

**PATTERN WISE CLUSTERING AND ENERGY ANALYTICS ON CAMPUS BUILDINGS**

**NAME: JOSEPH PABLO VARUN (A0163323N)**

**KE29 FT**

**SUPERVISOR: TIAN JING**

**CONTENTS**

1. Company Introduction……………………………………………………………….3

2. How buildings affect the environment ………………………………………4

3. Clustering the buildings based on area………………………………………5

4. Time series analysis of buildings…………………………………………….….8

5. Time Series Prediction Models…………………………………………………13

6. Observations And Conclusions…………………………………………………14

7.References………………………………………………………………………………..15

**COMPANY INTRODUCTION**

The Department of Building (DOB) at NUS has been involved with pioneering work in the building science domain since the late 70s, in the fields of Energy and Sustainability.

A strategic initiative grant from NUS in Energy Efficiency that led to the creation of the Centre for Integrated Building Energy and Sustainability in the Tropics (CiBEST), consisting of two research clusters:

1. Building Energy Efficiency (BEE) Hub Research Cluster,

2. Sustainable Buildings Research Cluster

BEE Hub (Building Energy Efficiency Hub) aims at addressing energy efficiency enhancements to more buildings on campus. BEE Hub has installed BTU meters, smart sensors, occupancy detection strategies and 72 IEQ sensor kits to collect data. All the data are in CSV format.

The ground floor area of 100 buildings, Electricity consumption data of 88 buildings and the cooling load data of 72 buildings are used for the study.

The below explain the different types of sensors used and the nature of data:

**NATURE OF DATA**

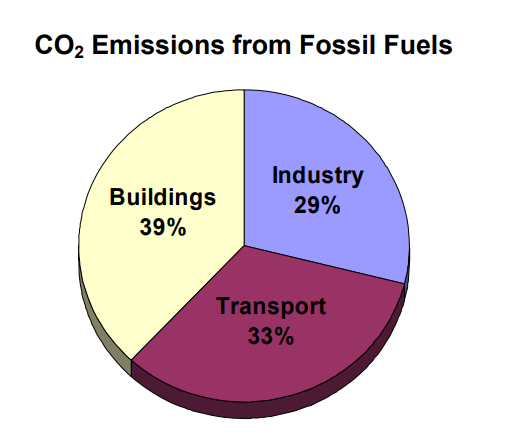
|  |  |  |
| --- | --- | --- |
| **SENSOR** | **PARAMETER** | **UNIT** |
| CO2 sensor | Measures CO2 concentration | ppm |
| Air Temperature | Measures the temperature of surrounding Air | Celcius or Fahrenheit |
| Humidity sensor | Moisture content in the air | grams of water vapor per cubic meter volume of air |
| Illuminance sensor | Measures the incident light concentration. | measuring luminous flux per unit area |

|  |  |  |  |
| --- | --- | --- | --- |
| **FEATURES** | **EXPLANATION** | **FREQUENCY** | **LOCATION** |
| PWM KWH(Power Watt meter) | Electricity measurement | 30 min | From 95 buildings around NUS |
| BTU Tonhr (British Thermal Unit) | Cooling Load measurement | 30 min | From 95 buildings around NUS |

**HOW BUILDINGS AFFECT THE ENVIRONMENT?**

* Most of the emissions come from heating, cooling and lighting to power appliances.
* Commercial and residential buildings are tremendous users of electricity, accounting for more than 70% of electricity use in the U.S.
* The building sector consumed 40 quadrillion BTUs of energy in 2005 at a cost of over $300 billion. Energy use in the sector is projected to increase to 50 quadrillion BTUs at a cost of $430 billion by the year 2025.
* The energy embodied in a single building’s envelope equals 8-10 times the annual energy used to heat and cool the building.
* If half of new commercial buildings were built to use 50% less energy, it would save over 6 million metric tons of CO2 annually for the life of the buildings the equivalent of taking more than 1 million cars off the road every year

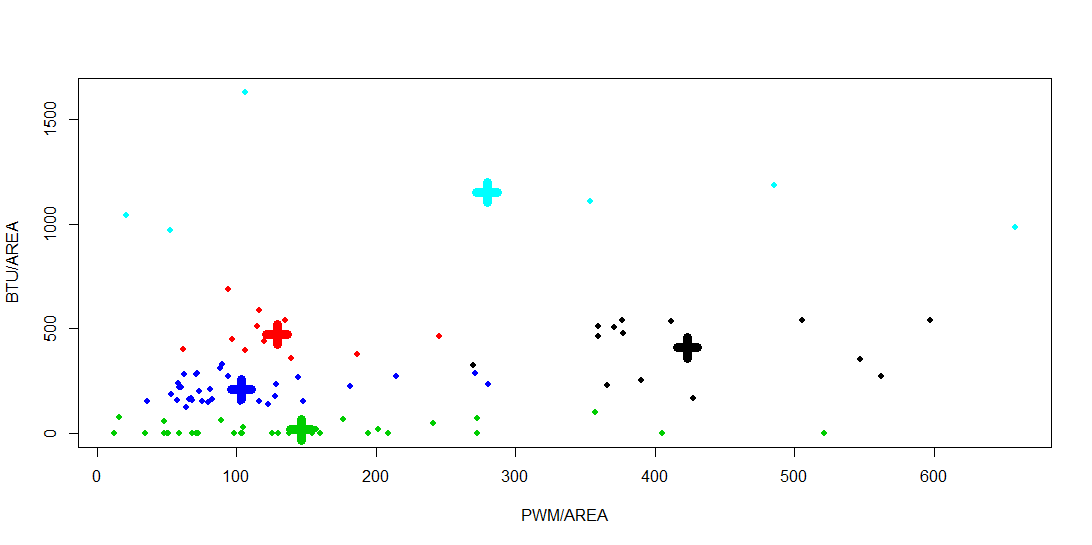
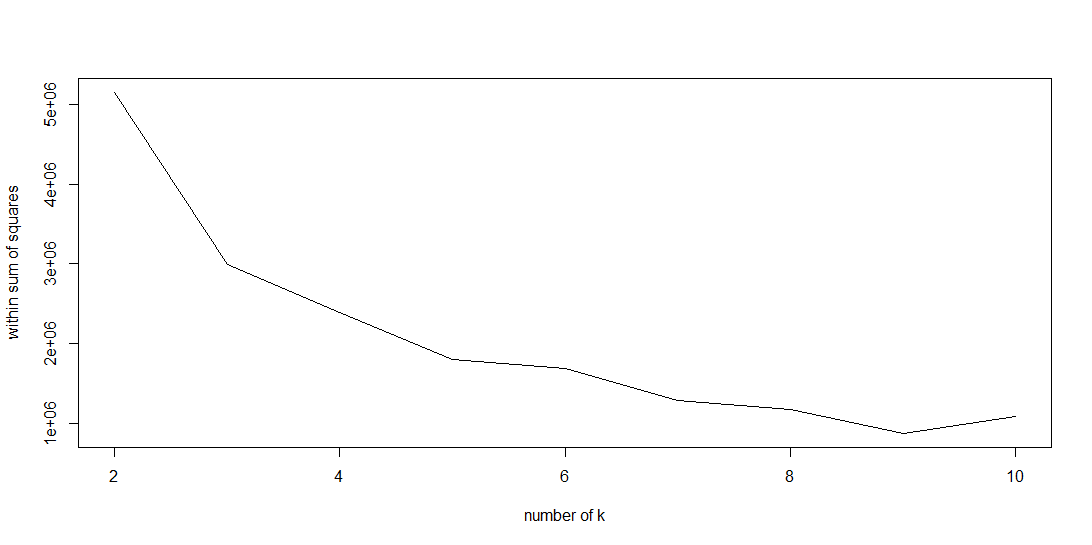
**COURTESY: ENVIRONMENTAL AND ENERGY STUDY INSTITUTE**



Hence ,energy optimisation of buildings is needed . By looking at the past trends we can analyse and model the future trend of the buildings. So the buildings are profiled into clusters and then analysed for planning and conserving the power.

**CLUSTERING THE BUILDINGS BASED ON THE AREA**

Initially the buildings are clustered according to their GFA(Ground floor area) according to their Electricity consumption and Cooling load consumption per unit area. The number of clusters are determined to be 5 using the below elbow plot. Then K-means clustering method has been used to cluster the buildings.



It is interesting to find that each cluster has diverse collection of buildings belonging to different faculties.

Every cluster has unique type of buildings belonging from many faculties and administration buildings that perform a unique function like office space,lecture hall,lab buildings etc.

The above plots show the average BTU and PWM consumption of each cluster.

**BUILDING COMPOSITION IN DIFFERENT CLUSTERS**

**FINDINGS:**

1. Cluster 1 has high BTU meeting halls.
2. Cluster 2 has all libraries and reading halls.
3. Cluster 3 has all the residences.
4. Cluster 4 mainly has administration buildings**.**

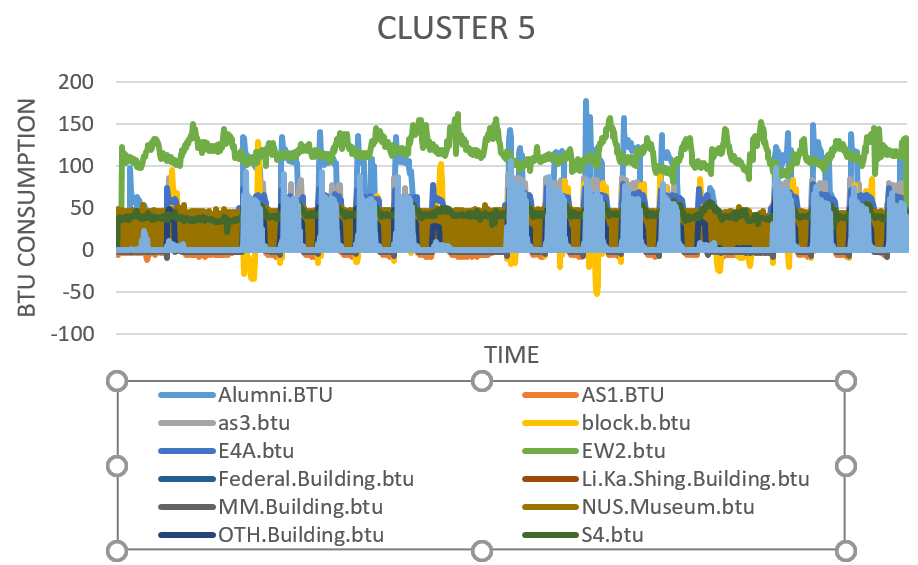
**TIME SERIES ANALYSIS OF BUILDINGS**

Using the time series data ,the BTU and PWM consumption over time ,the time series data is studied. The buildings are clustered based on their time series

pattern using K-shape clustering algorithm using DTW distance measure. The above plot shows the DTW based comparison of two buildings.

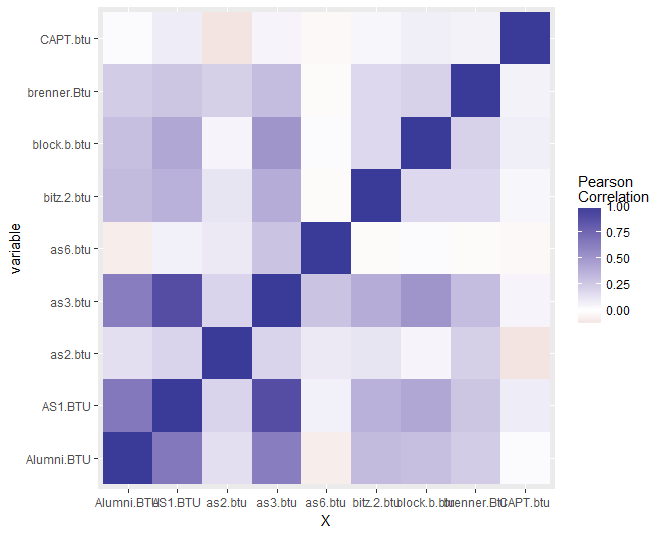
**TIME SERIES CLUSTERS OF BUILDINGS FOR AC LOAD**

The data points are split into weekdays and weekends. Office hours(9:00am-6:00pm) and non-office hours(06:00pm to 9:00am). The below data points show the AC load pattern during weekdays working hours for one semester(Jan-Apr).

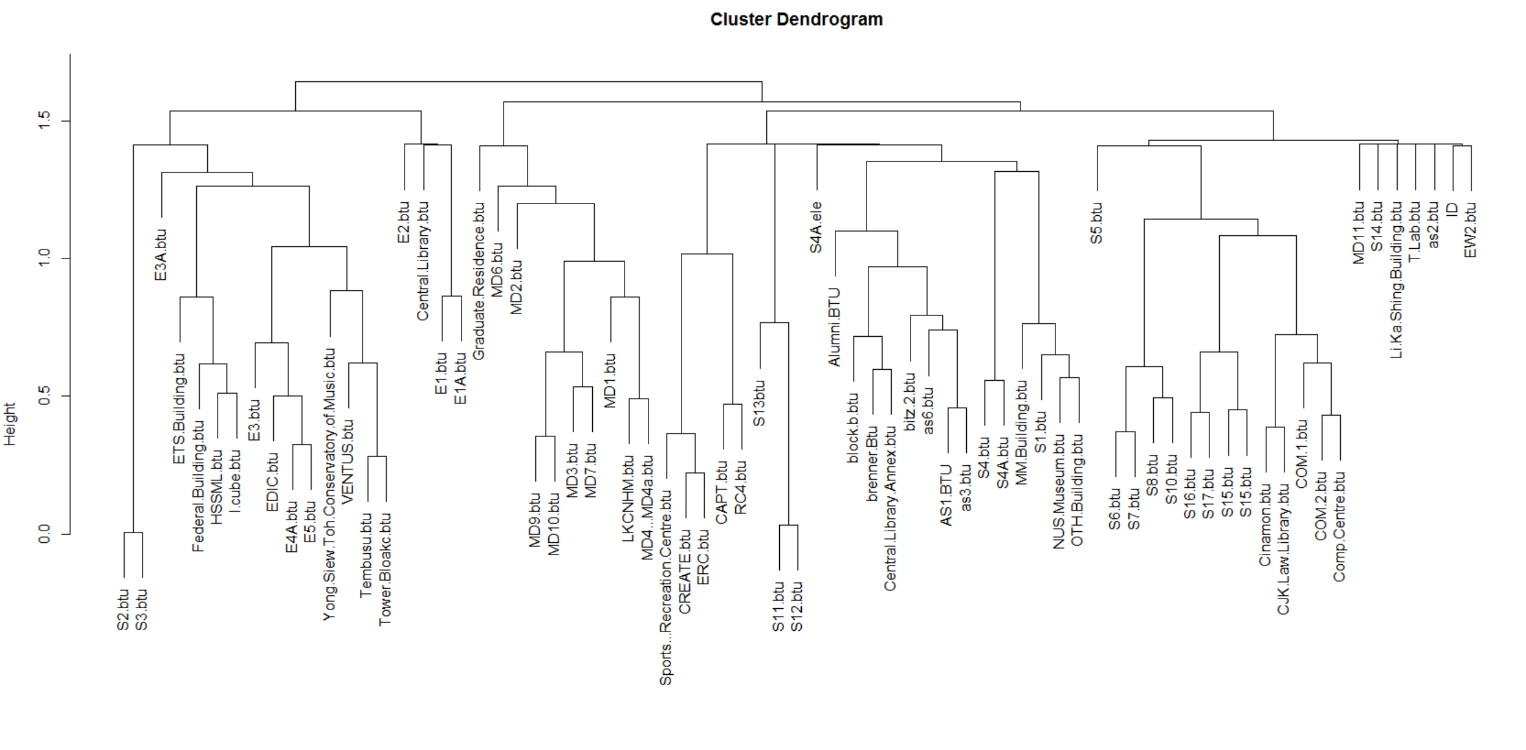


The buildings in the same cluster can be fitted in the same model .

**HEAT MAPS FOR STUDYING THE CORRELATION**

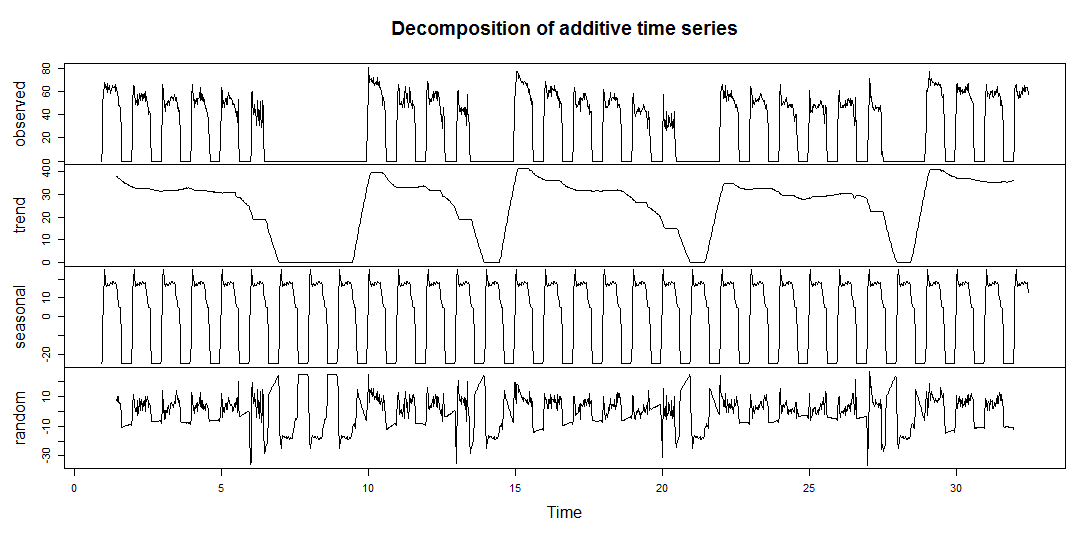


The correlation between one building with the other buildings is studied using the Pearson coefficient method.The result is represented using Dendograms.



In the above dendogram, the cut is made at 1.0.

**TIME SERIES DECOMPOSITION**

 The buildings in each cluster are decomposed to study their trend ,seasonality and irregularities(ramdonmness). Then they are fitted in the respective models.

**TIME SERIES PREDICTION MODELS**

The following models are used to model the data:

1.Holts Exponential Smoothing

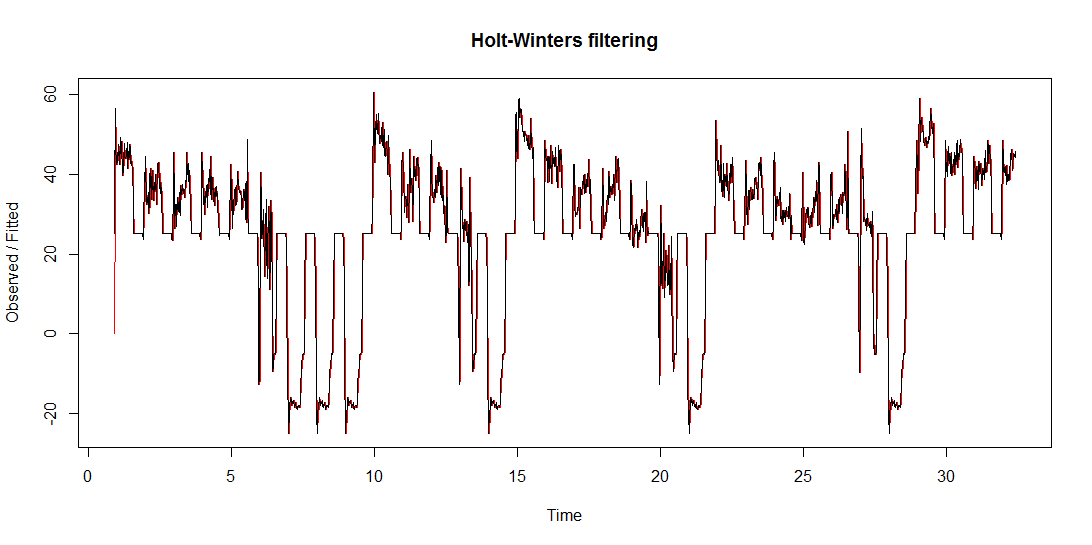
2.Holt Winters Model

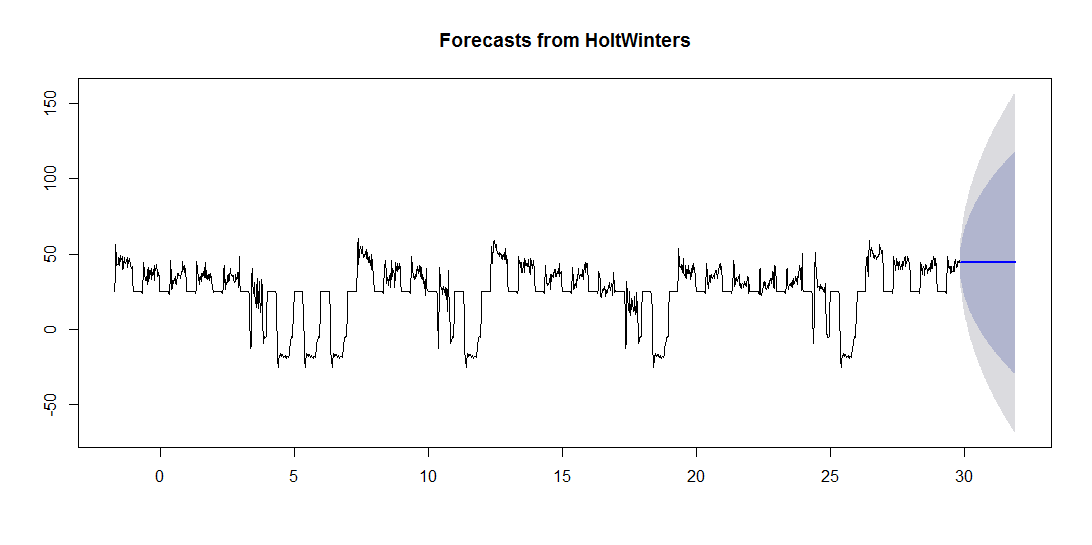
3.Support Vector Regression

4. ARIMA.

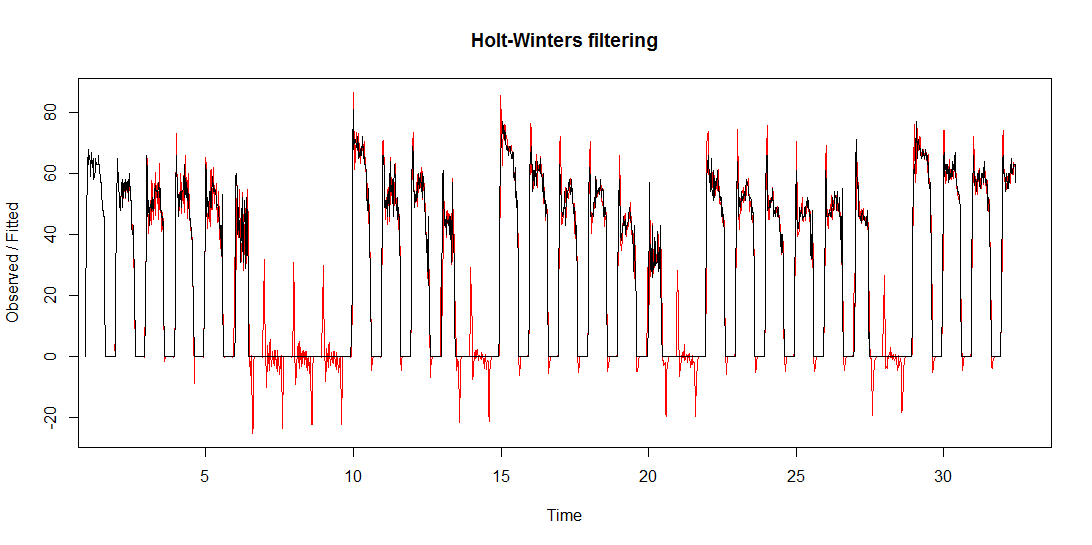
As the data is a univariate time series , only simple models like Holts Exponential Smoothing and Holt Winters Model fit the model best .The complex models like neural nets and SVR don’t fit well.

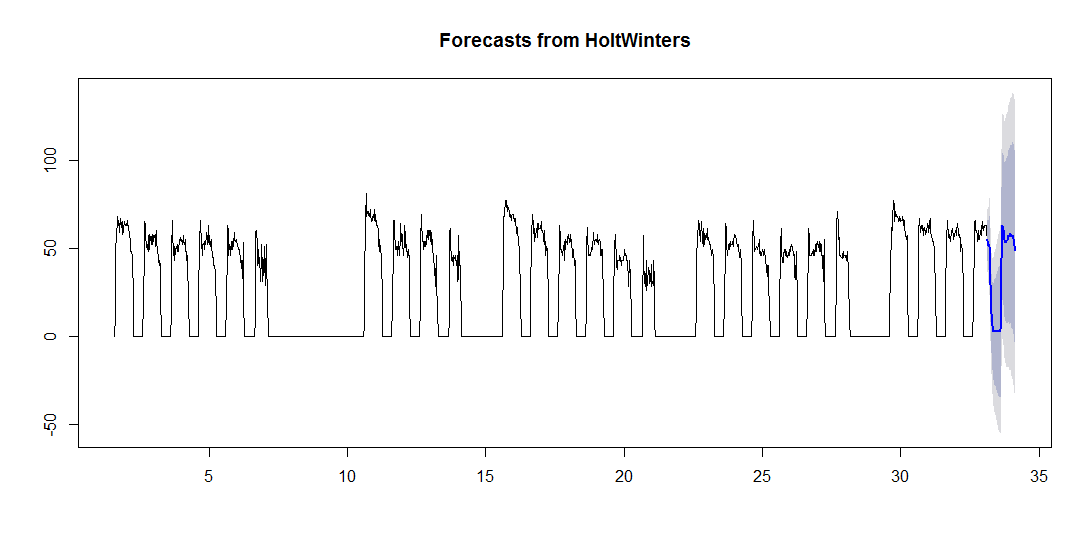
**HOLTS MODELLING (REMOVING SEASONALITY)**





**HoltWinters Exponential Smoothing (Includes Seasonality)**





Based on the different models , the performance index table below shows the fitness of different models.The values are calculated as a gross average over all the clusters.

**PERFORMANCE INDEX TABLE**

|  |  |  |  |
| --- | --- | --- | --- |
| S.NO | Model | RMSE | MAPE |
| 1. | SVR | 3.5 | 27.59 |
| 2. | Holts | 1.67 | 12.79 |
| 3. | Holt Winters | 2.04 | 13.09 |
| 4. | ARIMA | 2.81 | 19.014 |

**Observations And Conclusions**

From all the above models, we can find that Holts Exponential smoothing has low errors.

1. Hence, to improve the forecasting accuracy the seasonality has to be removed.
2. The high values of error is due to the Ramdomness in the Time series.
3. More attributes such as occupancy details, Temperature, C02 concentration can be helpful for fitting the model close to the data.

**Challenges**

1. Data is available only for one year
2. Lots of missing values
3. Irregularities are more in the data.
4. As the data is univariate, it is not suitable to fit advanced models.

**Future Works**

1. The clustering results will be used as an input for another research project. There is an approach for building one specific model for each time series. A single model can be trained using the data of time series and given the predictions for all the time series in the group.
2. Model used: Time delay recurrent neural networks . Unlike my approach, here each time series belongs to more than one group. The data is trained to learn more than one model.

**REFERENCES**

# K-Shape clustering algorithm for building energy usage patterns analysis and forecasting model accuracy improvement –Junjing Yang, Chirag Deb et.al

1. Forecasting diurnal cooling energy load for institutional buildings using Artificial Neural Networks.- Chirag Deb, Junjing Yang, Lee Siew Eang et.al
2. K-Shape : Efficient and Accurate Clustering of Time series- John Paparrizos and Luis Gravano.
3. Thorndike, Robert (1953) “Who Belongs in the family?”