CL 688 Initial Project Report

Forecasting Thermal and Electrical Loads for Efficient Building Management

by

Ashutosh Mulchandani - 200100037

&

Anika Matanhelia - 20d170003

Indian Institute of Technology Bombay

Mumbai - 400 076

September 2023

Abstract

A crucial element in the realm of building energy system design and the efficient management of power grids is the development of a precise and swift predictive model for building load. Traditionally, this demand has been addressed through physics-based simulation tools, but they often pose challenges due to their intricate model creation process and the need for engineering expertise. Here, machine learning algorithms, rooted in extensive data analysis, emerge as a solution to bridge this gap. These data-driven models offer a promising approach to tackle the complexities associated with building load prediction. We plan to use these data-driven approaches on data generated using complex simulation softwares

Keywords: Artificial Intelligence, Data-Driven Techniques, Multivariate Time Series, Thermal Load Prediction, HVAC.

Code files can be found here : `https://github.com/Scriea/CL688`

Contents

Introduction to the System

1.1 System Description

The energy efficiency of buildings is a critical point of interest in the field of sustainability. Building thermal load prediction plays an important role in energy management and efficiency. It has wide applications, such as in determining the capacity of heating, ventilation, and air conditioning (HVAC) systems in the phase of building design. The ability to predict the energy consumption of buildings can help to improve their design and operation, leading to significant savings in energy costs and emissions.

Big advantage of using data-driven models is that they do not require establishing thermal equilibrium equations and usually fewer inputs are required compared to physics-based simulation tools.

The idea is to find a mapping $F : X \to Y$, for predicting thermal loads and power loads without running heavy simulations which usually take hours of time. Here X is feature space.The widely used input variables included these three types:

1. Time-related information, such as the day type, occupancy, and equipment schedules
2. weather conditions, such as the temperature, humidity, and solar radiation;
3. building physical parameters, such as the window-to-wall ratio and R-value of the wall.

Y, the output targets are generally thermal loads or electricity consumption

1.2 Data Source

The dataset was obtained from internet. This dataset was generated using Energy simulation software, EnergyPlus and Python package eppy for co-simulation. A total of 230400 valid data samples were generated for analysis.

The dataset is hosted in a github repository, link.

Input/Output Data Description

2.1 Dataset Overview & Generation

This dataset encompasses 230400 samples, generated using EnergyPlus and Python simulations. Three categories of building energy models were developed:

1. Small office building - 1 floor - 511 m^2 area
2. Medium office building - 3 floor - 4980 m^2 area
3. Large office building - 12 floor + basement - 4632 m^2 area

Furthermore, the envelope types, HVAC system types, system operation schedules are different, which adds different distributions in the dataset, hence improves generalizability.
Seventeen key input variables were changed with the appropriate steps in each EnergyPlus. An uncertainity of ± 20% was introduced for each of the 17 variables, except for the cooling and heating temperature set points. For these two input variables, we used $\pm 1.11C$ ±2 of the default values as the lower and upper limits, so as to ensure that the indoor temperature was within a preferable range.

2.2 Input Features

The dataset comprises 17 distinct attributes. The primary objective of this dataset is to harness the information contained within the 17 features to make predictions. To elaborate on the specific nature of these features and responses:

No.	Input Feature	Unit
1	Dry Bulb	$^{\circ}C$
2	Relative Humidity	$\%$
3	Global Horizontal Radiation	Wh/m^2
4	Wind Speed	m/s
5	Total Floor Area	m^2
6	Aspect Ratio	-
7	Window-to-Wall Ratio	-
8	Floor Height	m
9	Exterior Wall Insulation R-value	$(m^2K)/W$
10	Roof Insulation R-value	$(m^2K)/W$
11	Specific Heat for Internal Thermal Mass	J/(kg·K)
12	Cooling Temperature Set Point	◦C
13	Heating Temperature Set Point	◦C
14	Fresh air volume	$m^3/s-m^2$
15	People Density	$m^2/person$
16	Lighting Power Density	W/m^2
17	Electric Equipment Power Density	W/m^2

Table 1: Input variables

2.3 Output Variables

The two output variables are:

1. **Total Thermal Load**: A real-valued output variable that quantifies the amount of electrical energy required to heat/cool a building to a desired range of temperature. It serves as a crucial metric for assessing a building's thermal performance and energy efficiency, aiding in the design of efficient heating systems and insulation.

2. **Total Electric Load**: Additionally, the dataset includes the Total Electic Load other than that required for HVAC. It is crucial for designing the electrical components for the building such as cable sizing, meters and breakers etc.

These values serve as ground truth for model evaluation and validation.

2.4 Data Visualization

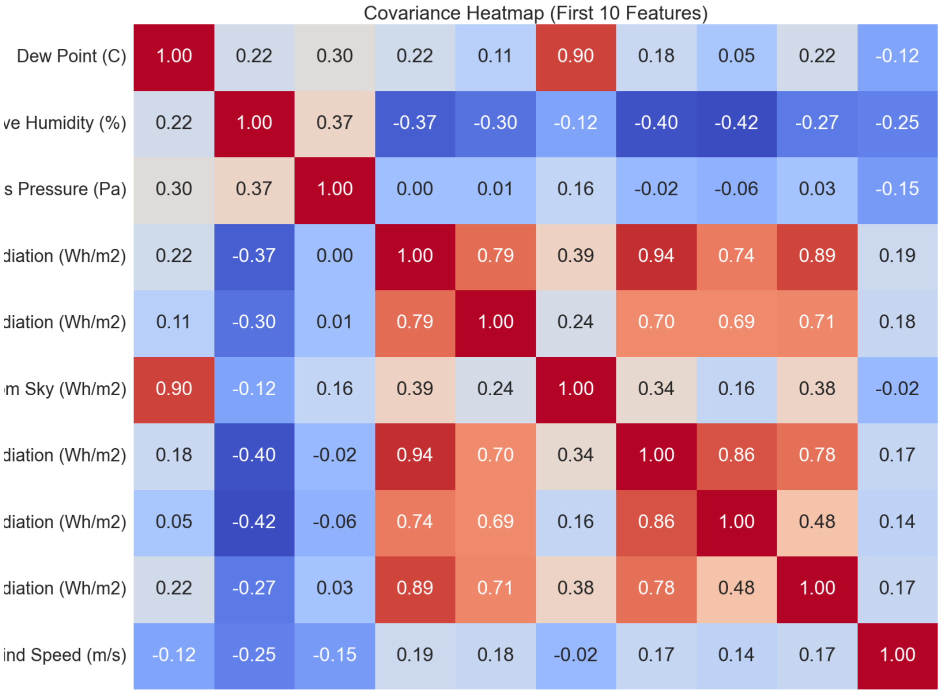

Figure 2.1: Correlation Heatmap of Features

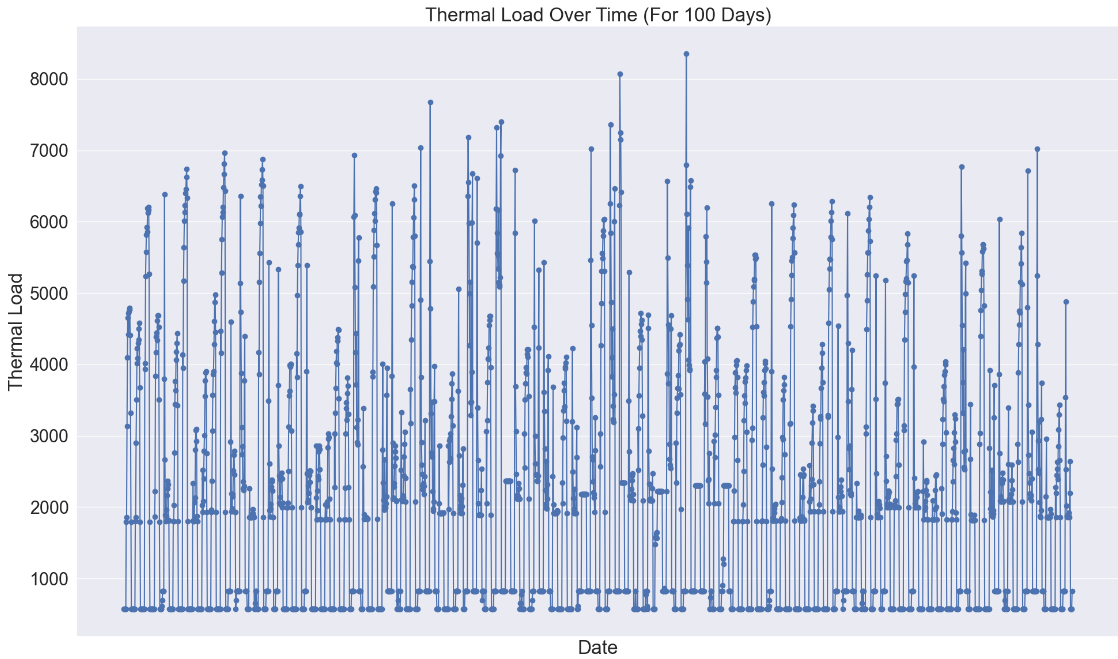

Figure 2.2: Line Chart

Problem Statements

This chapter outlines the problem statements that this project aims to address. Each problem statement is categorized into either regression or classification.

3.1 Problem Statement 1: Thermal Load Prediction

Objective: Develop a regression model to predict the thermal load of the building based on different features of the building.

3.2 Problem Statement 2: Electrical Load Prediction

Objective: Develop a regression model to predict the cooling load of the building based on the different features of the building. The model should provide accurate estimates of the electrical load for optimal design.

3.3 Problem Statement 3: Load Demand Classification

Objective: Implement a model to to categorize the demand for energy or thermal load into different four classes. The classes represent various demand characteristics, such as constant, increasing, decreasing, or cyclical demand. Since this is a sequential tasks we plan to use sequential models such as RNNs, LSTMs or GRUs.

3.4 Problem Statement 4: Load Anamoly Classification

Objective: An anomaly detection task to identify unusual or unexpected load patterns. We plan to train a model to detect outliers or anomalies in the Total Thermal Load and Total Power Load data, indicating when the load deviates significantly from the expected behavior.

Bibliography

[1] C. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006.

[2] J. VanderPlas, *Python Data Science Handbook: Essential Tools for Working with Data*, O'Reilly Media, Inc, Nov. 2016.