# Group #6 Assignment 1

Mahmut Osmanovic Isac Paulsson Sebastian Tuura
Mohamed Al Khaled

January 2, 2025

# Contents

**1 Introduction** **3**

**2 Methodology** **4**

**3 Results & Analysis** **5**

3.1 Air temperature v. Total energy production (Analysis A) . . . . . . 5

3.2 Wind power v. Wind Speed . . . . . . . . . . . . . . . . . . . . . . . . 5

3.3 Solar power v. Temperature, Sunshine Time & Irradians (Analysis C) 5

3.4 Irradiance, and Solar Time v. Total-Energy-Production (Analysis D) 6

# 1 Introduction

By understanding the relationship between weather conditions and energy output, energy producers and policy-makers can better predict production levels, optimize resource allocation, and improve the stability of the energy grid.

This is a study of the weather and its effect on energy production in Sweden, with data collected from *Sveriges meteorologiska och hydrologiska institut (SMHI)*, regarding weather from selected stations placed through out Sweden. Compared against energy production statistics gathered from the Swedish statistical agency *Statistikmyndigheten (SCB)*. The data spans a period of 12 to 90 months allowing us to capture seasonal patterns from throughout a typical Swedish year, seasonal variations, such as cold winters and long summer days, are expected to have a substantial impact on both energy demand and generation.

KNIME was used to analyze the data. KNIME is an advanced data analytics platform, which facilitated data integration and the application of clustering and regression techniques.

With the help of different visualization techniques and through the use of clustering in combination with regression techniques, we where able to both identify the patterns in the data as well as visualize the relationships between the weather variables and the energy usage, providing a clear view of how weather influences energy demand in Sweden.

Insights like these can help energy producers understand seasonal dependencies and optimize energy procured based on weather patterns.

# 2 Methodology

Data was gathered from two Swedish agencies being, *Sveriges meteorologiska och hydrologiska institut (SMHI)* which is a government agency under the Ministry of Climate and Enterprise, with the mission to serve as an expert body in meteorology, hydrology, oceanography, and climatology. And *Statistikmyndigheten (SCB)* which is responsible for developing, producing, and disseminating official statistics and other government statistics, as well as for coordinating the system for official statistics.

A period of between 12 to 90 months was chosen to encapsulate the climate changes of the entire Swedish year, as well as allowing us to view how the Swedish energy usage varies with the different seasons.

The data was aggregated by month, and the daily and/or hourly values were calculated as the mean.

For clustering, hierarchical clustering was used. KNIME the tool used for creating the workflow which was mentioned before has two types of clustering nodes, hierarchical clustering and k-means. Hierarchical clustering is useful when the number of clusters is unknown since k-means requires you to have previous knowledge of the number of clusters.

Linear regression in conjunction with a regression predictor was used to explore the different relationships between the weather variables and the energy usage as well as calculating the coefficient of determination ($r^2$).

The weather variables gathered from SMHI are Air Temperature, Rainfall amount, Sunshine time, Wind speed, Cloud percentage, Humidity percentage and Irridans. They are in different combinations compared and plotted against the total energy usage/consumption for a specific type of energy source i.e. Wind power or Solar power, (from SCB).

# 3 Results & Analysis

## 3.1 Air temperature v. Total energy production (Analysis A)

In Analysis A data was collected in the span of 90 months from the following 5 stations (Jönköping axamo, Stockholm observatoriekullen, Gubbhögen, Göteborg, Malmö). The analysis uses the mean air temperature between the stations. According to the analysis there exists a strong correlation between the air temperature and the total energy production seen in 1 where a low temperature during winter coincides with higher energy production. This is also supported by figure 2 showing a linear relationship between the parameters.

This suggests that a majority of the power production is incurred by lower temperatures which in turn raises the need for energy in heating applications, this is reasonable given the colder climate in Sweden. Another consideration is the energy consumption of cooling during hotter temperatures which does not seem to impact the energy production. It might still be an influence in hotter climates.

## 3.2 Wind power v. Wind Speed

The wind speed was gathered from four stations, Jönköping Axamo, Göteborg A, Malmö A and Bromma Flygplats for the last 90 months, ending in june 2024. The results can be viewed in figure 3. We can first observe that the wind power is increasing by each year, this is most likely due to the increase in the amount of wind turbines. Note that the amount of wind power that is produced varies in a cyclical manner. There is more energy produced during winter time. The second image in figure 3 highlights the relationship between wind power and wind energy. There is a positive correlation between wind power and the amount wind energy, albeit, not that strong. Nonetheless, by using a hierarchical plotting technique, one can split the data points into two groups, warmer and colder climate.

## 3.3 Solar power v. Temperature, Sunshine Time & Irradians (Analysis C)

Data regarding Temperature, Sunshine Time & Irridans are gathered from the four following stations, Göteborg, Stockholm, Jönköping* and Karlstad. Stations have some variations here for instance all stations regarding solar so sunshine time & irridans are seperate from the rest and are labeled as such *Göteborg - Sol* by SMHI.

* Jönköping swapped for Växjö regarding both Sunshine Time & Irridans since Jönköping did not have any solar stations & Växjö was the closest.

Clustering and analyzing a parallel coordinates plot with the named variables and the solar power usage provided by SCB shows a clear separation on the sunshine time where as the other variables seem to lack structure 4.

Plotting the cluster with regards to solar power usage & sunshine time shows very clear clustering. 5

### 3.4 Irradiance, and Solar Time v. Total-Energy-Production (Analysis D)

Analysis is performed on Global irradiance mean data collected from 5 stations in the span of 90 months. Hierarchical clustering shows two distinct clusters in figure 6. Plotting the data based on month (figure 8) gives a clear indication that the clusters are generally representative of the colder and the warmer months. In figure 6 a gap between the clusters is observed, a reason for this is not readily apparent.

An interesting find when comparing average temperature with irradiance is that the scatter plot takes on a circular shape which is cyclic in nature. The air temperature is seen to be out of phase with the irradiance. Colored by month the cycle is clearly visualized in figure 7.

Figure 9 shows prediction model for total energy based on additional parameters in Analysis D.

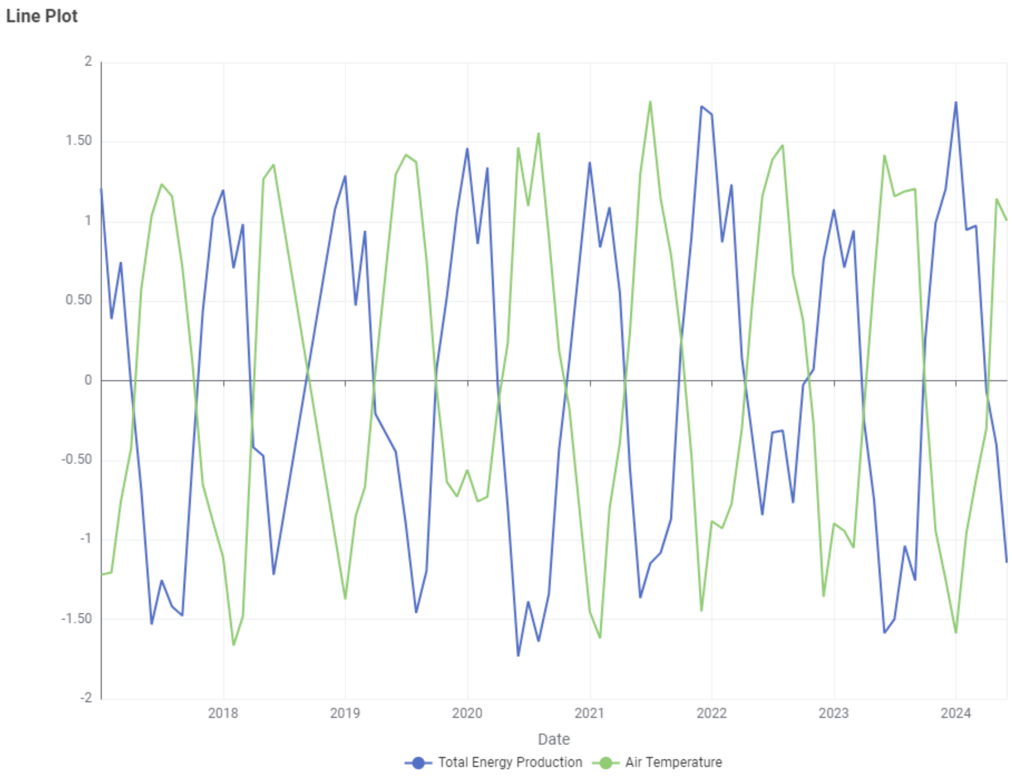


Figure 1: Air temperature vs Total energy production

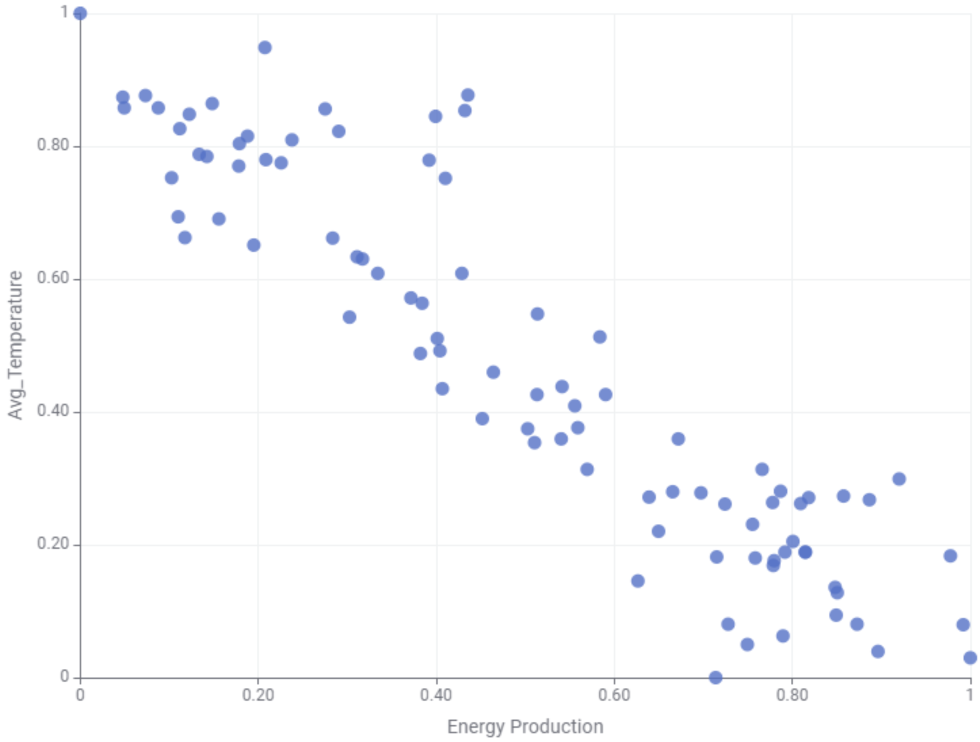


Figure 2: Average temperature vs total energy usage

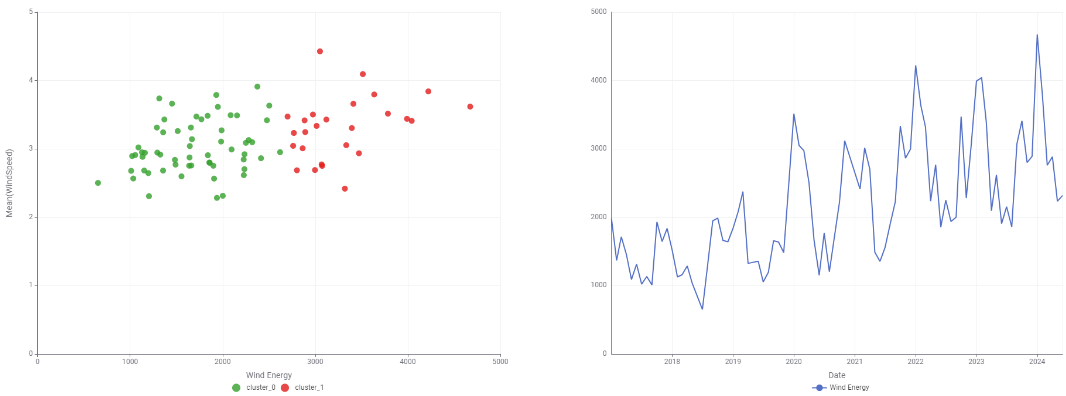


Figure 3: Wind speed vs wind power, Wind line graph

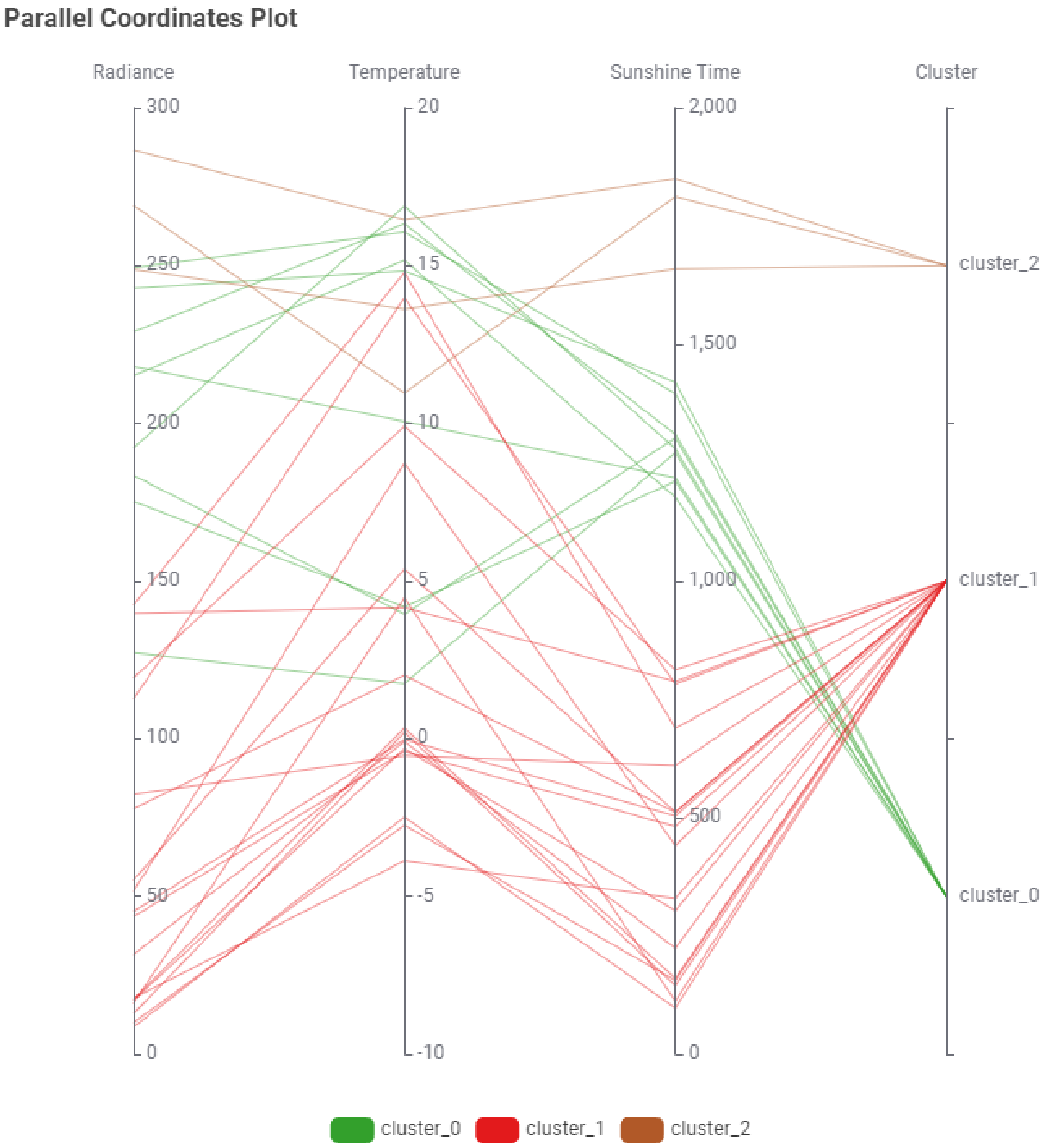


Figure 4: Parallel Coordinates Plot Solar

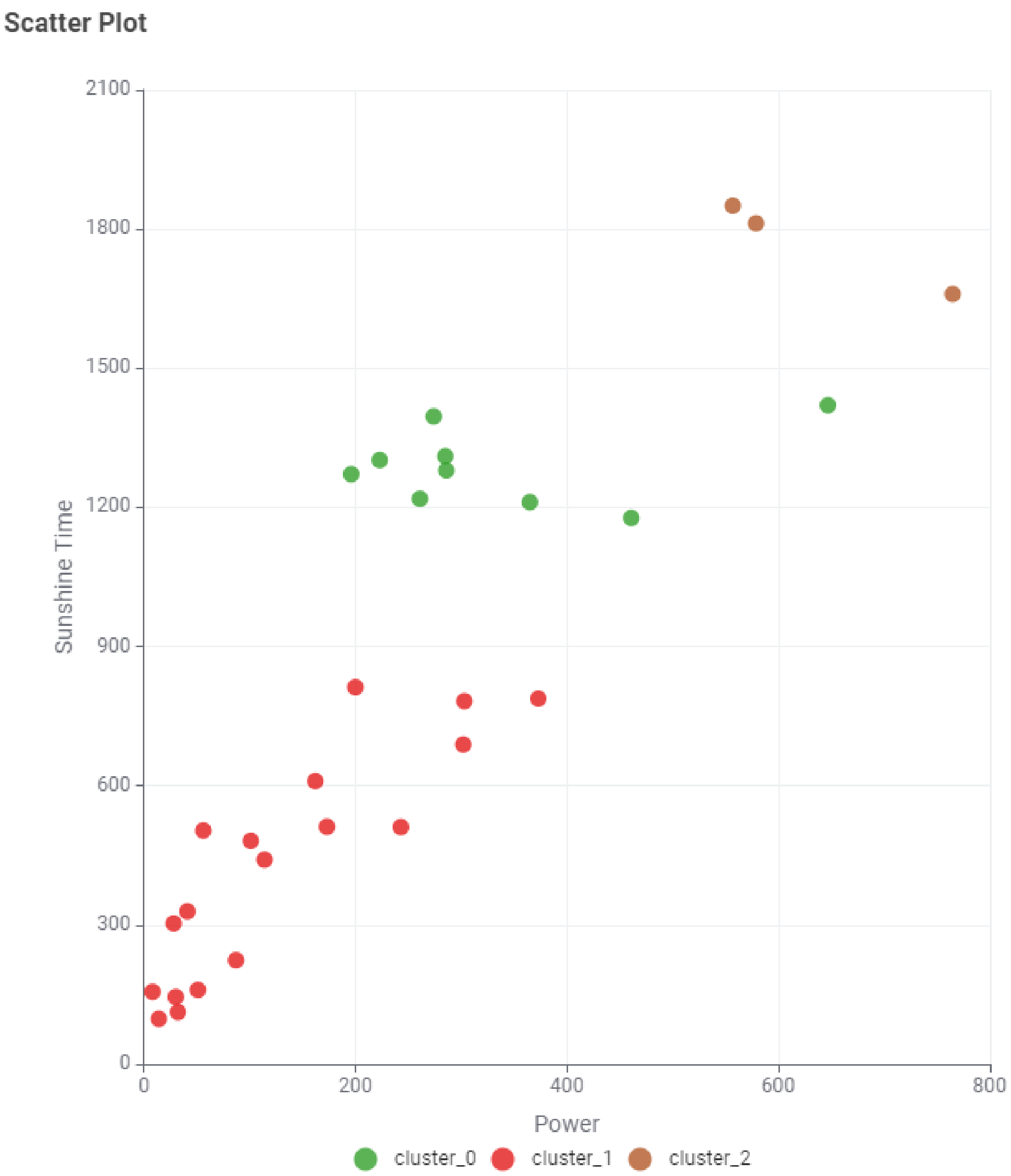


Figure 5: Scatter Plot Solar

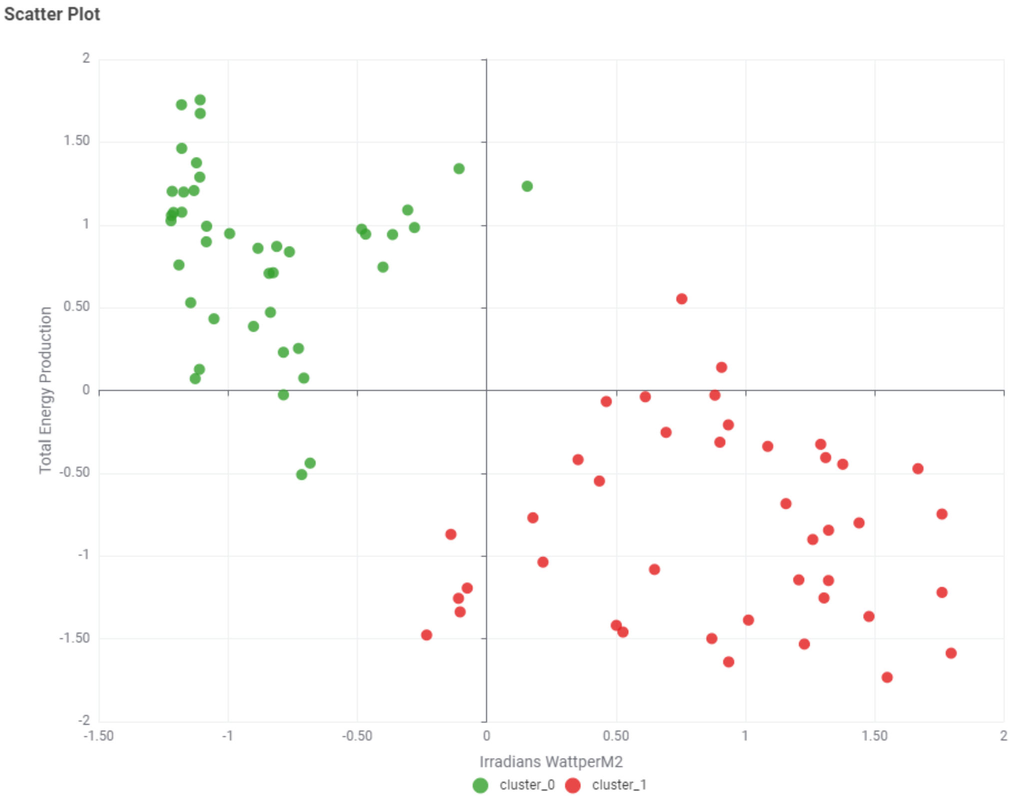


Figure 6: Averaged global irradiance vs total energy usage, clustered

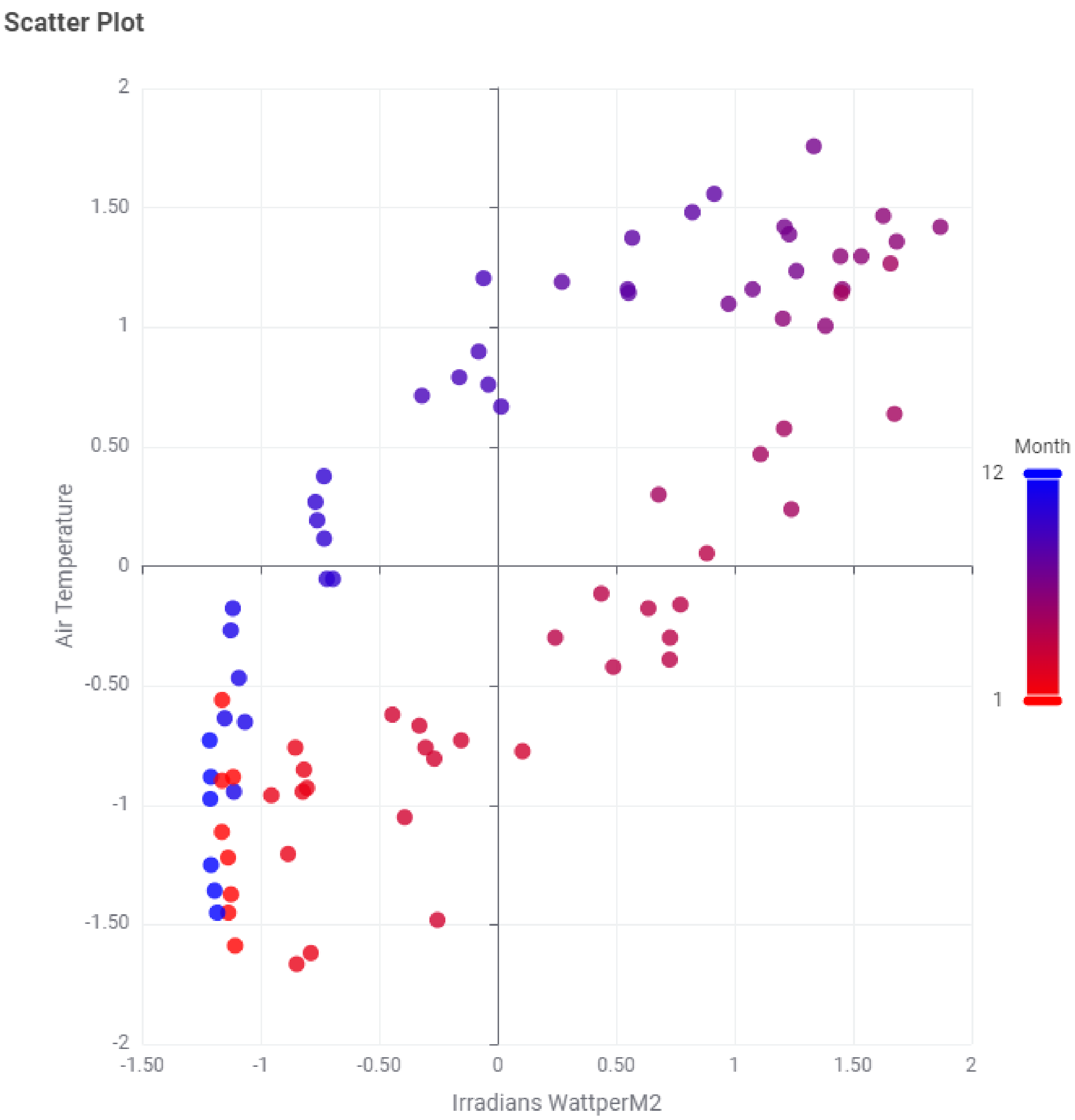


Figure 7: Average temperature vs irradiance, colored by month

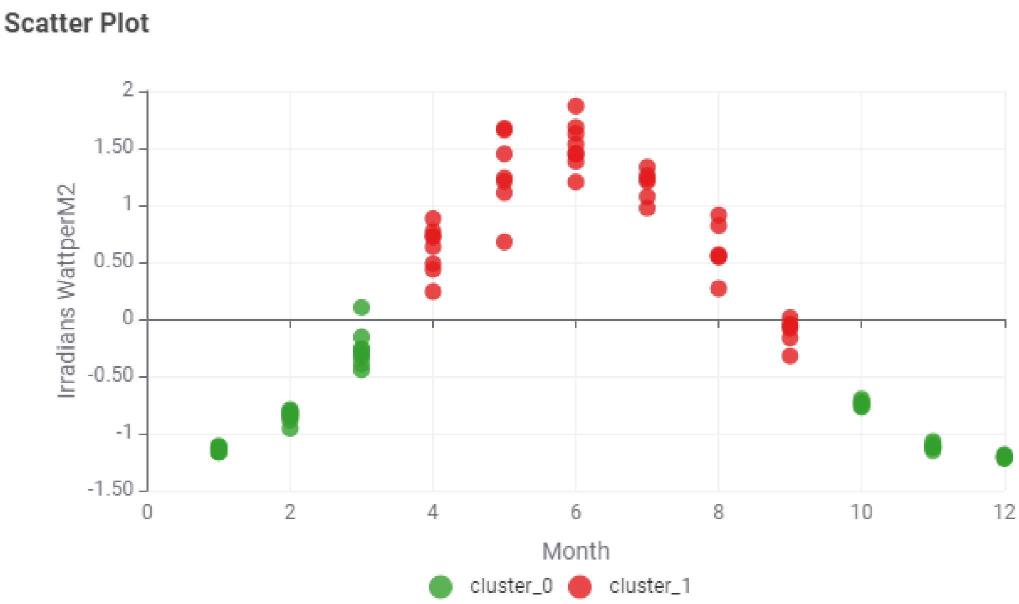


Figure 8: Averaged global irradiance sorted by month clustered

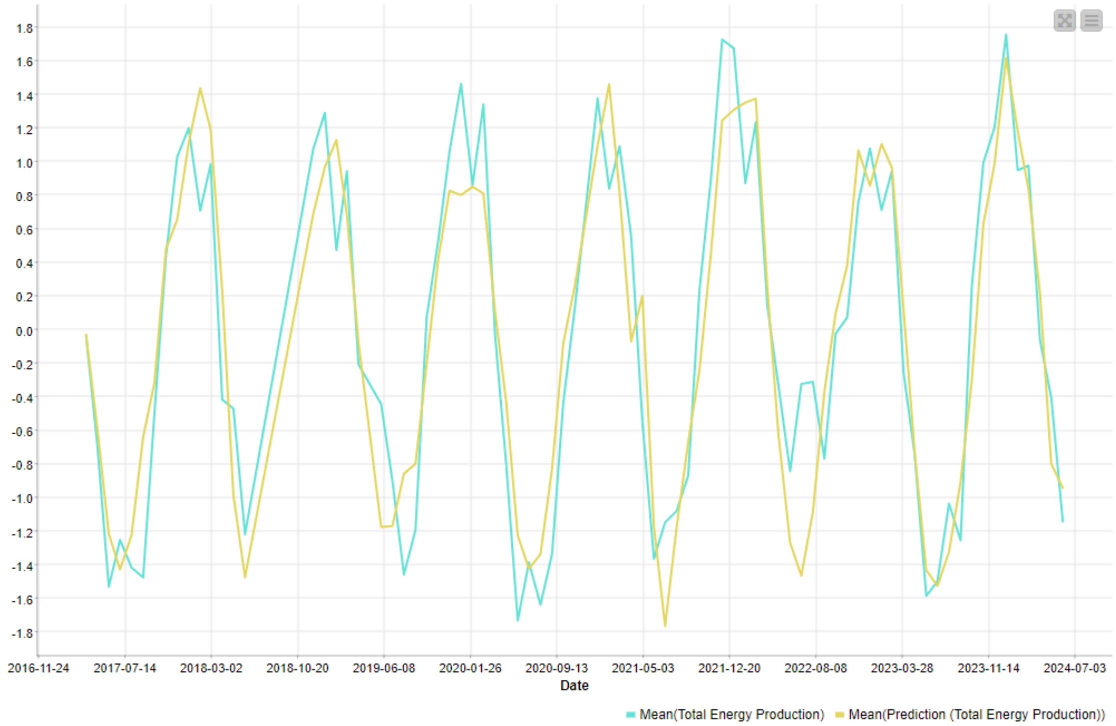


Figure 9: Total power prediction

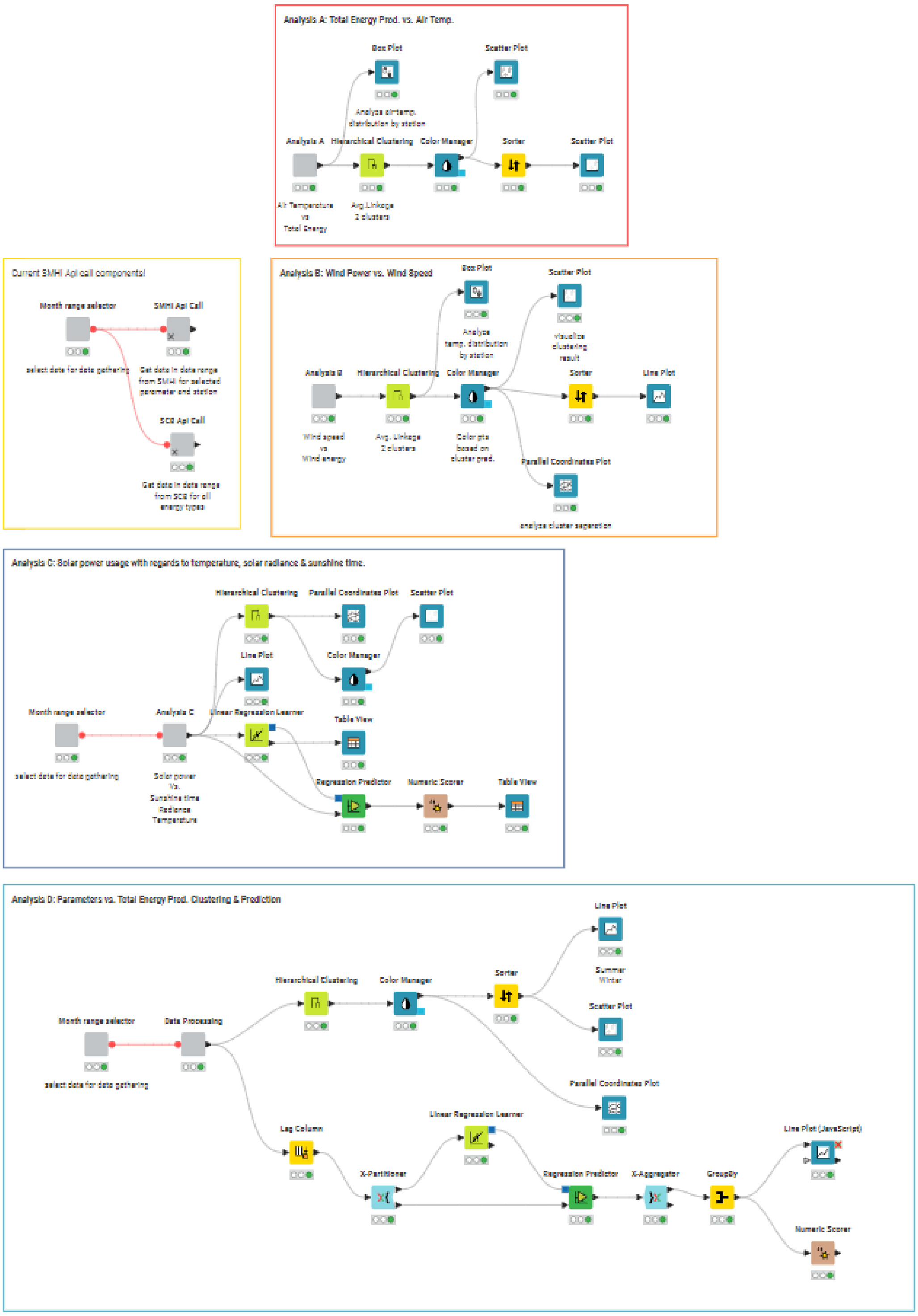


Figure 10: Workflow