2021 NYC

Tenant Energy Data Challenge

PROPTECH CHALLENGE

NEW YORK STATE OF OPPORTUNITY | NYSERDA

REBNY REAL ESTATE BOARD OF NEW YORK

Table of Contents

1

What is your forecasted consumption across all 18 tenant usage meters for the 24 hours of 8/31/20 in 15 minute intervals (1728 predictions)?

Provided Data Sources

Tenant Usage

Time series data of demand and consumption of electricity from one office tenant spanning from 1/1/18 12:00 AM - 8/30/20 12:00 AM where readings are taken at the end of 15-minute intervals.

- meter - unique meter ID
- date_time - date and time of reading
- consumption - electricity used since the last measurement (in kWh)
- max_demand - highest electrical power reading in the last 15 minutes (in kW)
- min_demand - lowest electrical power reading in the last 15 minutes (in kW)
- avg_demand - average electrical power reading over the last 15 minutes (in kW)

Occupancy

The number of unique entries to the building for a given day.

ConEd Electric

Time series data of building-wide (totalizer) demand and consumption of electricity spanning from 1/1/18 12:00 AM - 9/1/20 12:00 AM where readings are taken at the end of 15-minute intervals.

- meter - unique meter ID
- date_time - date and time of reading
- consumption - electricity used since the last measurement (in kWh)
- max_demand - highest electrical power reading in the last 15 minutes (in kW)
- min_demand - lowest electrical power reading in the last 15 minutes (in kW)
- avg_demand - average electrical power reading over the last 15 minutes (in kW)

Time series data of external temperature and humidity readings taken on the hour from the building's rooftop weather station.

- temp - hourly readings of external temperature (in Fahrenheit)
- humidity - hourly readings of external humidity (in %)

ConEd Steam

1 - Time series data of building-wide demand and consumption of steam spanning from 1/1/18 12:00 AM - 9/1/20 12:00 AM where readings are taken at the end of 15-minute intervals.

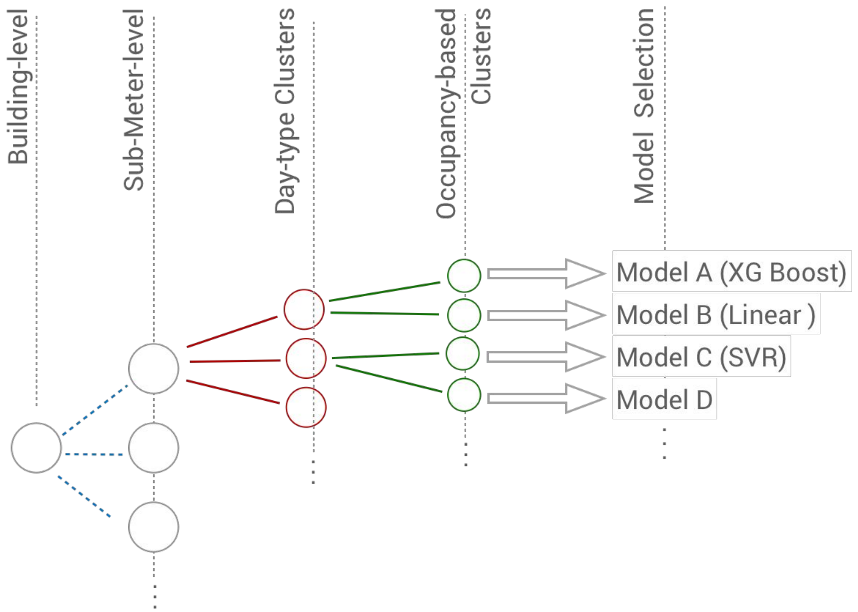
- date_time - date and time of reading
- consumption - steam used since the last measurement (in lbs)
- max_demand - highest steam reading in the last 15 minutes (in lbs/hour)
- min_demand - lowest steam reading in the last 15 minutes (in lbs/hour)
- avg_demand - average steam reading over the last 15 minutes (in lbs/hour)

2 - Time series data of external temperature and humidity readings taken on the hour from the building's rooftop weather station

- temp - hourly readings of external temperature (in Fahrenheit)
- humidity - hourly readings of external humidity (in %)

Consumption Forecasts Hierarchical Modeling Framework

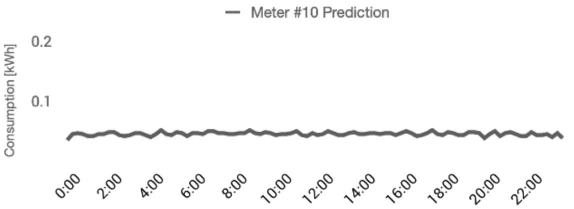
Here, a framework that combines Clustering and Model selection algorithms is developed to construct ML-based models to predict tenant-level electricity demand for different phases observed by commercial buildings (i.e. before the pandemic, during the shutdown, and after reopening) in highest fidelity.

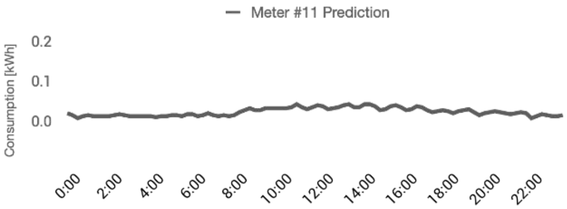


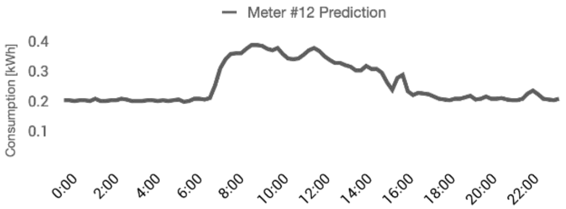
Hierarchical Modeling Framework

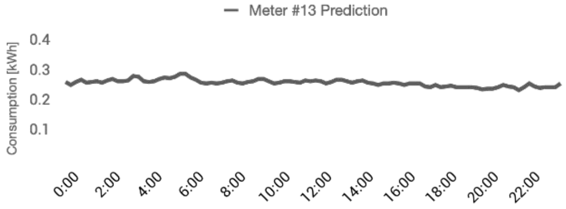
Consumption Forecasts (kWh): Results Submeter 1-9

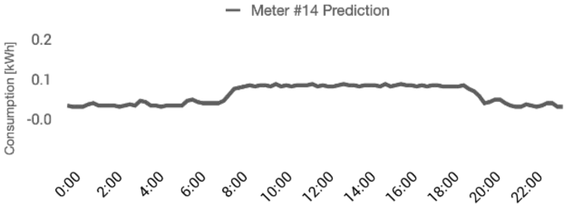
Consumption Forecasts (kWh) : Results Submeter 10-18

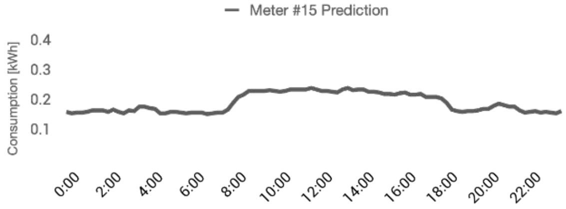


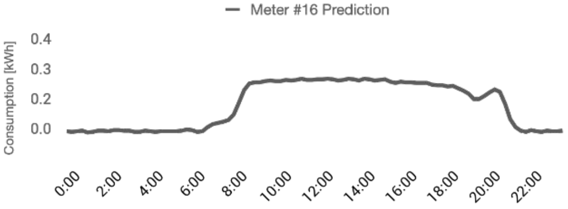


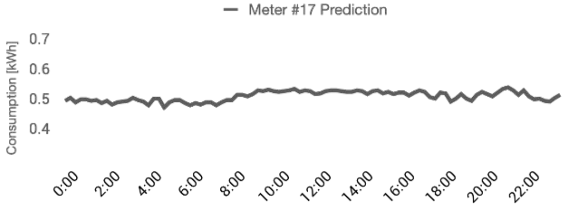


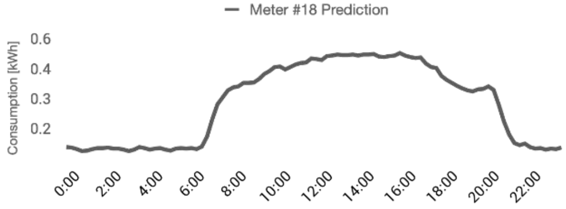




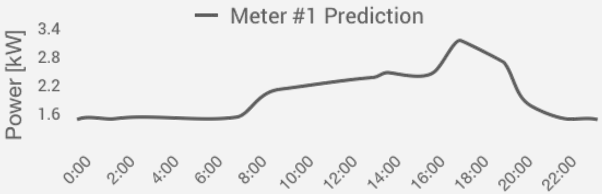


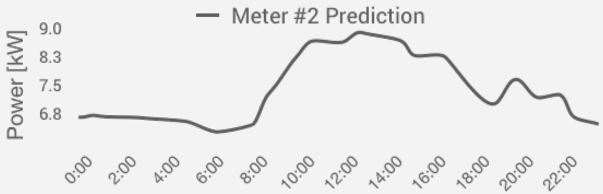


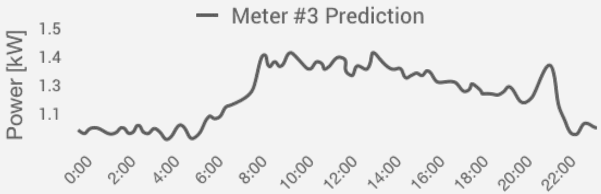


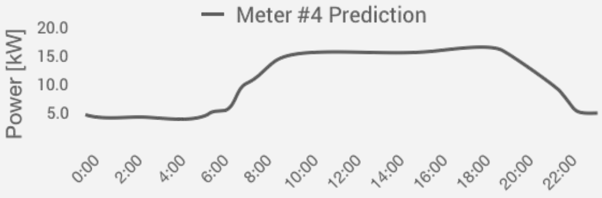


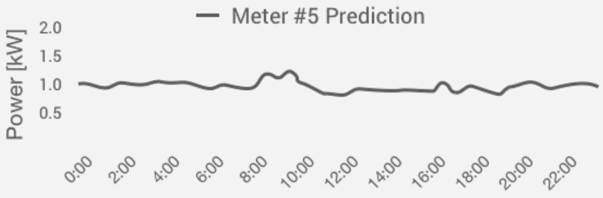
Consumption Forecasts (kW): Results Submeter 1-9

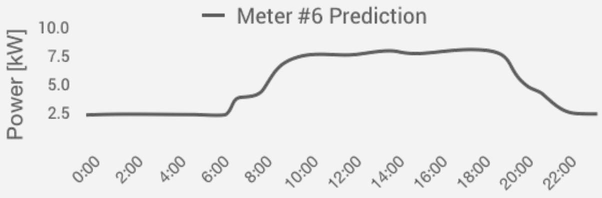


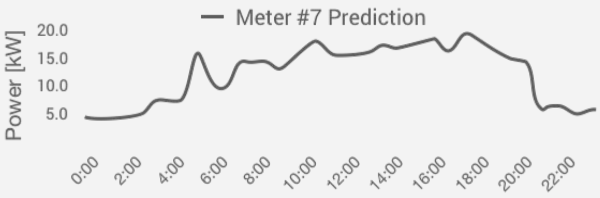


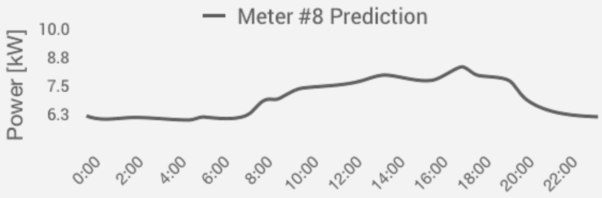


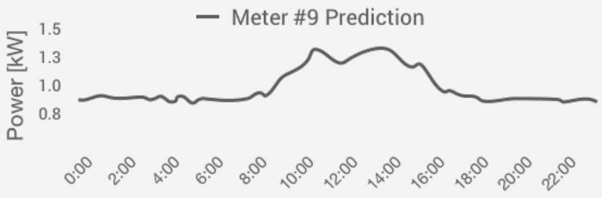




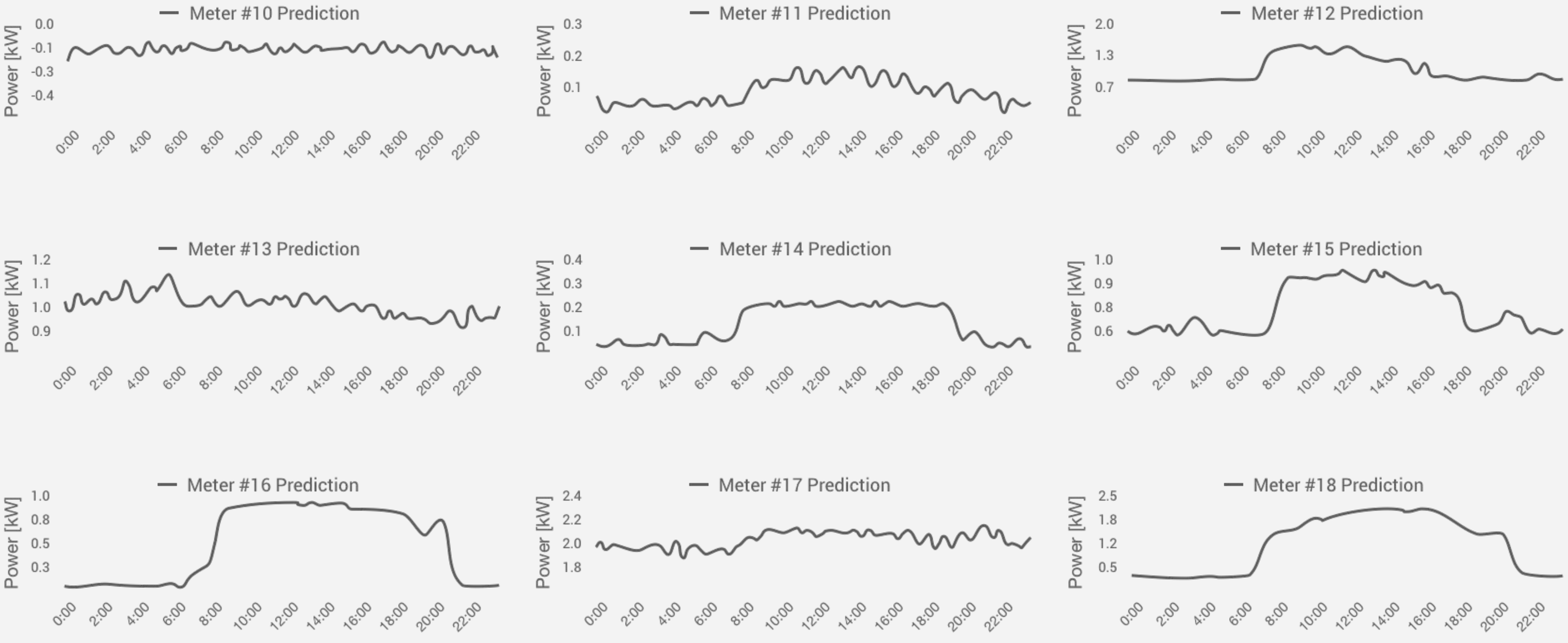








Consumption Forecasts (kW): Results Submeter 10-18



2

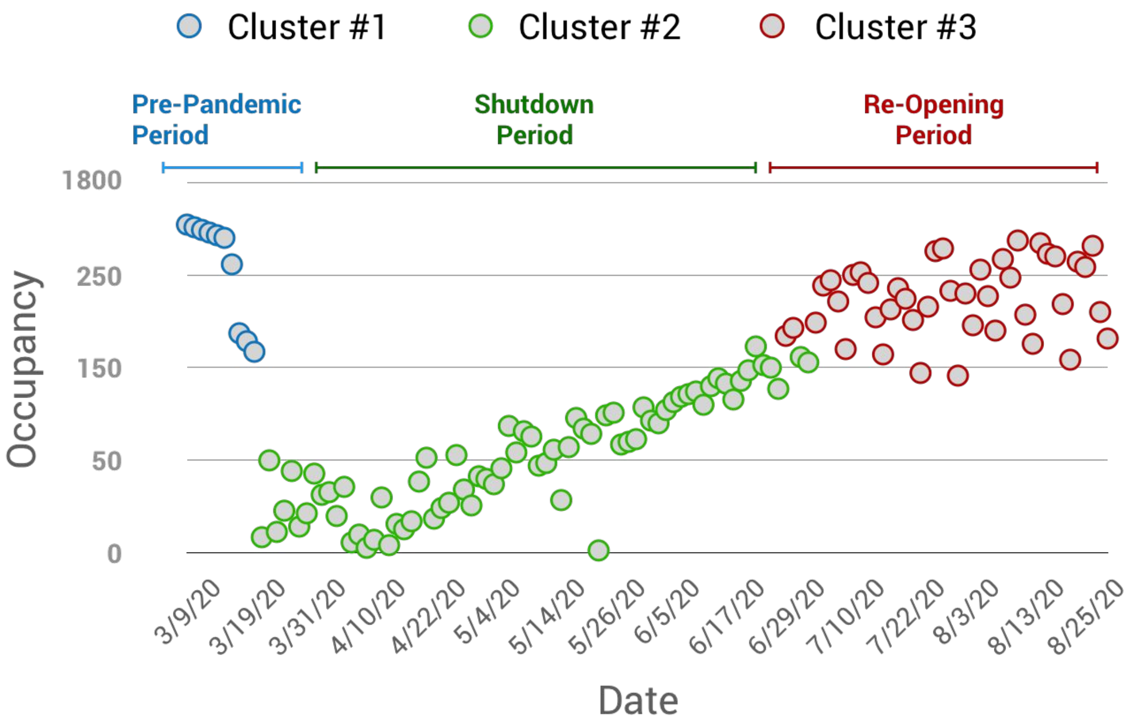
How correlated are building-wide occupancy and tenant consumption?

Occupancy-Based Clustering

To understand the correlations between building occupancy and tend consumption, occupancy data is clustered based on occupancy level and date using expectation maximization algorithm.

As illustrated in the Figure here, the occupancy data is clustered into three clusters. Interestingly, the clustered occupancy follows the real-world behavior and represents the three phases that are already observed in New York including:

(i) pre-pandemic period
(ii) shutdown period
(iii) reopening period



Submeter-Level Heat-Map Analysis

To investigate the impact of occupancy-based clustering on hourly tenant energy consumption (tenant sub meter data), heatmap analysis is performed on each submeter data as illustrated below.

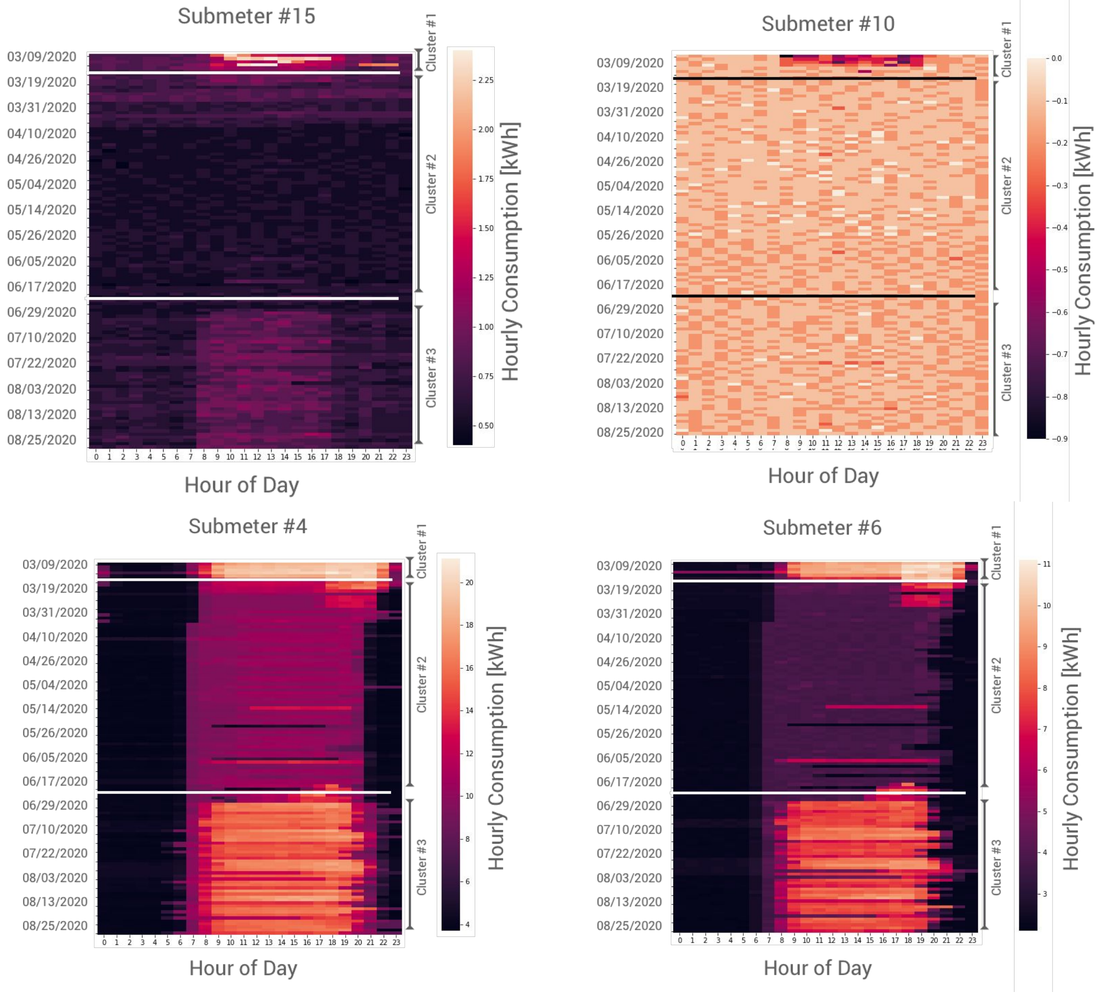
From heatmap analysis it can be clearly observed that

(i) there is a strong correlation between daily tenant-level energy consumption trend during operational hours and the occupancy-based clusters in floors associated with submeters No. 2, 4, 6, 14, and 18,

(ii) there is a clear correlation between tenant consumption trend and occupancy-based clusters in floors associated with submeter No. 1, 3, 7, 8, 11, 12, and 15, and

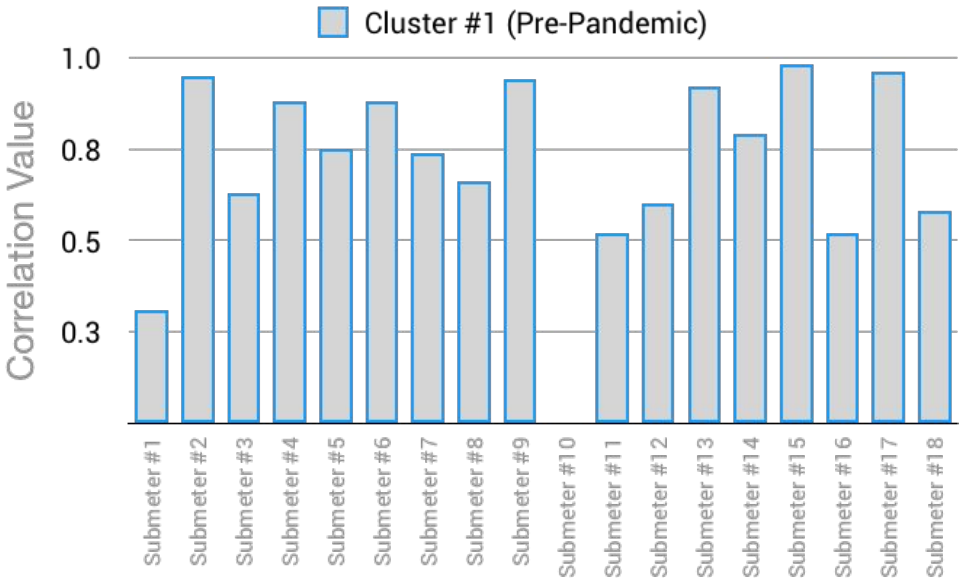
(iii) there is no significant correlation between tenant consumption trend and occupancy-based clusters in floors associated with submeters No. 5, 9, 10, 13, 16, and 17.

Again, occupancy-based clusters represent the three phases that the city experienced during the c-19 pandemic including (i) pre-pandemic phase, (ii) shutdown period, and (iii) reopening phases in New York.



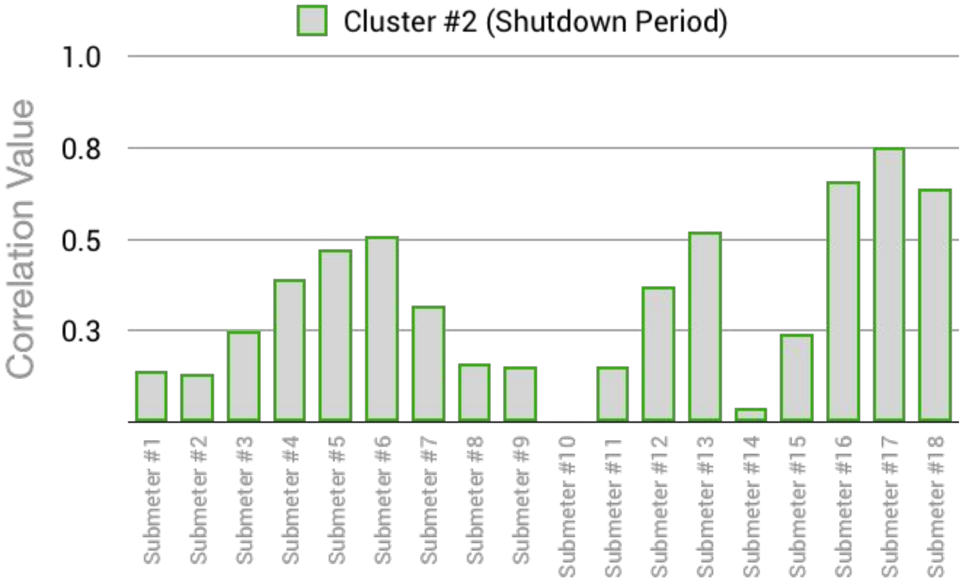
Correlation Analysis - Pre-Pandemic

In this section, the correlation analysis is performed to understand the relationship between tenant energy consumption (submeter data) and building-wide occupancy during the three different phases in 2020 .



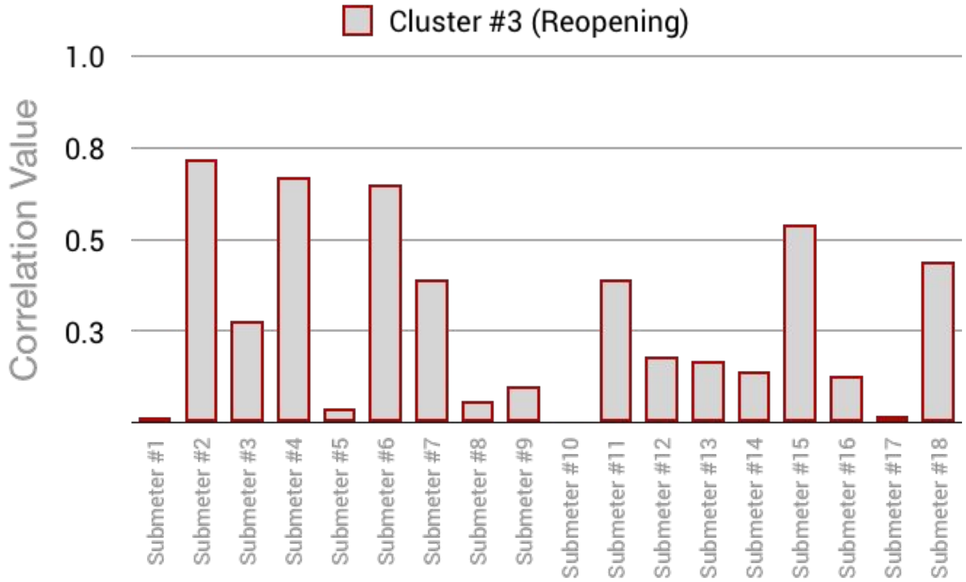
Correlation Analysis - Shutdown

In this section, the correlation analysis is performed to understand the relationship between tenant energy consumption (submeter data) and building-wide occupancy during the three different phases in 2020.



Correlation Analysis - Reopening

In this section, the correlation analysis is performed to understand the relationship between tenant energy consumption (submeter data) and building-wide occupancy during the three different phases in 2020.



Correlation Conclusions

From heatmap and correlation analysis provided in section 1-3 in different occupancy-based clusters, it can be concluded:

Occupancy has the strongest correlation with electricity before pandemic and after reopening. Almost no / weak correlation during the shutdown period

Note:
Strong: Correlation higher than 0.75
Clear: Correlation between 0.5 and 0.75
Weak: Correlation less than 0.5

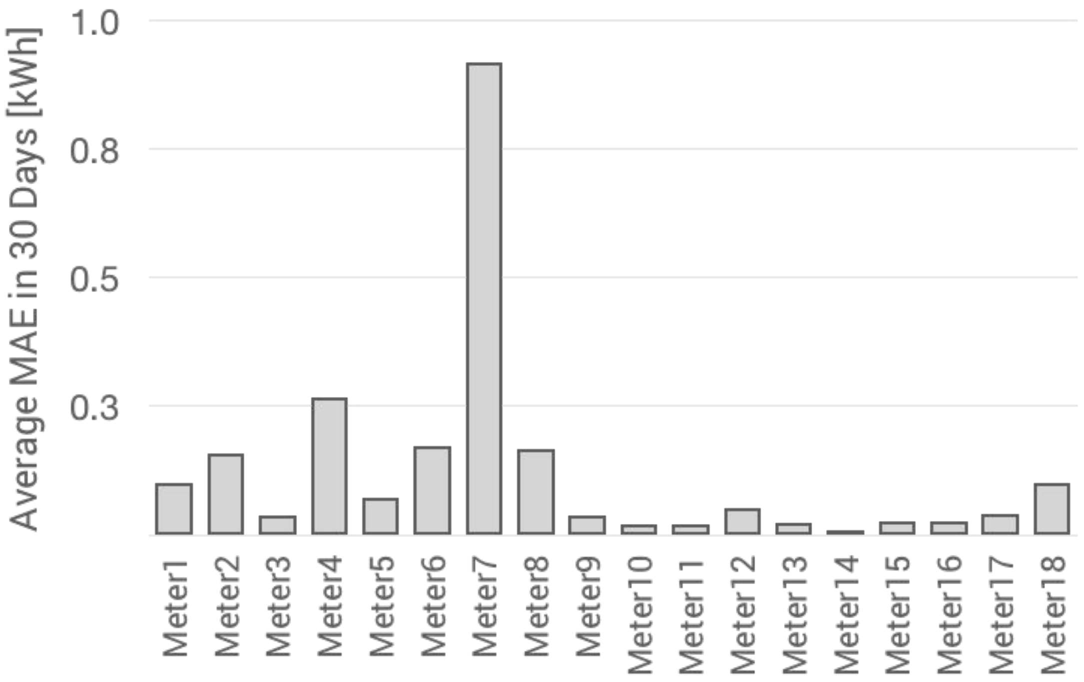
	Correlation between pandemic phases (occupancy-based clustering) and tenant energy consumption trend associated with this meter	**Pre-Pandemic:** Tenant energy consumption and building-wide occupancy in Cluster No. 1	**Shutdown:** Tenant energy consumption and building-wide occupancy in Cluster No. 2	**Reopening:** Tenant energy consumption and building-wide occupancy in Cluster No. 3
Submeter 1	Clear	Weak	Weak	Weak
Submeter 2	Strong	Strong	Weak	Strong
Submeter 3	Clear	Clear	Weak	Weak
Submeter 4	Strong	Strong	Weak	Strong
Submeter 5	Weak	Clear	Weak	Weak
Submeter 6	Strong	Strong	Weak	Clear
Submeter 7	Clear	Clear	Weak	Weak
Submeter 8	Clear	Clear	Weak	Weak
Submeter 9	Weak	Strong	Weak	Weak
Submeter 10	Weak	Weak	Weak	Weak
Submeter 11	Clear	Weak	Weak	Weak
Submeter 12	Clear	Clear	Weak	Weak
Submeter 13	Weak	Strong	Weak	Weak
Submeter 14	Strong	Strong	Weak	Weak
Submeter 15	Clear	Strong	Weak	Clear
Submeter 16	Weak	Clear	Clear	Weak
Submeter 17	Weak	Strong	Clear	Weak
Submeter 18	Strong	Clear	Clear	Clear

3

What is the mean absolute error for your model?

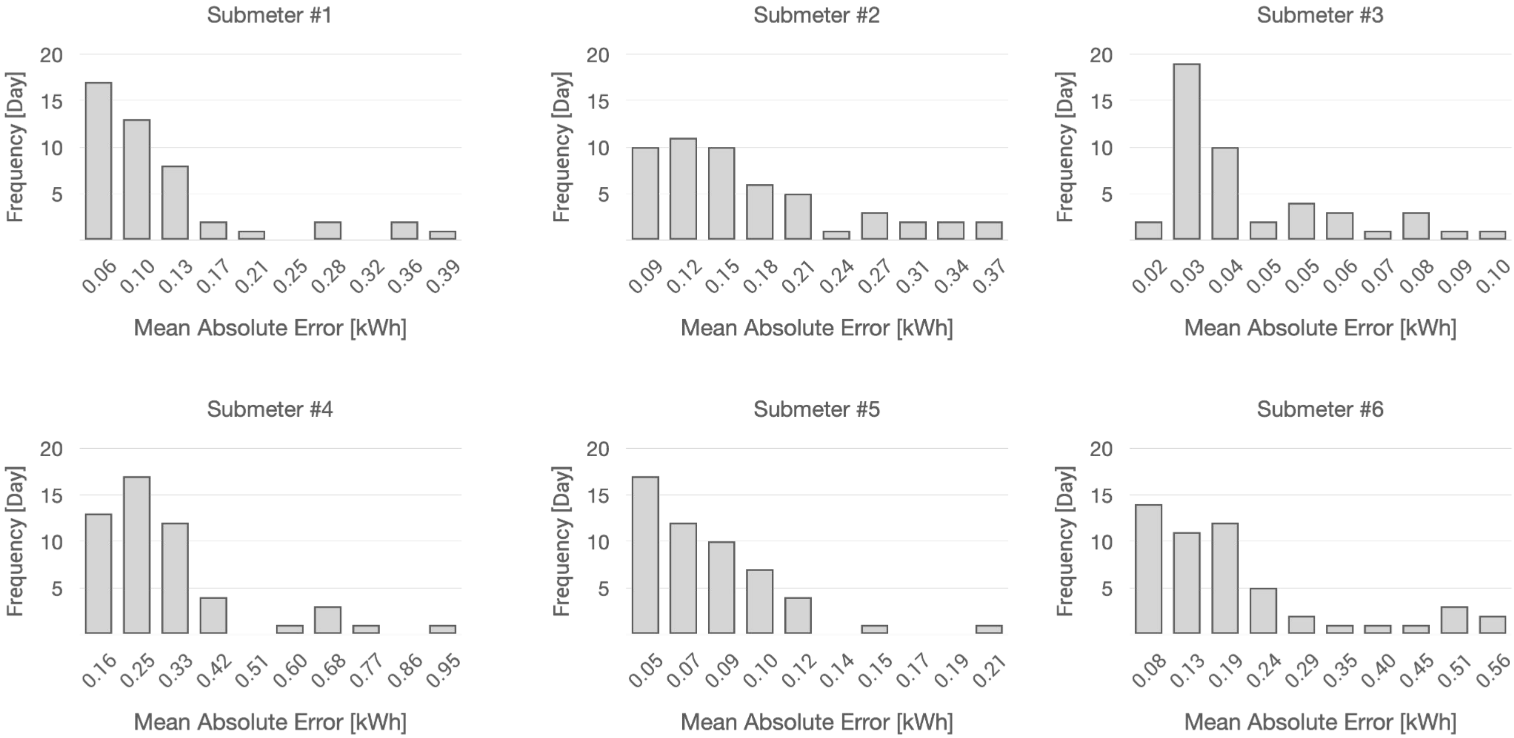
Mean Absolute Error (kWh)

Figure to the right illustrates the average of the MAE in different submeter models; it is observed that the MAE varies between ~0.009 kWh (in the model constructed over submeter #14) to ~0.9 kWh (in the model constructed over submeter #7).



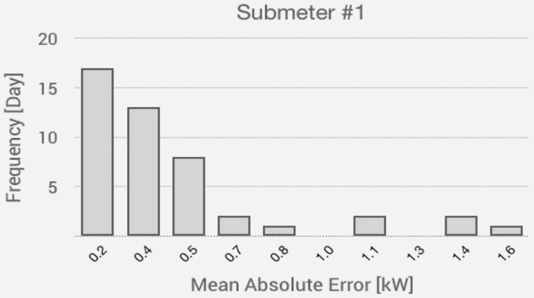
Mean Absolute Error (kWh)

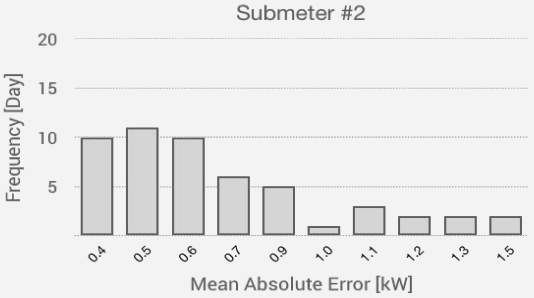
Here the daily mean absolute error (MAE) is calculated by comparing the predicted behavior of submeter models (18 models) and the observed data from these submeters in the past 30 days of data.

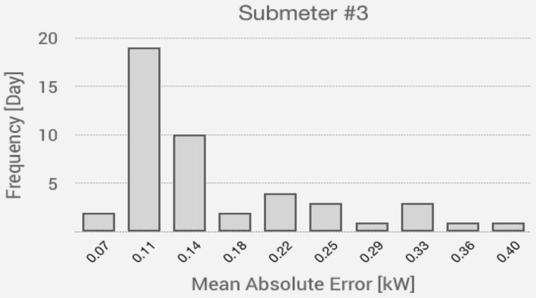


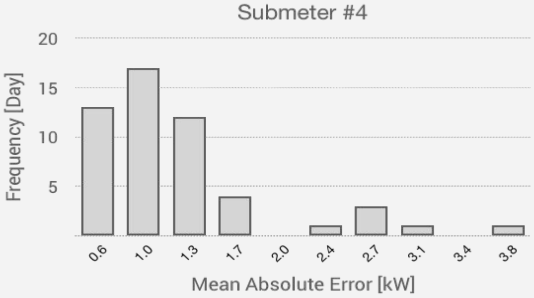
Mean Absolute Error (kW)

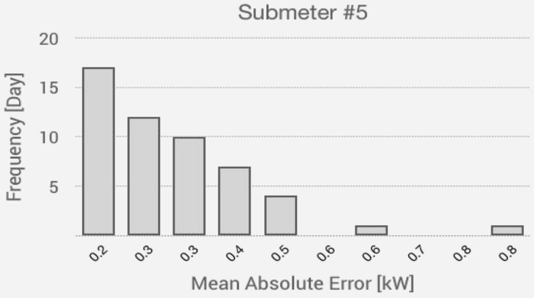
Here the daily mean absolute error (MAE) is calculated by comparing the predicted behavior of submeter models (18 models) and the observed data from these submeters in the past 30 days of data.

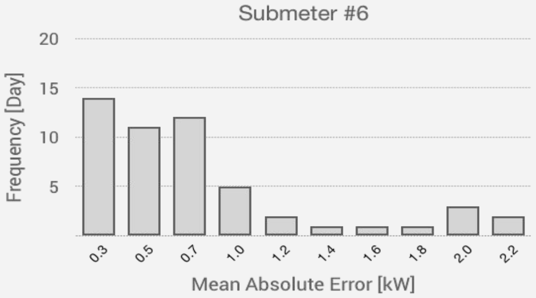






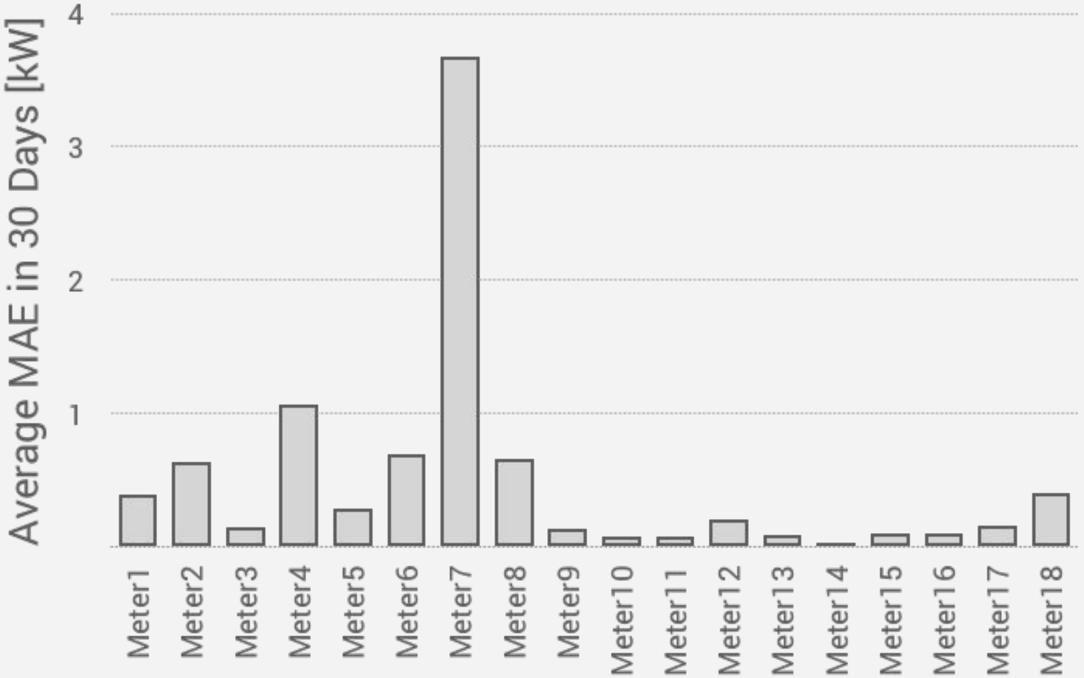






Mean Absolute Error (kW)

Figure to the right illustrates the average of the MAE in different submeter models; it is observed that the MAE varies between 0.03kW (in the model constructed over submeter #14) to 3.6kW (in the model constructed over submeter #7).

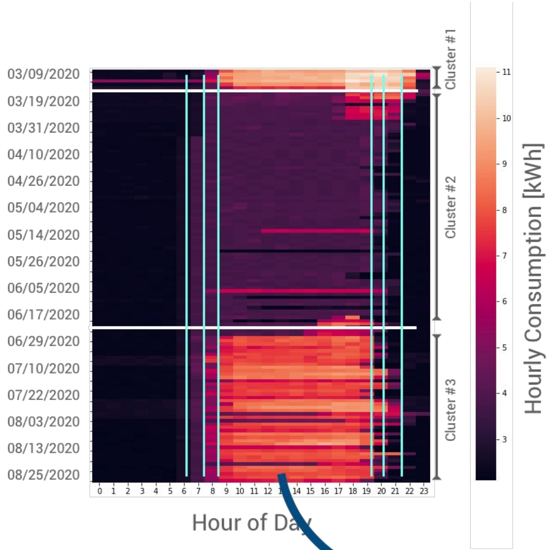


4

What features / predictors were most important in determining energy efficiency?

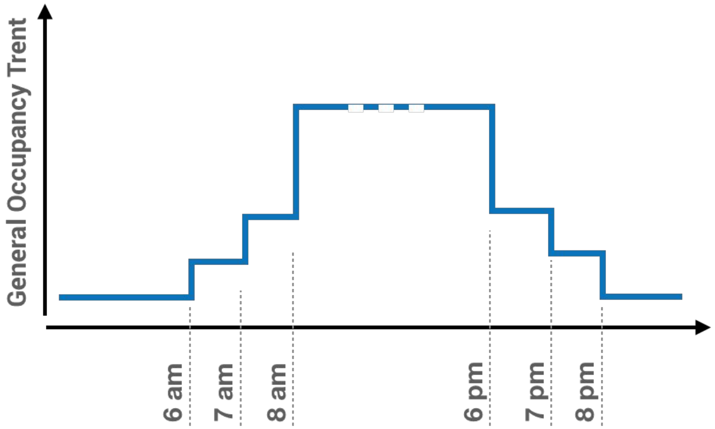
The Power Of Occupancy-Based ECMs

Building occupancy patterns facilitate the successful implementation of advanced real-time Energy Conservative Measures (ECM) in buildings. To investigate the potential impact of occupancy-based ECMs on the provided building, here, occupancy detection is conducted through the use of floor-level submeter data.



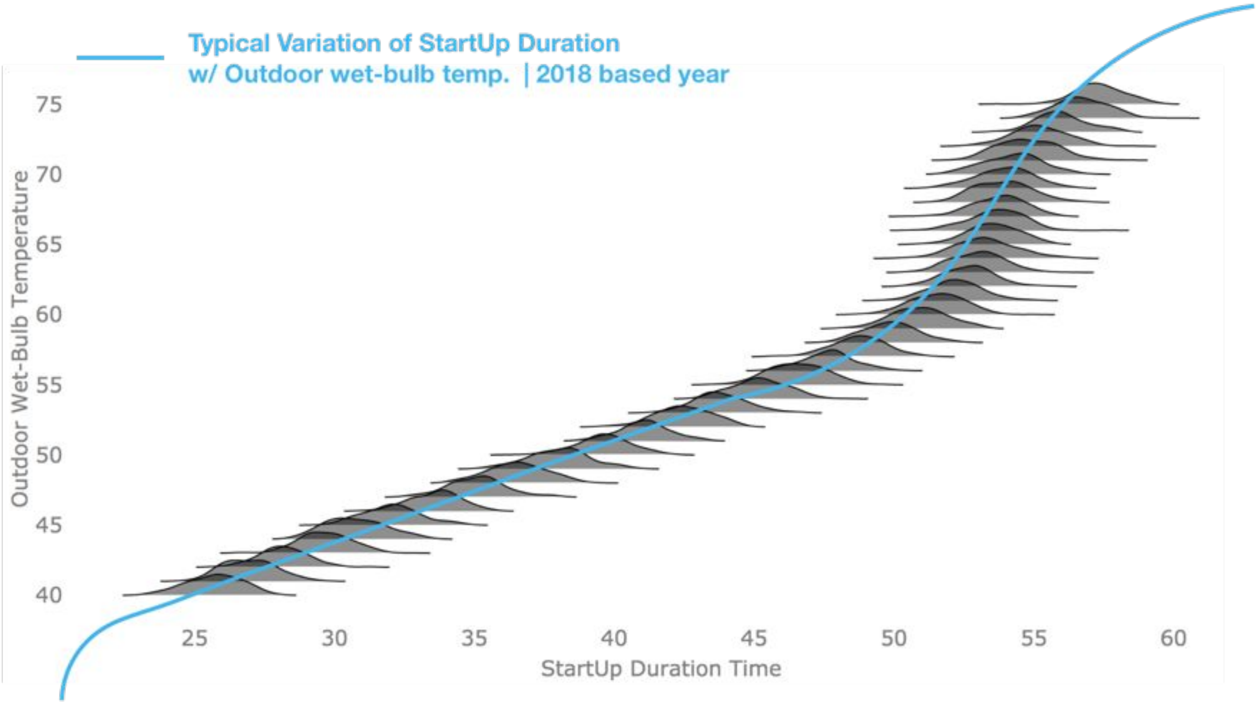
Occupancy trend detection from submeter data

To this end, the heatmap analysis is used to find the submeters that are highly correlated with tenant occupancy:



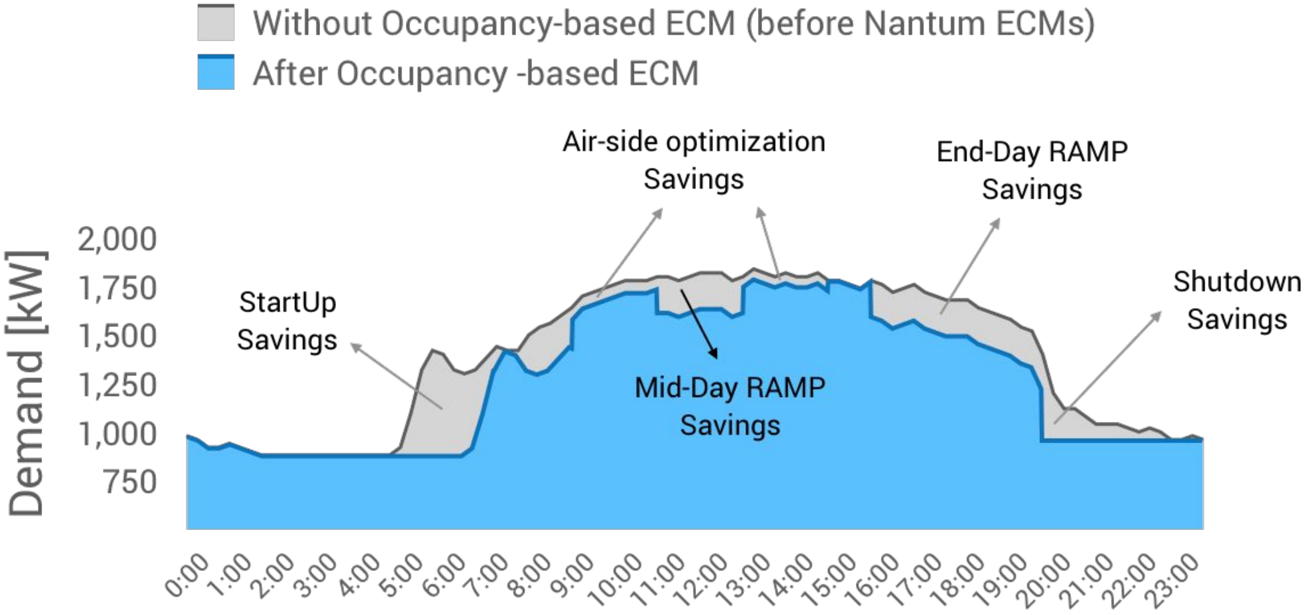
The Power Of Occupancy-Based ECMs

Next, the occupancy-level, the daily building level energy consumption and the climate location of the building is used to find the similar buildings in Nantum DB. The selected similar buildings are then used to approximate the specifications needed to simulate the behavior of the buildings in different outdoor conditions as illustrated in the figure here:



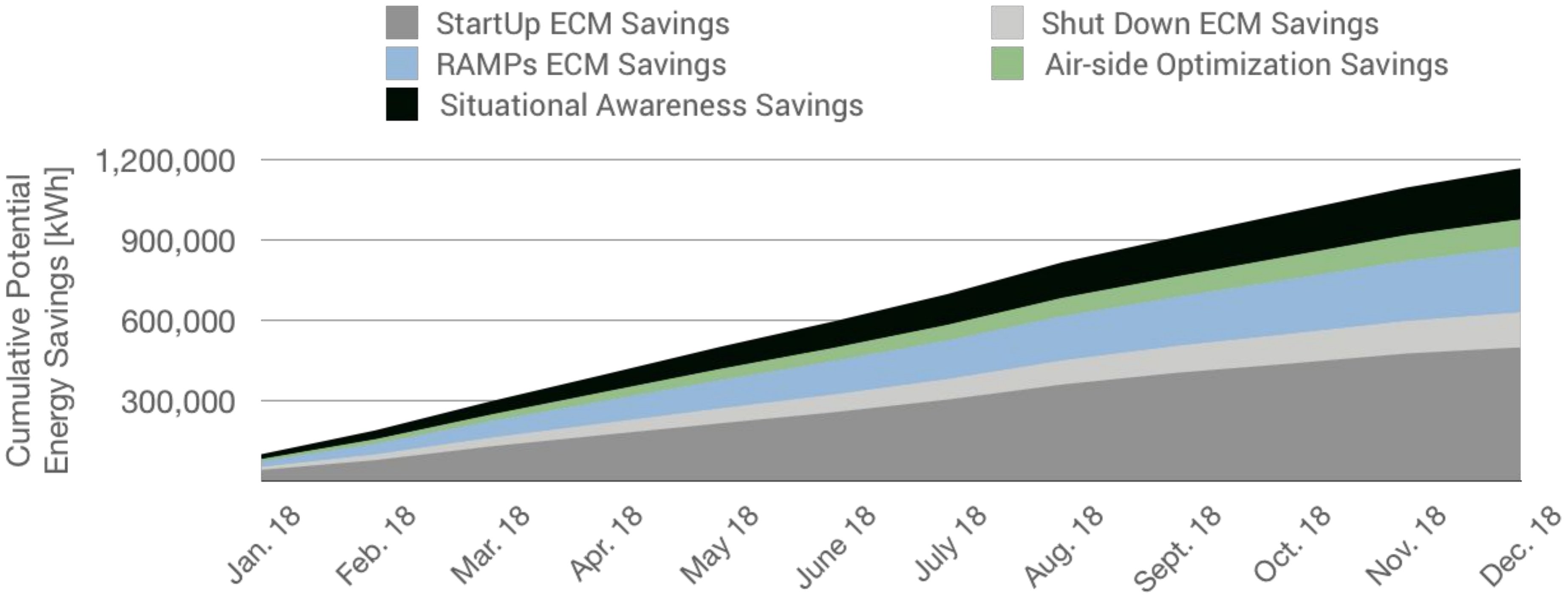
The Power Of Occupancy-Based ECMs

The detected occupancy trend along with the models constructed from similar buildings are then used to approximate the impact of occupancy-based ECMs already developed in Nantum including startup, shutdown, air-side optimization, mid-day ramp, and end of day ramp. Figure to the right illustrated the potential impact of Nantum Occupancy-based ECMs on 2019-08-16:



The Power Of Occupancy-Based ECMs

It is observed that using daily occupancy trends (on a building or floor-level), the building can save about 10% energy on average (between 9% to 15% daily energy savings) by implementing start-up, shut-down, and RAMP ECMs.

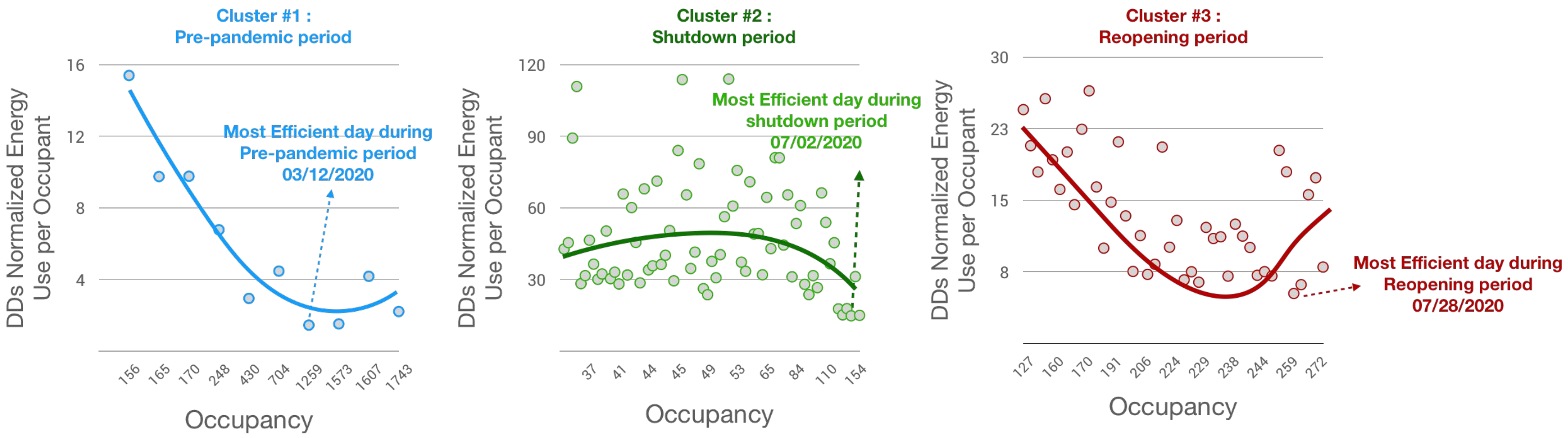


5

What is the most energy-efficient occupancy level as a percentage of max occupancy provided?

Ideal Occupancy vs. Electricity

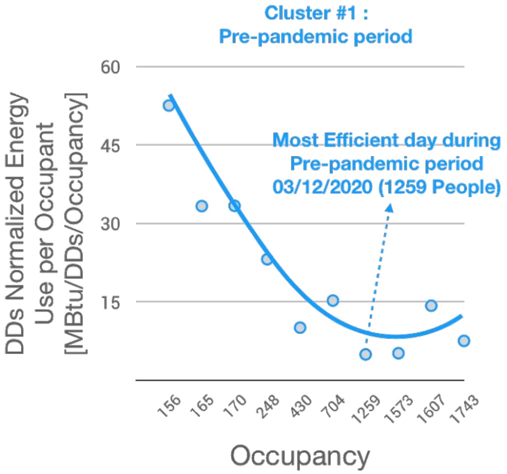
To understand the most efficient days in different occupancy-based clusters, the variation of Normalized Daily Electricity Consumption (Degree Days (DDs) normalized daily electricity consumption) per occupant with building-wide occupancy is modeled. It can be clearly observed that the building has the highest efficiency (electricity use efficiency) during the pre-pandemic on **03/12/2020** with **1,259 people** inside the building and **33,600 kWh** energy consumption.

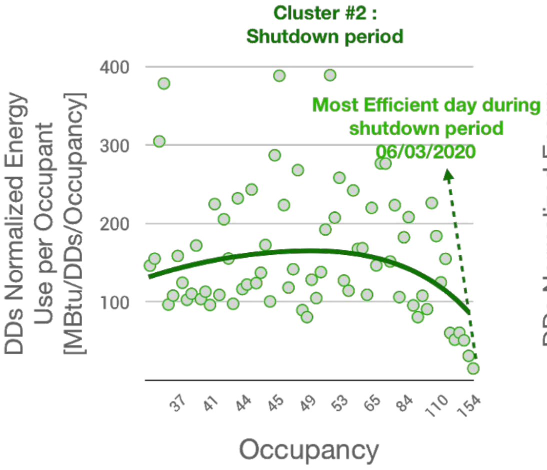


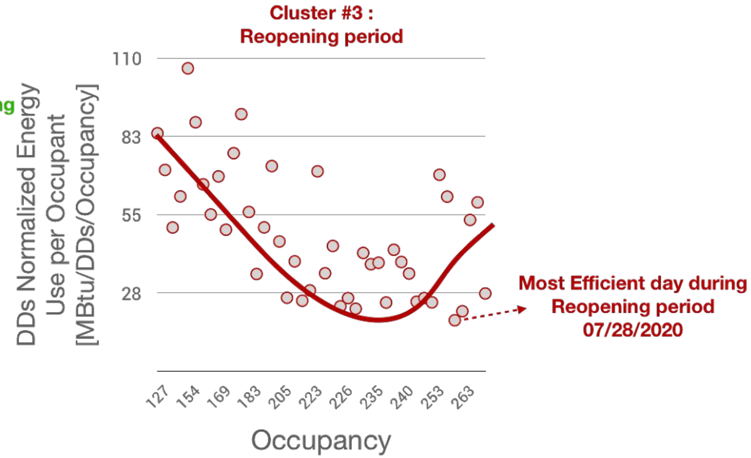
Ideal Occupancy vs. (Electricity + Steam)

Below the variation of Normalized Daily Energy Consumption (Degree Days (DDs) normalized daily electricity and steam consumption in MBtu) per occupant with building-wide occupancy is also investigated. Again, It can be clearly observed that the building has the highest efficiency during the pre-pandemic on 03/12/2020 with 1,259 people. This number is different in other clusters.

In the normal operation of the building, the energy efficiency of the building is maximized when the occupancy is between 1200 to 1500.







- Normal Building Operations
- Normal Occupancy
- High Energy Efficiency

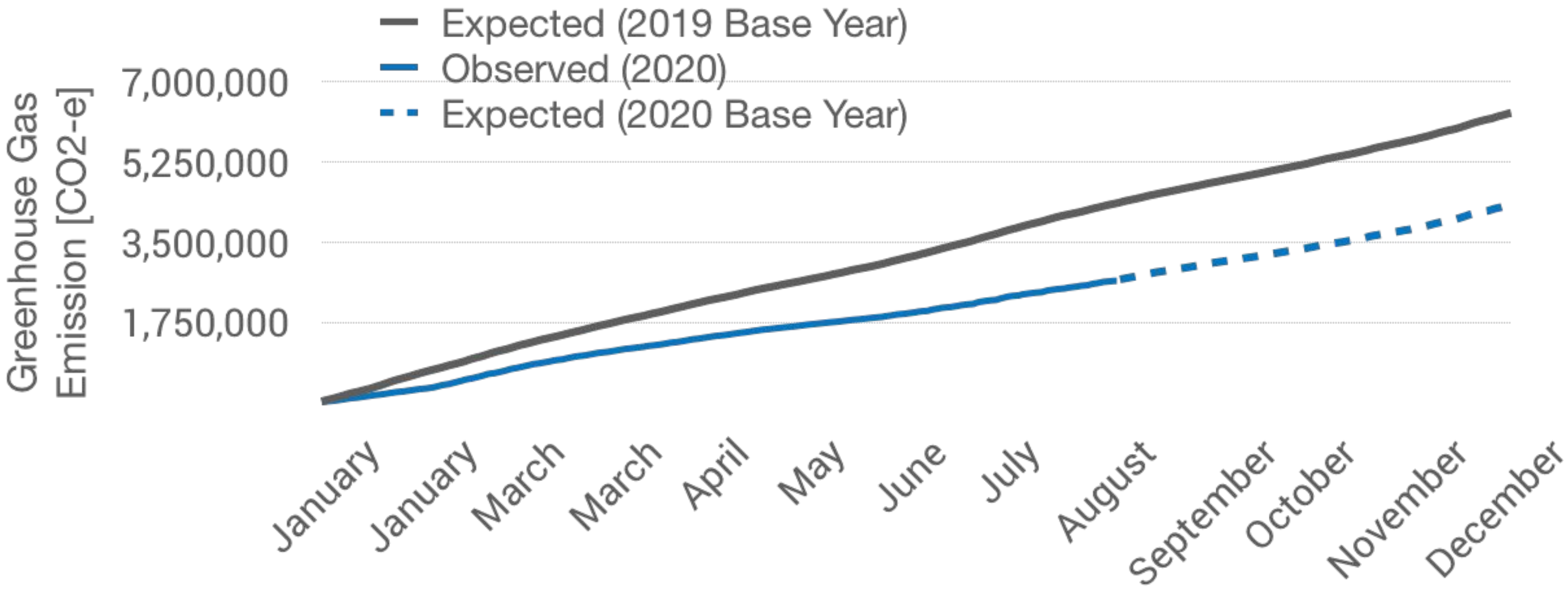
- Lease Obligation Based Operations
- Low Occupancy
- Low Energy Efficiency

- Close-to-Normal Operations
- Low Occupancy
- Medium Energy Efficiency

6

What else, if anything, can be concluded or constructed from your model?

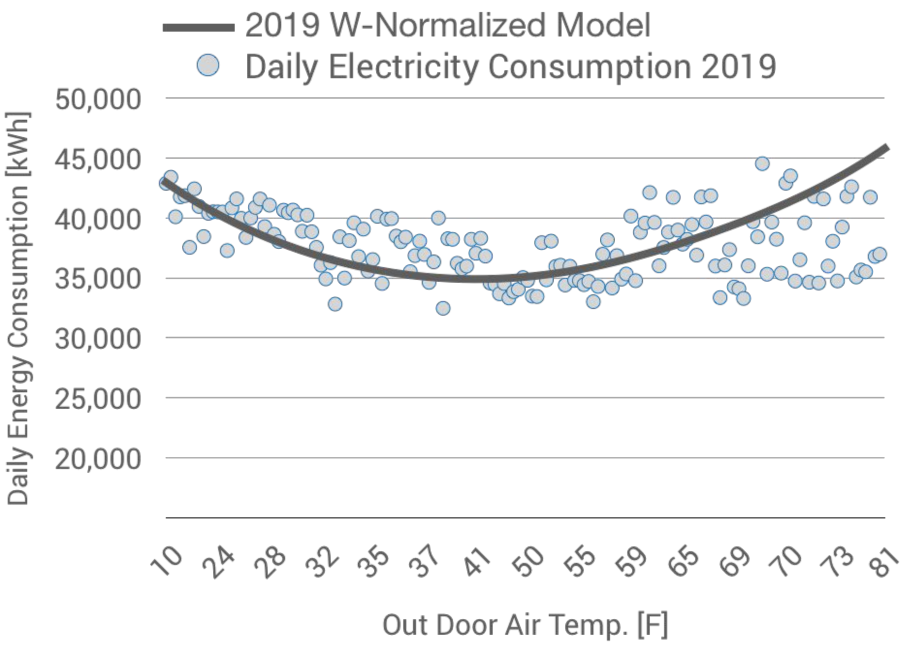
LL97: 2020 Greenhouse Gas Emission

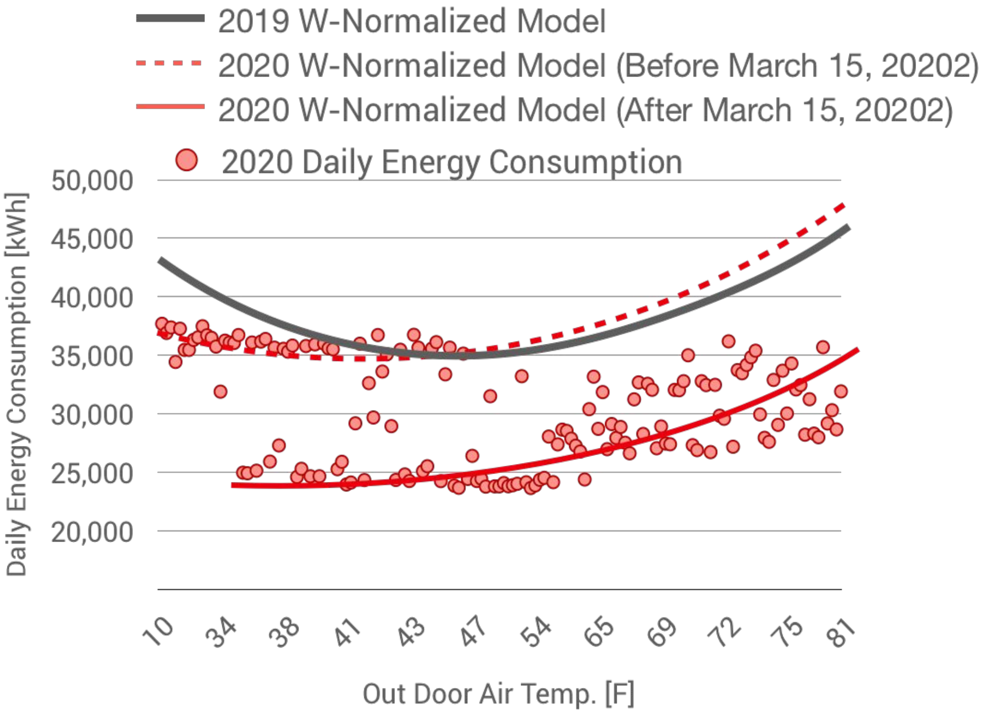


Without COVID impact, the expected GHG emission for 2020 based on historical data is the red line. Because of the impact of COVID, observed 2020 emission until August and expected emission until December should drop about 32.02%. Further, if this trend continues, there is potential for reaching and surpassing goals set for LL97 as well as avoiding penalties.

Weather Normalized Model

While 2019' energy consumption can be modeled by one single weather normalized model, two separate weather normalized models are needed for representing 2020's energy consumption. One model(dashed red line) closely represents the normal pattern which resembles 2019 model (grey line). Another model(solid red line) represents energy consumption affected by pandemic where daily kWh is significantly reduced.

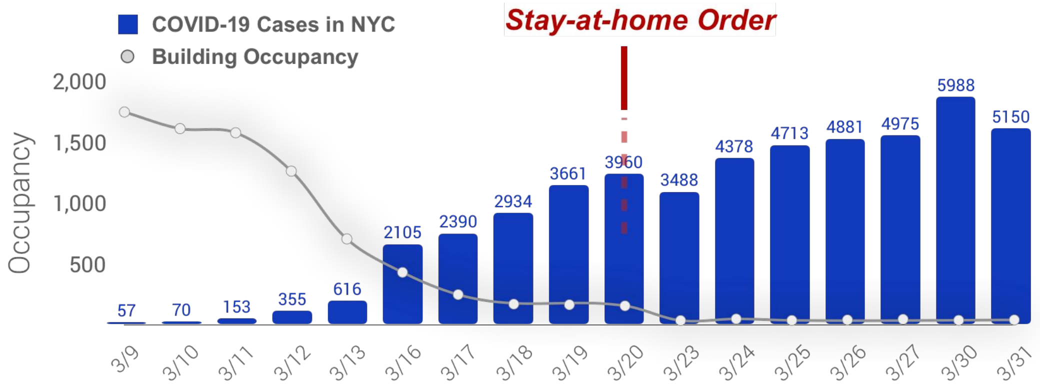




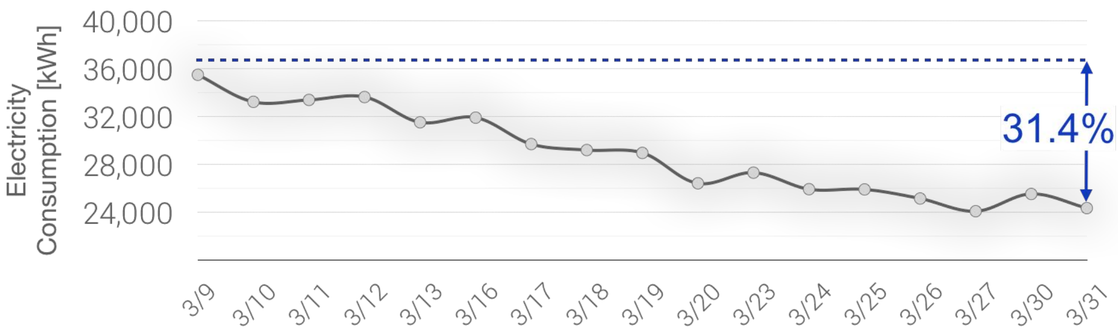
Shutdown Period & Stay-at-home order

During shutdown period, the magnitude of occupancy reduction is much more significant than energy reduction. This is mainly because buildings have to continue operations for property safety and tenant Service Level Agreements (SLA) regardless of occupancy level. It can be understood that 70% of electricity used by building is essentially independent of occupancy. It purely depends on building envelope and weather to keep it at certain service level.

- NYC Office Building Occupancy and COVID-19 Cases

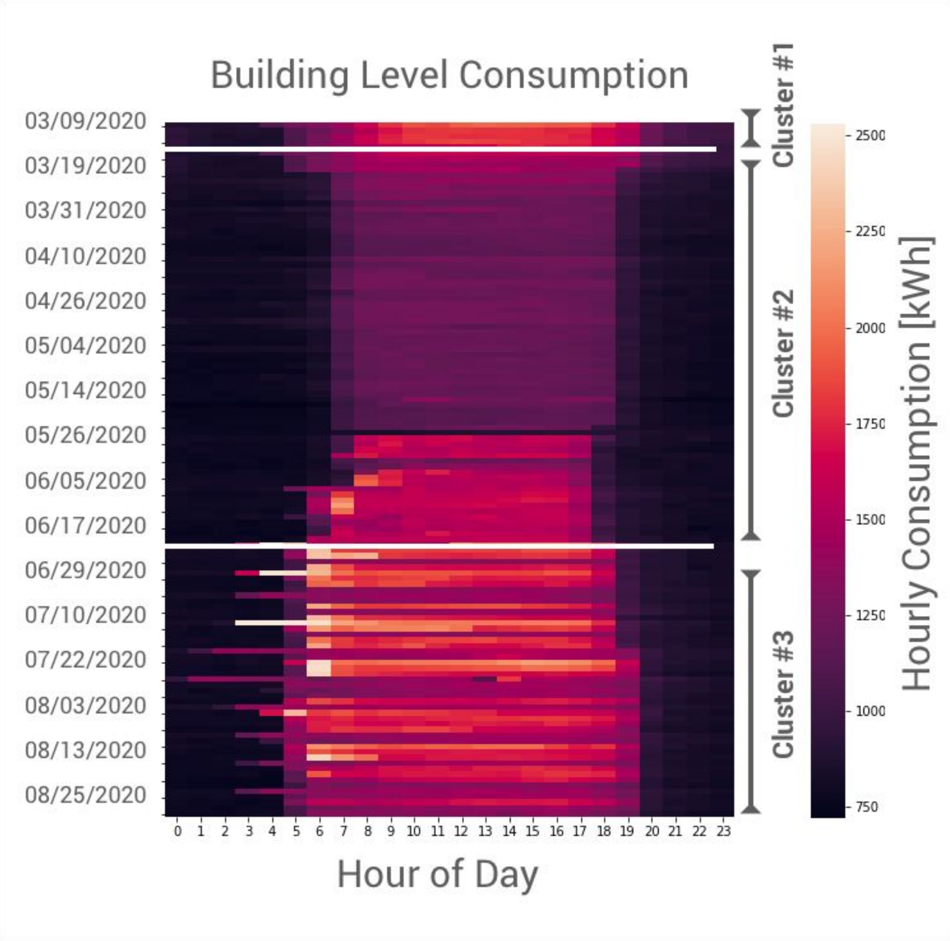


- Impact of COVID-19 in Building Electricity Consumption



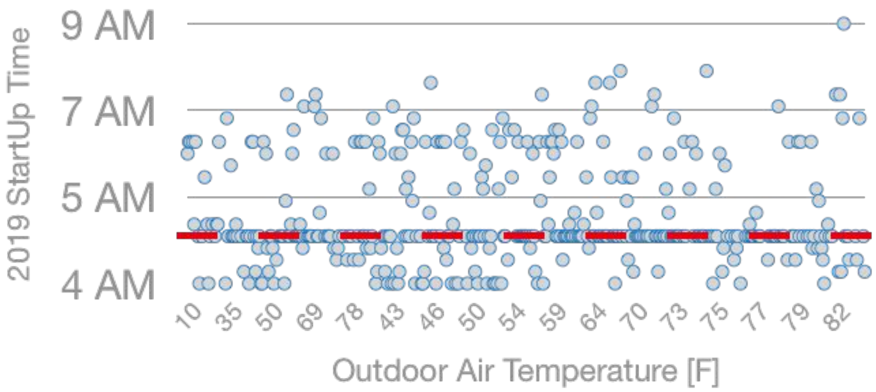
Heatmap Analysis on Building Level Electricity Consumption

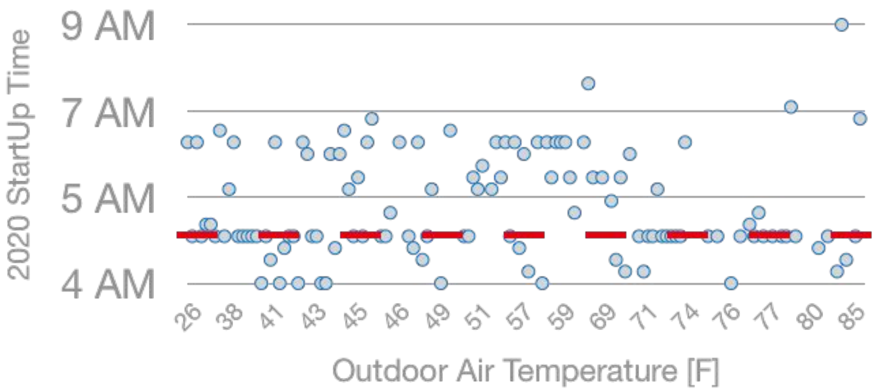
There is a clear separation of three clusters which represents the pre-COVID period, shutdown period and reopening period by observing the heat map of building level consumption as well. At the same time, the operational hours of each day and the differences between the three clusters can be clearly observed. One key observation that further supports the previous finding that building are still using a significant amount of energy because of SLA during shutdown period.



Variation of StartUp w Outdoor Air Temp.

From building electricity data, the startup times of the building can be interpolated and plotted as following comparing against the weather conditions. It can be observed that the startups mostly happen around 4:30 am and is independent of weather and occupancy conditions. For controlling the optimal startup times for buildings considering the interior space conditions, weather predictions, and occupancy trends, Nantum can be extremely helpful and reliable to provide these scheduling recommendations and automate these controls to improve energy efficiency and generate energy and monetary savings.





7

What other information, if any, would you need to better your model?

Additional Data

High Frequency Occupancy	• High frequency occupancy (i.e., 5-min) is critical for predicting energy consumption of a building. • High frequency occupancy is critical for real-time energy-efficiency control systems.
Tenant level occupancy	• High frequency tenant level occupancy is critical for predicting energy consumption of the submeter data. • High frequency tenant level occupancy is critical for real-time zone-level energy-efficiency control systems. This leads to more savings from the ECMs.
Operational Hours	• Daily operational hours data can provide us a new dimension that can be used to improve the fidelity of the prediction models.
Lease SLA times and obligation on operational hours	• Lease SLA times is the time that the human-thermal comfort and indoor air quality should be in acceptable band based on ASHRAE 55. • Having SLA times is significantly important to perform schedule optimization on building operation and maximize the energy consumption of buildings.
Building Square Footage	• The size of the building is an important factor in calculating the energy efficiency metric. • Considering the LL97, the square footage of the building can be used to calculate the 2024 and 2030 Greenhouse gas emission limits [in CO2-e], and the possible fines associated with them.
Disaggregated electricity consumption	• Disaggregated electricity data or equipment-level electricity consumption is a key for detecting anomaly and performing fully automated Preventive and Predictive maintenance.
Indoor Air Quality	• High indoor air quality can and should be an objective for building control systems to optimize human comfort along with temperature and humidity while maximizing the energy efficiency of the building environment.
Building Manifests	• A building manifest includes performance specs and static data accessible on each piece of equipment. This information not only enables energy-efficiency applications, it is also essential for anomaly detection and prediction.