attributed to inaccurate temperature delta values obtained with constant setpoint than would be obtained with variable temperature values..

3.2. Post-Electrification Results

The results shown here depict the increase in electricity requirement when heating systems are electrified across the state.

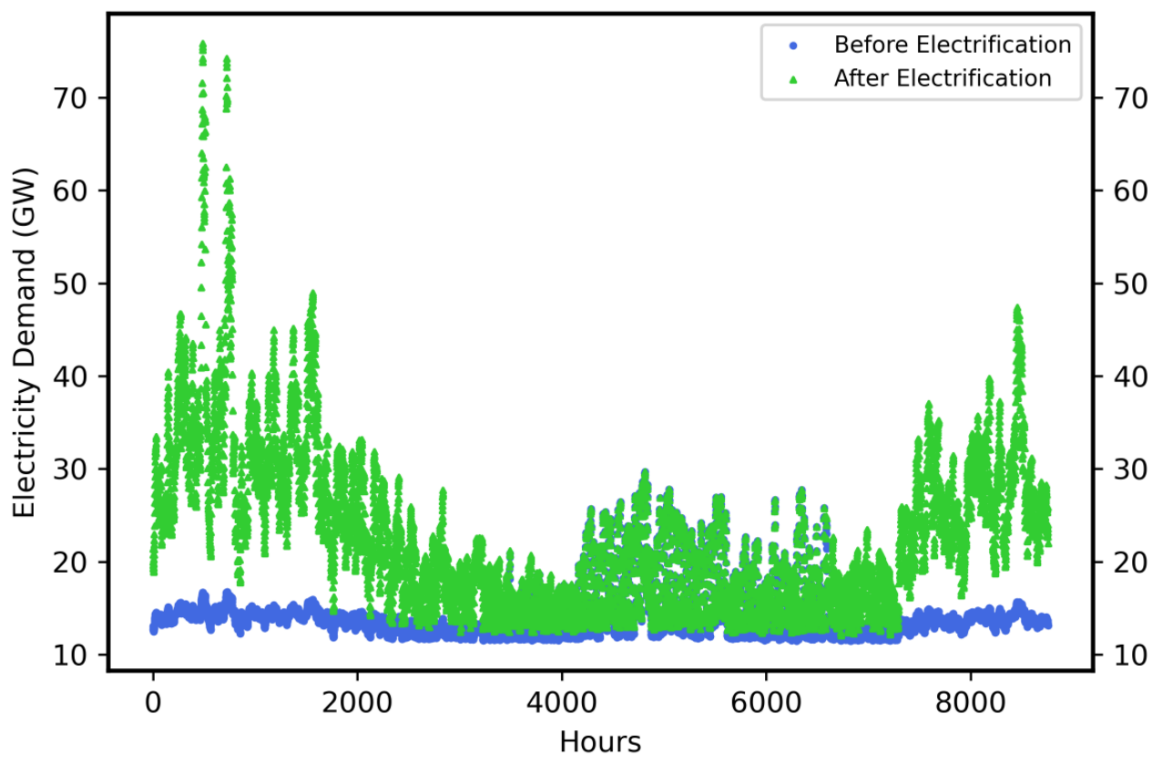
As seen previously in figure 4a, New York State has an apparent dearth of electric heating. On the other hand, as per data published by the American Community Survey, there are more electric cooling systems in NYS as compared to heating systems. Therefore, in the current scenario, the demand for electric cooling is much higher than the demand for electric heating, with the demand for electricity peaking during the summer months because of a surge in the usage of electric cooling.

Figure 6a compares the total electricity demand in residential and commercial buildings before and after the electrification of heating systems. The difference between before and after shows an approximate increase of 48% in the total electricity demand in the residential and commercial sectors. Under this scenario, which highlights an increase in the adoption of electric heating systems, it is likely that our previously noted demand peak will shift from summer to winter. Additionally, the winter peak would be 2.5 times the summer peak. A direct consequence of this demand shift would be a greater likelihood of dispatching peaker plants during winters, changes in wholesale electricity pricing, grid emissions profiles, and power plant maintenance schedules. As a result of this increase in demand, NYS will be tasked with finding the cleanest electricity supply sources.

However, an offset would also be provided if the newly installed ASHPs were to be used for cooling because their improved efficiency would further lower the summer demand. For our study, we have assumed that the already existing cooling systems would remain in place, and therefore this offset is not accounted for in the overall yearly demand.

Figure 6b shows the daily average temperature values in NYS to show how cold the day was for which the demand was calculated.

a)



b)

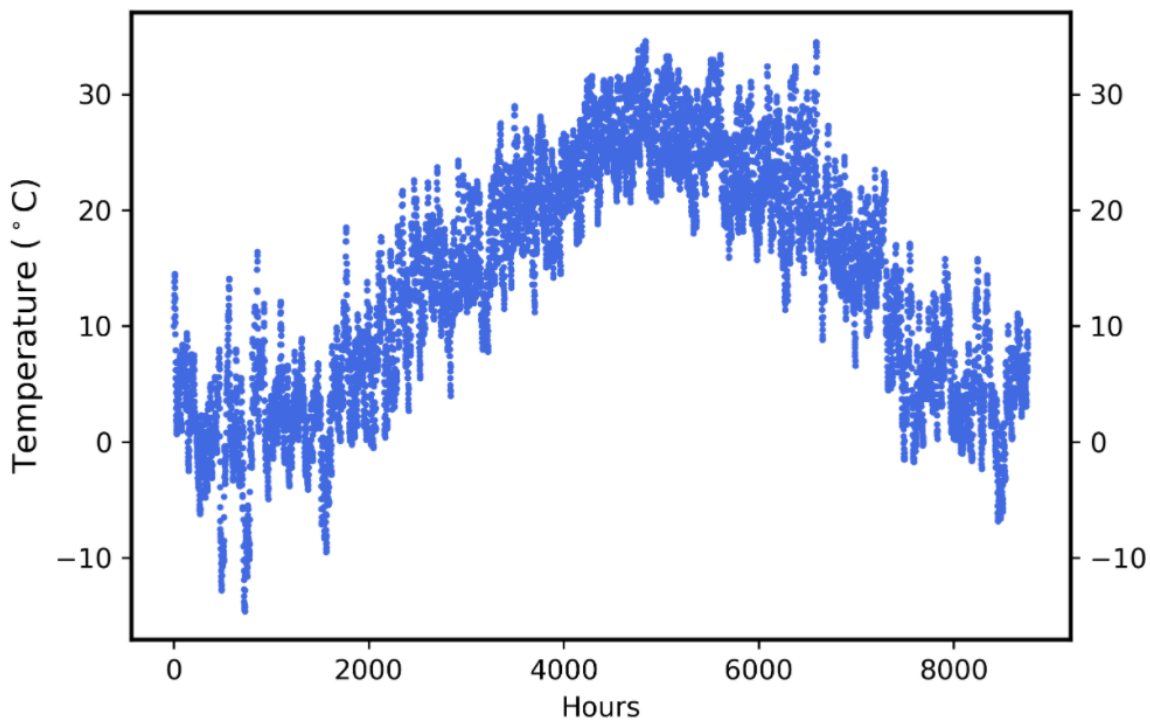


Figure 6 a) calculated hourly electricity demand before and after electrification at 100% HP penetration b) the temperature trend for the year 2019 for which the demand was calculated.

To represent the change in electricity demand values when advanced high-efficiency heating systems are used, we generate two scenarios of electrification using a high-efficiency ASHP and an ultra-high efficiency MSHP. The efficiency of these advanced HPs was derived from a study performed by White et al. [7] on the electrification of heating systems in Texas. The high-efficiency heat pump is a high-end variable speed ASHP that is currently available with a heating seasonal performance factor (HSPF) of 10.0, whereas the ultra-high efficiency MSHP has a heating seasonal performance factor (HSPF) of 14 and represents the most efficient electric space heating technology currently available.

We used the same method as the one described in Section 2.2.2.2 to obtain the temperature-dependent COP of the high-efficiency ASHP and the ultra-high efficiency MSHP from Figure 3 in Ref. [7]. We use the set of equations (indicated by distinct colors) in **figures 7a and 8a** to calculate temperature-dependent COP for the various heating system operating temperatures.

Figures 7b and 8b below show the approximate values for electricity demand before and after electrification in the scenarios where high-efficiency variable-speed HPs were used at a 100% penetration. We can see that the overall increase in demand is 44% when using a high-efficiency ASHP, which is lower than the 48% increase with a standard current HP. With high-efficiency ASHP, the peak demand also reduces by 42%. On the other hand, when an ultra-high efficiency MSHP is employed, the

overall increase in demand is 31% as compared to a 48% increase with a standard current HP, and the peak demand reduces by 45%. Therefore, these calculations clearly represent the extent to which improvements in HP efficiency can reduce demand and benefit in moving electrification goals forward.

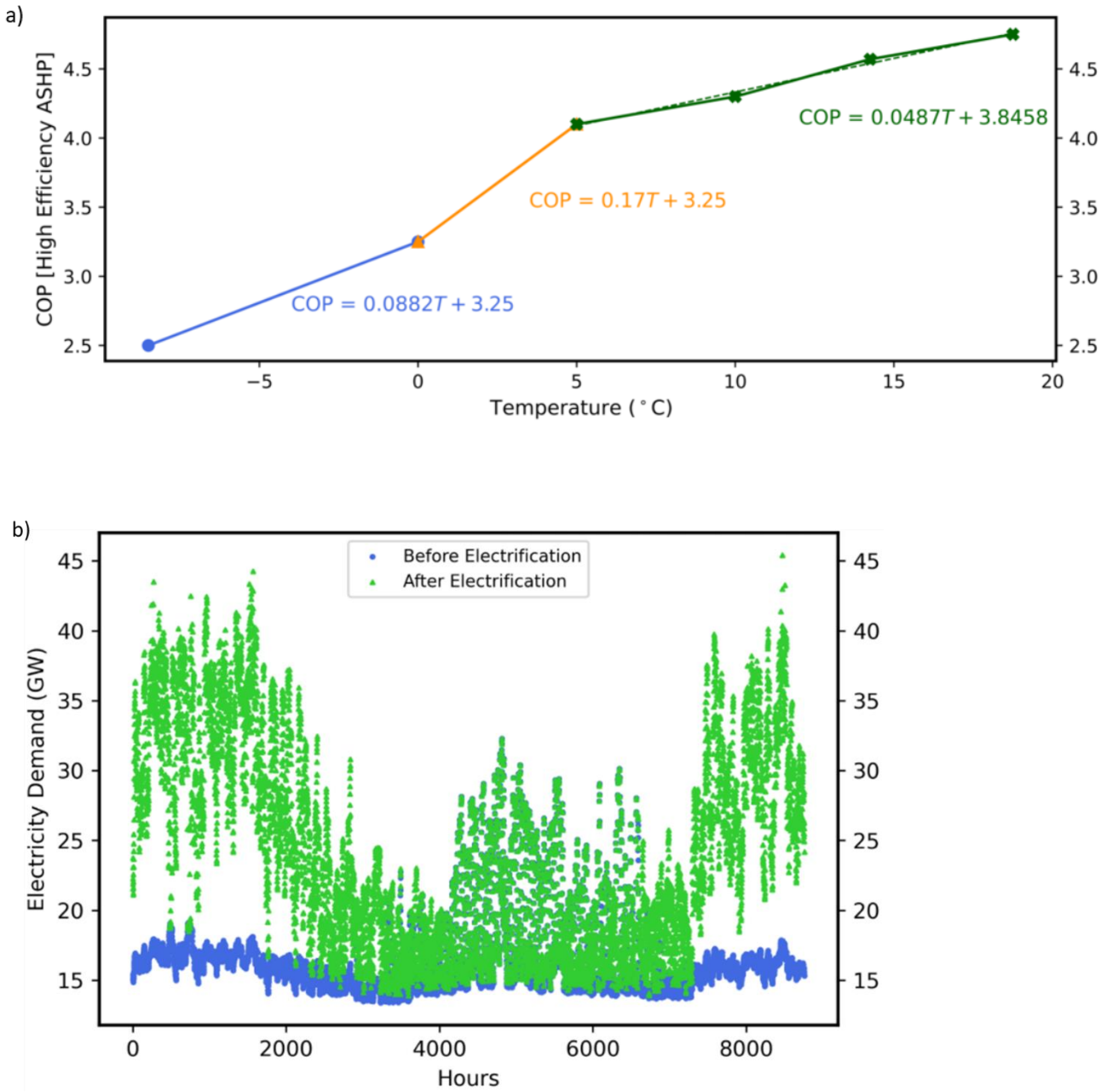
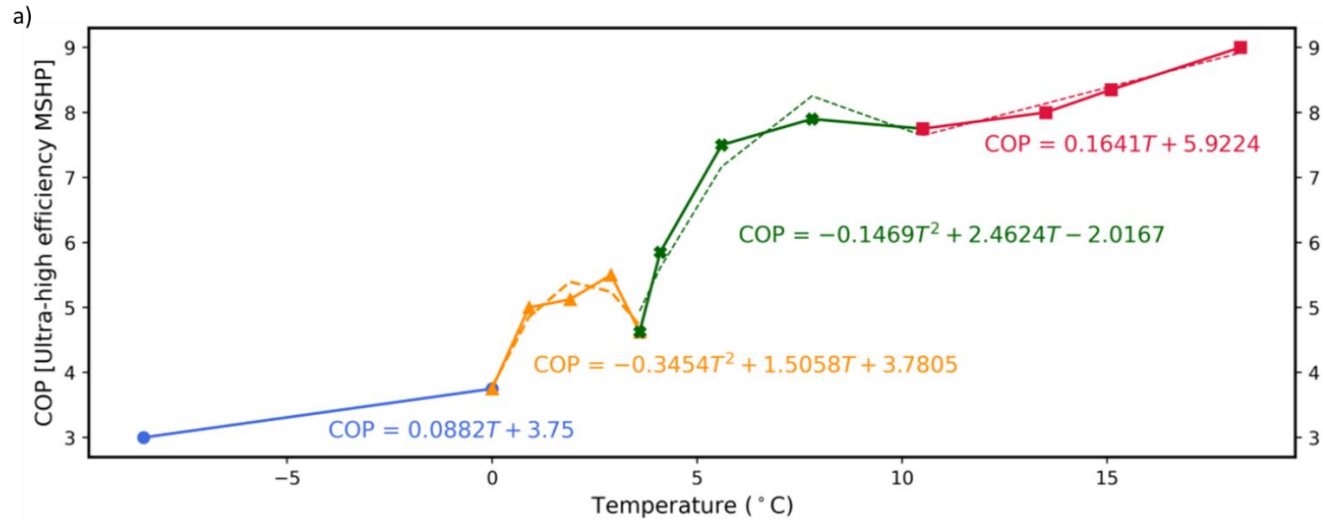


Figure 7 a) shows the COP of a variable-speed high-efficiency ASHP with an HSPF of 10.0. To calculate temperature-dependent COP for the various heating system operating temperatures, we used a set of equations indicated by

different colors in the graph. b) shows that by using a high-efficiency ASHP, the overall increase in demand reduces to 44% as compared to 48% using a standard current HP. The peak demand reduces by 42%.



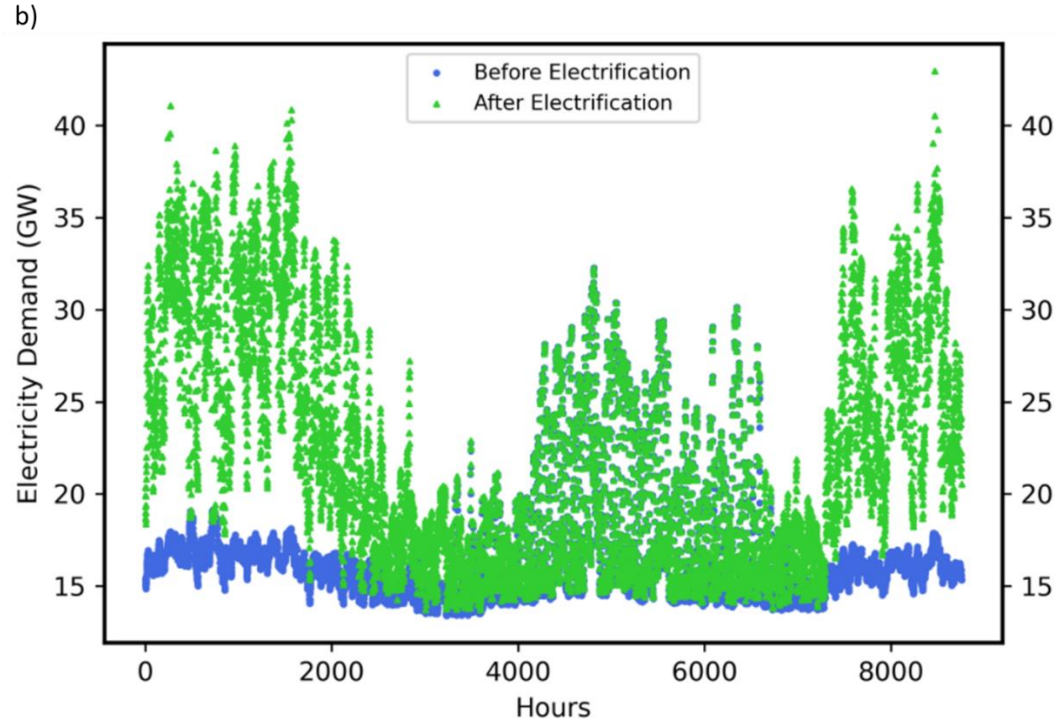


Figure 8 a) shows the COP of a variable-speed ultra-high efficiency MSHP with an HSPF of 14. To calculate temperature-dependent COP for the various heating system operating temperatures, we used a set of equations indicated by different colors in the graph. b) shows that by using an ultra-high efficiency MSHP, the overall increase in demand reduces to 31% as compared to 48% using a standard current HP. The peak demand reduces by 45%.

We generate three scenarios inspired by a report on ramping up heat pump adoption in New York State [29] prepared by Vermont Energy Investment Corporation (VEIC) and funded by the Natural Resources Defense Council (NRDC). We develop three scenarios of heat pump adoption rate to attain the NYS energy efficiency goals: baseline, moderate, and high growth rate scenarios. All these electrification scenarios are developed using the current standard HP.

The **baseline** scenario considers the average heat pump penetration rate of 5% for the year 2019 as per a report prepared by NYSERDA [30]. We do not consider any enhanced growth due to changes in policies in this scenario. **Figure 9** shows the electricity demand upon electrification at this rate.

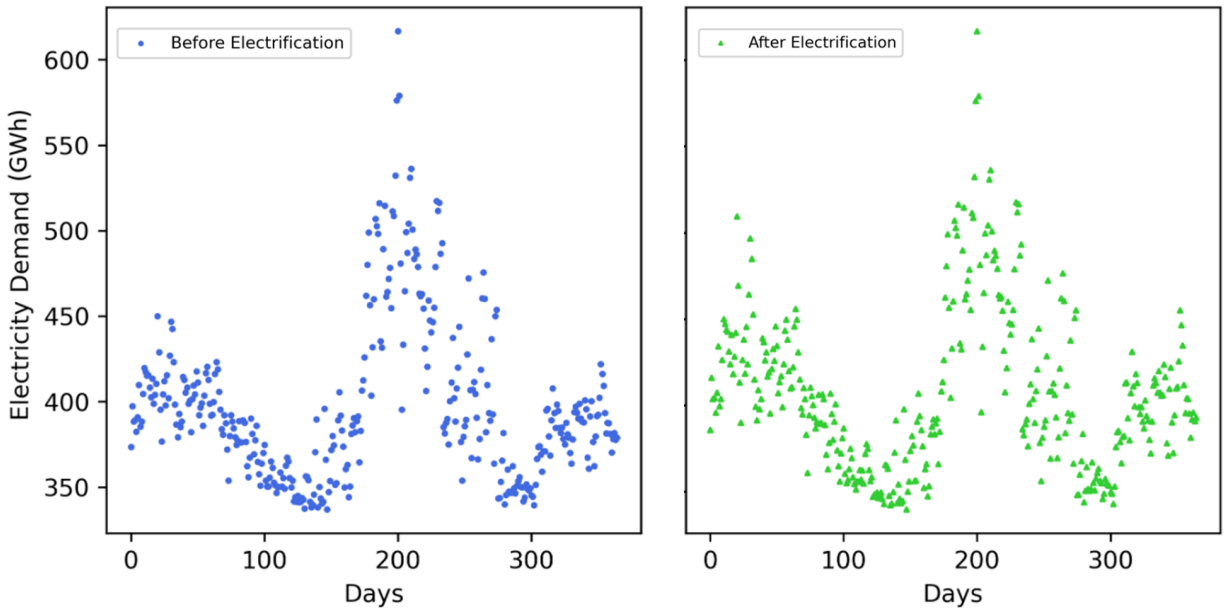


Figure 9 Baseline scenario with a projected penetration of 5% in NYS for the year 2019 as per a study by NYSERDA

In the **moderate** growth scenario, we consider a cumulative average growth rate of 12% calculated based on the adoption potential trajectory from 2019 to 2025 assumed in the study by NYSERDA [31]. Because the cold climate heat pump market in New York is still in its infancy, data to support adoption trajectory assumptions is scarce. **Figure 10** shows the electricity demand upon electrification at this rate.

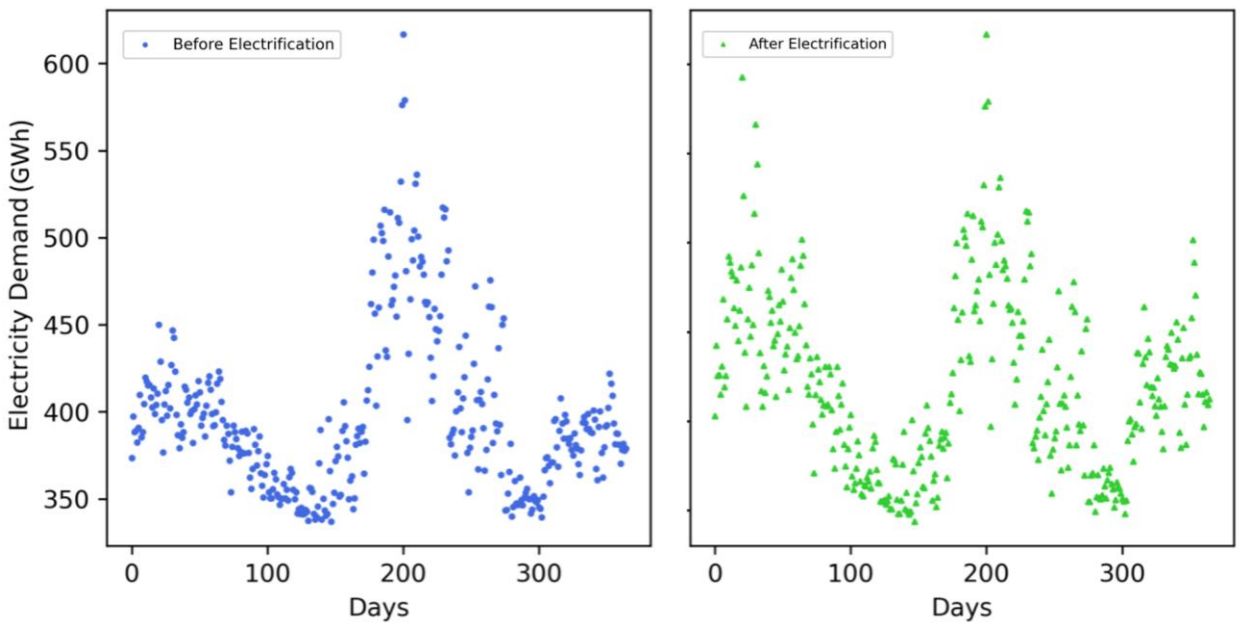


Figure 10 Moderate growth scenario with 12% annual heating system electrification in accordance with the adoption potential trajectory assumed in a report by NYSERDA.

The **high** growth scenario is based on a report by VEIC which assumes that the heat pump market is supported by aggressive program initiatives [29]. The adoption rate is capped at 59%, which is the national average annual growth rate in the solar photovoltaic (PV) sector over the last ten years (from 2018). **Figure 11** shows the electricity demand upon electrification at this rate.

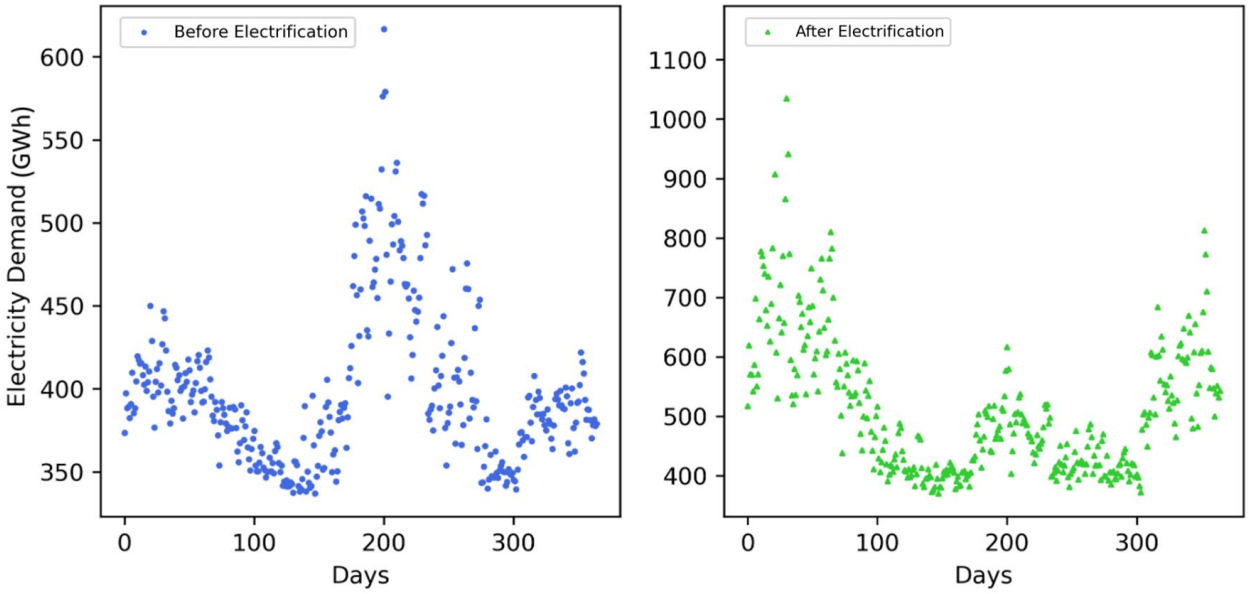


Figure 11 High growth scenario with a growth rate of 59% based on a report by VEIC which assumes the heat pump growth rate to be the same as the growth rate of solar PV adoption in NYS for the next decade

Figure 12 shows load duration curves of pre- and post-electrification levels. This curve shows a relationship between load and time. The coordinates indicating load are shown in decreasing order of magnitude. We see that the exponentially high demand hours are fewer in number, and it is for those hours that the peaker plants would need to be used to make up for the excess demand.

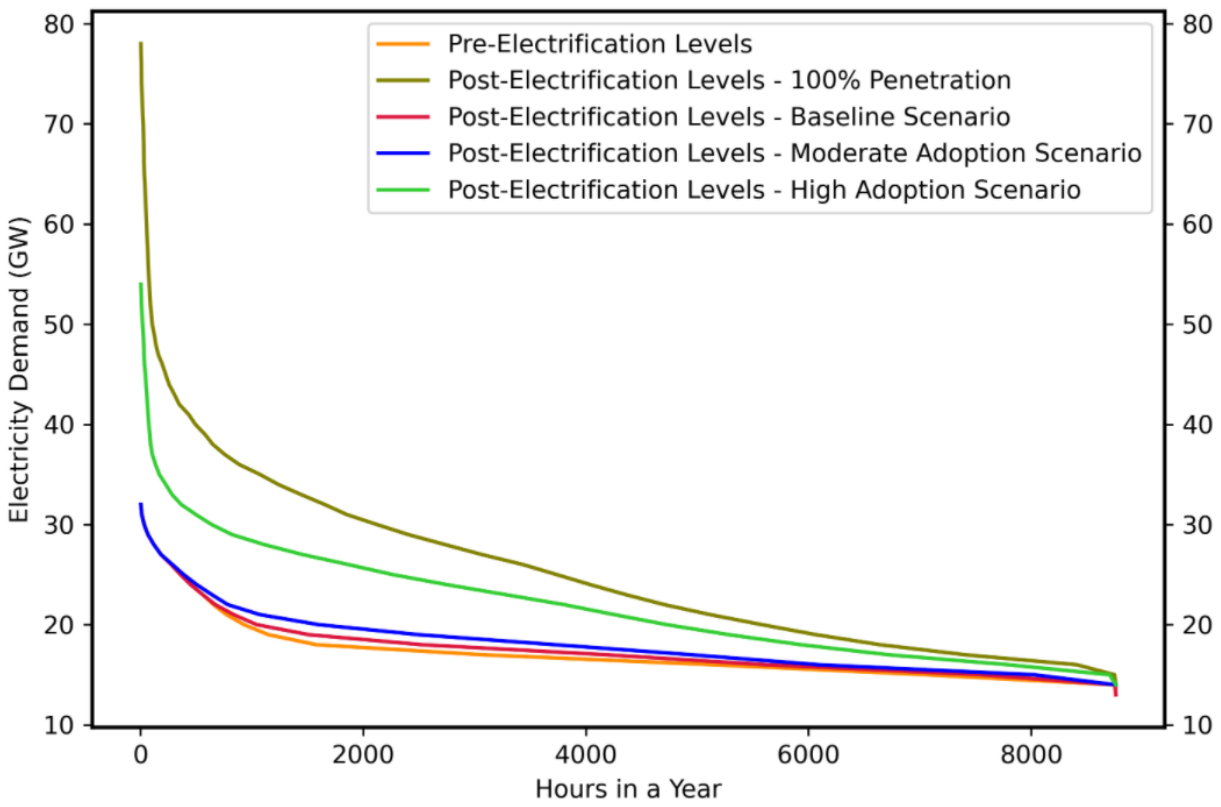


Figure 12 Load duration curves for calculated values of before and after electrification (using standard HP) levels. These curves show the load required to meet every hour of the State's heating demand from highest demand hour to lowest demand hour for each of the HP adoption scenarios.

4. Conclusion

In this analysis, we utilized New York State's Tax Parcel Data and used optimization and statistical techniques to quantify the electricity usage required post the electrification of heating systems. The results were generated for both residential and commercial sectors and summed to give the statewide demand values and show a staggering 48% increase in the total electricity required in NYS after 100% HP penetration. Another finding that should not be overlooked by policymakers and utility service providers is the apparent demand shift from a summer peak to a winter peak that is forecasted to occur because of an electrified New York State.

We also evaluated various electrification scenarios using highly efficient HP models. We show that technologically advanced HPs that are currently available can reduce electrification demand by as much as 17% as compared to the standard HPs. In addition, we also develop scenarios of different HP penetration values based on a number of studies on heating electrification conducted in NYS. We use these values to help us visualize the potential electricity demand profiles under various rates of HP adoption.

Most fundamentally, this study provides a base to generate insights into the grid implications of heating electrification initiatives and regulations. Therefore, it is valuable to grid planners and policymakers, particularly those in regions where demands are largely driven by summer cooling systems. These regions need to prepare for the shift in seasonal peaks. Our analysis also showed how the demand can vary across geographies, and thus it can assist in the development of specialized electrification plans with numerous localized benefits. Moreover, the approach used in this study might be applied to other areas to investigate the effects of space heating electrification with various other heat pump types (for example, GSHPs) which might bring the demand down from the current levels.

The results of this work can be used by future studies to determine the impact of such an increase in electricity demand. They can also be used to determine the amount of energy that can be provided by renewable resources, along with helping stakeholders evaluate if the current renewable infrastructure is enough to support the demand. If the current installations are not enough and new installations are needed, studies building on our research can also determine the most effective composition of renewable energy sources—such as wind [32], hydropower, solar etc.—that can be used to power the grid in New York State.

5. Future Work

There have been several studies on the calculation of heating demand. However, it is challenging to create one study that follows an end-all-be-all technique to solve this problem. This is because heating demand depends on several variables, and a single study cannot address all of them. Some internal factors like age of the house [33],[34], occupant behavior [35–37], inside setpoint temperature [38–40], insulation and ventilation [42],[43], type of heating equipment used and its efficiency [44]; and external factors like building shape [45], ambient temperature [40], solar irradiation [46], and sometimes even wind speed [47],[48] impact the amount of electricity that will be required to heat the buildings.

Some areas to improve upon in this study would be the introduction of daylight hours while making demand calculations and considering occupancy hours. Making the model dependent on these factors would improve its accuracy by bringing it closer to a real-life scenario. Additionally, using different HP types (e.g., GSHPs) to model the temperature-dependent COP will make the calculations more inclusive. Because there is constant ongoing research to make GSHPs [51] and technologically advanced HPs [52] easily adoptable by the public, using newer models and variations of HPs will ensure subsequent research to be more future-oriented. Our study does not take into consideration the offset of electricity demand by using HPs for cooling in a real-world scenario. This will add another layer of complexity that can be analyzed further in future studies.

References

[1] "Residential Energy Consumption Survey," *https://www.eia.gov/consumption/residential/data/2015/index.php/index.cfm?view=consumption).* .

[2] Z. E. Lee and K. M. Zhang, "Scalable identification and control of residential heat pumps: A minimal hardware approach," *Applied Energy*, vol. 286, Mar. 2021, doi: 10.1016/j.apenergy.2021.116544.

[3] M. Lynch, "The Building Electrification Primer for City-Utility Coordination Acknowledgements."

[4] "Meeting Local Law 97 through Energy Efficiency and Beneficial Electrification." [Online]. Available: http://nyc.gov/nycha

[5] M. C. Baechler, T. L. Gilbride, P. C. Cole, M. G. Hefty, and K. Ruiz, "Guide to Determining Climate Regions by County PREPARED BY Pacific Northwest National Laboratory," 2015. [Online]. Available: www.buildingamerica.gov

[6] T. Mai *et al.*, "Electrification Futures Study: Scenarios of Electric Technology Adoption and Power Consumption for the United States." [Online]. Available: https://www.nrel.gov/docs/fy18osti/71500.pdf.

[7] P. R. White, J. D. Rhodes, E. J. H. Wilson, and M. E. Webber, "Quantifying the impact of residential space heating electrification on the Texas electric grid," *Applied Energy*, vol. 298, Sep. 2021, doi: 10.1016/j.apenergy.2021.117113.

[8] M. Waite and V. Modi, "Electricity Load Implications of Space Heating Decarbonization Pathways," *Joule*, vol. 4, no. 2, pp. 376–394, Feb. 2020, doi: 10.1016/j.joule.2019.11.011.

[9] Z. Verbai, Á. Lakatos, and F. Kalmár, "Prediction of energy demand for heating of residential buildings using variable degree day," *Energy*, vol. 76, pp. 780–787, Nov. 2014, doi: 10.1016/j.energy.2014.08.075.

[10] M. Olonscheck, A. Holsten, and J. P. Kropp, "Heating and cooling energy demand and related emissions of the German residential building stock under climate change," *Energy Policy*, vol. 39, no. 9, pp. 4795–4806, Sep. 2011, doi: 10.1016/j.enpol.2011.06.041.

[11] N. Fumo and M. A. Rafe Biswas, "Regression analysis for prediction of residential energy consumption," *Renewable and Sustainable Energy Reviews*, vol. 47. Elsevier Ltd, pp. 332–343, 2015. doi: 10.1016/j.rser.2015.03.035.

[12] D. B. Crawley, J. W. Hand, M. Kummert, and B. T. Griffith, "Contrasting the capabilities of building energy performance simulation programs," *Building and Environment*, vol. 43, no. 4, pp. 661–673, Apr. 2008, doi: 10.1016/j.buildenv.2006.10.027.

[13] M. T. Ke, C. H. Yeh, and J. T. Jian, "Analysis of building energy consumption parameters and energy savings measurement and verification by applying eQUEST software," *Energy and Buildings*, vol. 61, pp. 100–107, 2013, doi: 10.1016/j.enbuild.2013.02.012.

[14] Nyserda, “Toward a Clean Energy Future: A Strategic Outlook 2021-2024.”

[15] “CLCPA-Fact-Sheet”.

[16] “Ecobee Smart Thermostat,” *https://www.ecobee.com/en-us/*.

[17] “NYISO Load Data,” *https://www.nyiso.com/load-data*.

[18] “PLUTO Data,” *https://www1.nyc.gov/site/planning/data-maps/open-data/dwn-pluto-mappluto.page*.

[19] “EIA Monthly Consumption Data,” *https://www.eia.gov/totalenergy/data/monthly/previous.php*.

[20] “Propane Consumption Data,” *https://www.eia.gov/state/seds/archive/seds2017.pdf*.

[21] “Distillate Oil Consumption Data,” *https://www.eia.gov/dnav/pet/pet_cons_821dst_dcu_SNY_a.htm*.

[22] “NSRDB Weather Data Website,” *https://nsrdb.nrel.gov/*.

[23] “Energy Tips for Commercial Buildings,” *https://doee.dc.gov/service/energy-tips-commercial-buildings*.

[24] “NYSERDA Commercial Baseline Study,” *https://www.nyserda.ny.gov/About/Publications/Building-Stock-and-Potential-Studies/Commercial-Statewide-Baseline-Study*.

[25] “American Community Survey Data,” *https://data.census.gov/cedsci/*.

[26] T. Hoyt, E. Arens, and H. Zhang, “Extending air temperature setpoints: Simulated energy savings and design considerations for new and retrofit buildings,” *Building and Environment*, vol. 88, pp. 89–96, Jun. 2015, doi: 10.1016/j.buildenv.2014.09.010.

[27] “EnergyPlus and SEEM Modeling Enhancements via Software-to-Software Comparison Using NREL’s BEopt Test Suite,” *https://www.nrel.gov/docs/fy16osti/65858.pdf*.

[28] “Heating and Cooling (Temperature Set Point) Policy.” [Online]. Available: www.plu.edu/facilities-management

[29] “Ramping Up Heat Pump Adoption in New York State Targets and Programs to Accelerate Savings Prepared for NRDC by VEIC,” 2018. [Online]. Available: https://energyplan.ny.gov/-/media/nysenergyplan/final/greenhouse-gas-inventory.pdf.

[30] Nyserda, “New Efficiency: New York - Analysis of Residential Heat Pump Potential and Economics,” 2019.

[31] “New York ISO Climate Change Impact Study Phase 1: Long-Term Load Impact,” 2019.

[32] M. Z. Jacobson, “On the correlation between building heat demand and wind energy supply and how it helps to avoid blackouts,” *Smart Energy*, vol. 1, p. 100009, Feb. 2021, doi: 10.1016/j.segy.2021.100009.

[33] L. J. French, M. J. Camilleri, N. P. Isaacs, and A. R. Pollard, “Temperatures and heating energy in New Zealand houses from a nationally representative study-HEEP,” *Energy and Buildings*, vol. 39, no. 7, pp. 770–782, Jul. 2007, doi: 10.1016/j.enbuild.2007.02.004.

[34] M. Aksoezen, M. Daniel, U. Hassler, and N. Kohler, “Building age as an indicator for energy consumption,” *Energy and Buildings*, vol. 87, pp. 74–86, Jan. 2015, doi: 10.1016/j.enbuild.2014.10.074.

[35] T. Catalina, V. Iordache, and B. Caracaleanu, “Multiple regression model for fast prediction of the heating energy demand,” *Energy and Buildings*, vol. 57, pp. 302–312, 2013, doi: 10.1016/j.enbuild.2012.11.010.

[36] I. Gaetani, P. J. Hoes, and J. L. M. Hensen, “Estimating the influence of occupant behavior on building heating and cooling energy in one simulation run,” *Applied Energy*, vol. 223, pp. 159–171, Aug. 2018, doi: 10.1016/j.apenergy.2018.03.108.

[37] Y. S. Lee and A. M. Malkawi, “Simulating multiple occupant behaviors in buildings: An agent-based modeling approach,” *Energy and Buildings*, vol. 69, pp. 407–416, Feb. 2014, doi: 10.1016/j.enbuild.2013.11.020.

[38] K. U. Ahn and C. S. Park, “Different occupant modeling approaches for building energy prediction,” in *Energy Procedia*, Jun. 2016, vol. 88, pp. 721–724. doi: 10.1016/j.egypro.2016.06.050.

[39] E. J. Hutchinson, P. Wilkinson, S. H. Hong, and T. Oreszczyn, “Can we improve the identification of cold homes for targeted home energy-efficiency improvements?,” *Applied Energy*, vol. 83, no. 11, pp. 1198–1209, 2006, doi: 10.1016/j.apenergy.2006.01.007.

[40] M. Berger and J. Worlitschek, “A novel approach for estimating residential space heating demand,” *Energy*, vol. 159, pp. 294–301, Sep. 2018, doi: 10.1016/j.energy.2018.06.138.

[41] J. Suhonen, J. Jokisalo, R. Kosonen, V. Kauppi, Y. Ju, and P. Janßen, “Demand Response Control of Space Heating in Three Different Building Types in Finland and Germany,” *Energies*, vol. 13, no. 23, p. 6296, Nov. 2020, doi: 10.3390/en13236296.

[42] C. Carpino, R. Bruno, and N. Arcuri, “Statistical analysis of the heating demand in residential buildings located in Mediterranean climate and proposals for refurbishment,” in *Energy Procedia*, 2017, vol. 133, pp. 16–27. doi: 10.1016/j.egypro.2017.09.365.

[43] R. Haas, H. Auer, and P. Biermayr, “The impact of consumer behavior on residential energy demand for space heating,” 1998.

[44] O. Ruhnau, L. Hirth, and A. Praktiknjo, “Time series of heat demand and heat pump efficiency for energy system modeling,” *Scientific Data*, vol. 6, no. 1, Dec. 2019, doi: 10.1038/s41597-019-0199-y.

[45] I. Y. Choi, S. H. Cho, and J. T. Kim, “Energy consumption characteristics of high-rise apartment buildings according to building shape and mixed-use development,” in *Energy and Buildings*, Mar. 2012, vol. 46, pp. 123–131. doi: 10.1016/j.enbuild.2011.10.038.

[46] M. de Rosa, V. Bianco, F. Scarpa, and L. A. Tagliafico, “Heating and cooling building energy demand evaluation; A simplified model and a modified degree days approach,” *Applied Energy*, vol. 128, pp. 217–229, Sep. 2014, doi: 10.1016/j.apenergy.2014.04.067.

[47] T. A. Little, “Wind Induced Energy Demand in Cold Climates,” 1993.

[48] G. Papaefthymiou, B. Hasche, and C. Nabe, “Potential of heat pumps for demand side management and wind power integration in the German electricity market,” *IEEE Transactions on Sustainable Energy*, vol. 3, no. 4, pp. 636–642, 2012, doi: 10.1109/TSTE.2012.2202132.