Electrical Grid Cyber-Monitoring

Dec. 3, 2020

CMPT318 Cyber Security

Dr. U Glässer

Group 29

Richard Swann ID 823902440

Matt Fung ID 301329703

Jason Cao ID 200107583

Table of Contents

Table of Figures

1. Abstract

Concerns about the vulnerability of the Electrical Grid to cyber attacks is growing. The sophistication of cyber attacks is continuously increasing and the conventional defenses against these attacks are now of questionable value. Artificial Intelligence plays and will increasingly play an important role in the next generation of defense solutions. This report evaluates the performance of Hidden Markov Models to detect a cyber incursion and considers the applicability of Reinforcement Learning to the problem of protecting the Electrical Grid.

2. Introduction

The presence of an Advanced Persistent Threat (APT) in the Supervisory Control And Data Acquisition (SCADA) system for an Electrical Generation and Distribution Grid represents a serious threat to security, health and economic wellbeing of the Canadian population.

The Supervisory control and data acquisition (SCADA) for an Electrical Grid is controlling huge forces, equivalent to tons of TNT exploding per second e.g., BC Hydro is producing and distributing power equivalent to 2 tons of TNT exploding per second. Losing control of that power could lead to enormous economic damage, both to the grid itself and to the tens of thousands of pieces of economically vital equipment connected to it.

There are many different threat scenarios to consider each of which might require a different approach. A few of these are:

- Detecting an attack on the SCADA system and blocking the early phases of the "kill chain" to stop infection of the system.

- Identifying the presence of an established APT perhaps by sensing communications with its offsite controller.
- Discovering a quiescent APT as it becomes active and attempts to act against the electrical grid.

The use of Hidden Markov Models (HMMs) for protection in these scenarios is evaluated in reasonable depth and the potential of Reinforcement Learning (RL) is considered.

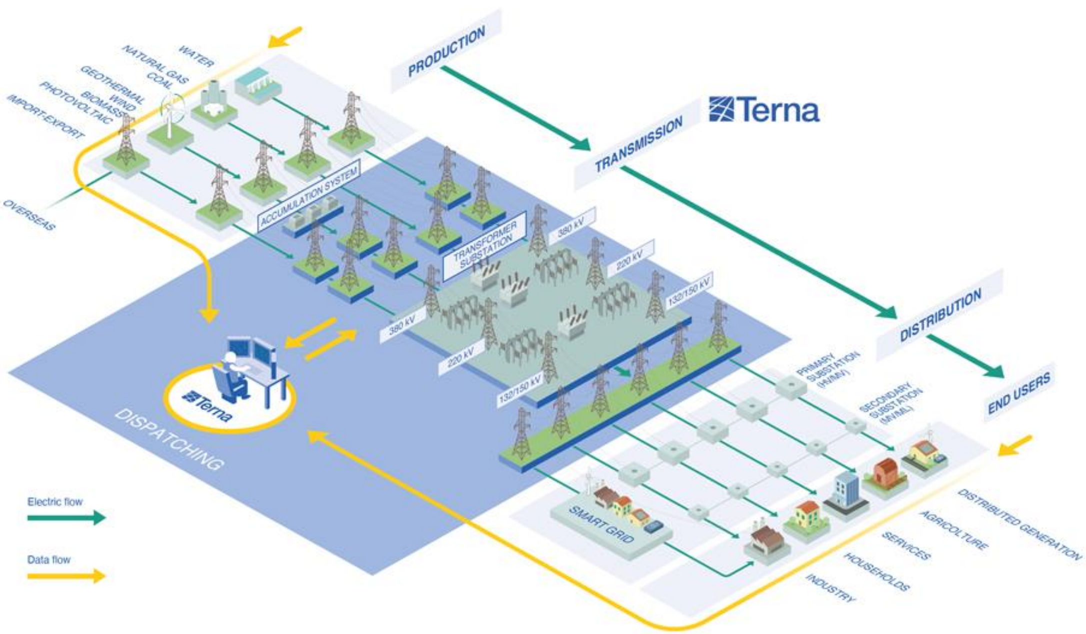


Figure 1 Generation and Distribution with Centralized Control

3. Electrical System Basics

In order to appreciate the breadth of the Threat Surface it is important to understand some of the key elements of the Electrical Grid and its relative vulnerability to attack

3.1. Voltage, Power, Current

The basic relationship is $Voltage\ (V) = \frac{Power\ (P)}{Current\ (I)}$

Current, measured in Amperes, is controlled by the customers and is the Load on the electrical system. As customers turn electrical equipment on and off, they change the current in the system. (Turning something "on" increases current.)

Power, measured in Watts, is controlled by the generating system. At any given time, a constant amount of power is being produced. The generating control system can change how much power is being produced but there is normally a lag measured in tens of seconds to minutes depending upon the generating technology in use.

Voltage, measured in Volts, is the "uncontrolled" element of the equation. For a given Power being produced the Voltage will fluctuate in a manner inversely proportional to the Load i.e., as Load (Current) increases Voltage in the system will decrease. This is what causes "brownouts" when the load exceeds the generation capacity of the system. Voltage drops until many pieces of equipment can no longer function. The reverse occurs if load is removed faster than generating capacity can be shed such as when a substation or a transmission line is suddenly taken down by component failure or natural disaster. Voltage will spike as the load is suddenly reduced. This can cause significant damage to equipment attached to the grid and to the grid itself, however because this is a not an uncommon issue the transmission system is designed to deal with it. Various types of automatic voltage limiters such as Metal Oxide Varistor Arresters are incorporated into the transmission and generation systems. These systems can absorb tremendous amounts of energy thus instantly increasing the Load and lowering the Voltage but only for a short period of time, seconds or fractions of a second. Therefore, power production must be reduced very quickly to eliminate the overvoltage otherwise the Arresters will be destroyed and Voltage across the system will rise unchecked causing enormous damage. Large scale power generation equipment generally has some automated mechanism for "crashing" the generator to reduce power output almost instantly when necessary, but these actions often carry a risk of damaging the generator so they are extreme measures that power companies would prefer to avoid.

3.2. Reactive Power

The Electrical Supply System also must deal with additional complexities. The basic equation $Voltage\ (V) = \frac{Power\ (P)}{Current\ (I)}$ is only true for purely "Resistive" loads. Loads can also be "Inductive" and "Capacitive". Inductive and Capacitive loads shift the phase of current and voltage relative to each other. This phase shift impacts the Power that can be delivered to Resistive Loads. If we define θ as the angle describing the phase shift between Voltage and Current, then $Effective\ Power\ (P) = \cos\theta * Voltage\ (V) * Current\ (I)$. If the phase shift were 30° the generators would have to produce 13% more power to compensate and this would create additional stress on the distribution system. Clearly power companies want to avoid the phase shift. This can be done by balancing Inductive and Capacitive Loads as they cause phase shifts of opposite sign. Unfortunately, almost all loads are either Resistive, heating, lighting or Inductive, motors, transformers, so the distribution networks include large banks of high voltage capacitors that can be switched into the network as required to correct the phase shift caused by Inductive Loads.

Capacitance is added and removed from the distribution network in a measured fashion as sudden changes in the network reactance (amount of Inductive or Capacitive Load) can induce large transients that wreak havoc on sensitive equipment and potentially portions of the grid.

3.3. Synchronization

The vast majority of the electrical grid operates on Alternating Current (AC) power. In North America that is the familiar 60Hertz cycle. The use of AC power imposes a major constraint on the grid, all the many hundreds of generators connected to the grid must be maintained in synchronization. This means producing AC power at exactly the same frequency and the same phase so that the power from many sources can be summed to meet the demand.

For the North American Western Grid that means generators from BC to New Mexico operated by 50 different major and many more minor suppliers must be synchronized. Failure to synchronize the generators result in major losses of usable power and large currents flowing that can damage transmission lines and generator substations.

3.4. Threat Attack Surface

An attacker with control of the SCADA system for an electrical grid has many options for doing damage as every major component is under computer control. Tripping substation breakers will generate large transients in the rest of the system. Switching banks of capacitors in and out can generate instability and currents that cause transmissions lines to fail. Forcing generators out of synchronization can damage generators and substations. These are all things that would be difficult to diagnose quickly particularly if the SCADA is generating false status information for the operators (à la Stuxnet).

4. Data Analysis

Electrical systems are complex, and the dominant characteristics are defined by the load that they are serving at any given time. The nature of the load varies continuously throughout the day and by the day of the week. Thus, a set of observations that are entirely normal at 13:00 on Monday might be highly anomalous at 01:00 on Sunday.

A detailed analysis of the data was undertaken to ensure that a meaningful set of variables was selected for the Hidden Markov Model Analysis and generally to understand the variability of the observations.

Several approaches to using Hidden Markov Model Analysis to identify anomalies were investigated for a selected time period and immediately adjacent periods.

4.1. Cleaning Procedure

The dataset was found to have approximately 1% of records with NA values, as these would have caused issues with subsequent analysis steps. It was decided to set the NA values to normal values selected to reasonably reflect the values displayed by the various variables in the adjacent time periods. The data was scanned column by column. Where an NA value was located it was replaced with the most recent good value from that column. For NA values that occurred in the first row of the dataset (one value) it was replaced with the mean of the column.

4.2. Principal Component Analysis

Previous work related on this project had dealt with a single arbitrarily selected time period. While the scope of this report is also limited to a single period, in order to understand what issues were likely to be encountered in developing a practical system to address the 24/7 requirement, an in-depth analysis of the complete data set was undertaken. The data was segregated by the day of the week, with Monday being day 1 to Sunday day 7. Then separated into six daily periods with each period spanning four hours, 00:00 – 03:59, 04:00 – 07:59, … 20:00 – 23:59. Finally a Principal Component Analysis (PCA) was done on each of the resulting 42 distinct time periods.

The expectation was that weekday periods covering the same time frame would be similar across the five days while Saturday and Sunday periods would differ from weekdays and possibly from each other. Principal Component Analysis of the 42 time periods revealed that this was not the case. The Heat Maps of the contribution of PC1 and PC1 and 2, Table 1 and Table 2, showed that even for the same weekday period, the differences from day to day are quite large. The complete results are tabulated in Appendix A – PCA Results for Complete Dataset.

PC1 Importance

Period	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
00:00 - 03:59	0.63	0.61	0.50	0.69	0.56	0.56	0.38
04:00 - 07:59	0.79	0.79	0.63	0.76	0.77	0.75	0.54
08:00 - 11:59	0.64	0.59	0.45	0.59	0.52	0.50	0.66
12:00 - 15:59	0.58	0.66	0.57	0.65	0.62	0.35	0.53
16:00 - 19:59	0.76	0.66	0.80	0.82	0.64	0.70	0.76
20:00 - 23:59	0.74	0.79	0.79	0.76	0.76	0.83	0.88

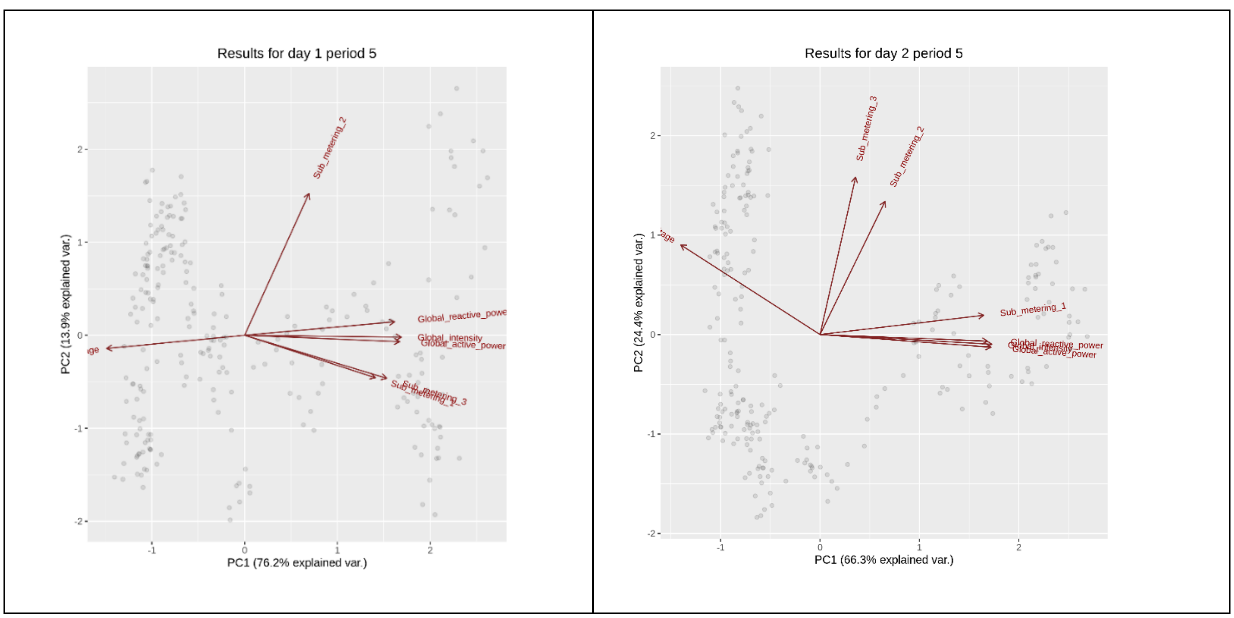
Table 1 Heatmap of PC1 Variance

PC1 and PC2 Importance

Period	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
00:00 - 03:59	0.82	0.77	0.70	0.85	0.72	0.72	0.71
04:00 - 07:59	0.89	0.91	0.81	0.91	0.89	0.89	0.73
08:00 - 11:59	0.80	0.80	0.72	0.76	0.74	0.68	0.80
12:00 - 15:59	0.78	0.79	0.75	0.80	0.79	0.69	0.71
16:00 - 19:59	0.90	0.91	0.88	0.91	0.79	0.89	0.91
20:00 - 23:59	0.89	0.89	0.91	0.89	0.89	0.93	0.95

Table 2 Heatmap of PC1 and PC2 Variance

Even the contribution of different variables to the Principal Components differed significantly from day to day for the same period as shown in Figure 2. The complete results are tabulated in Appendix B – PCA Plots for Complete Dataset.



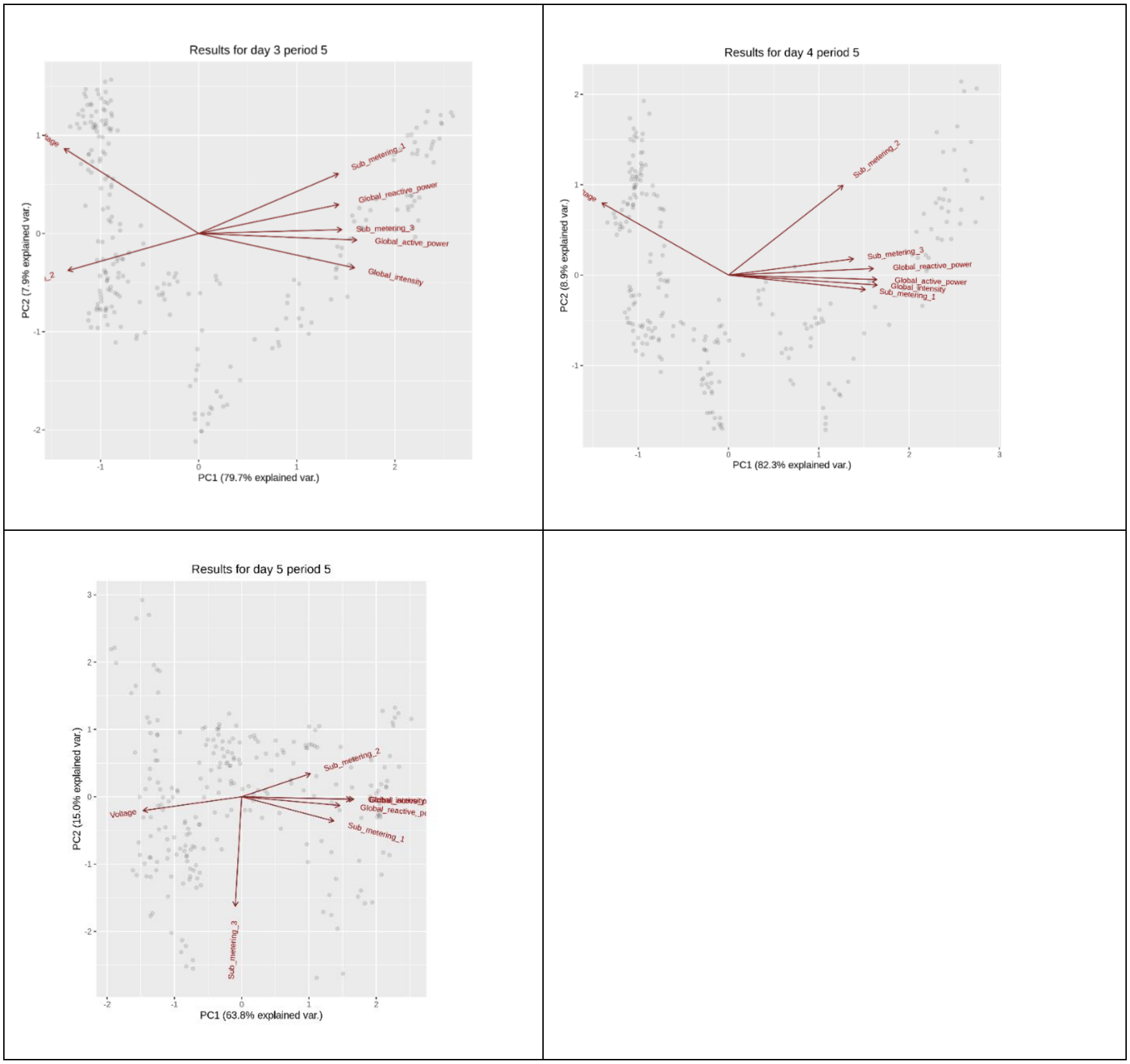


Figure 2 BiPlots of Weekday Period 5

From Table 2 the combination of PC1 and 2 explained 91% of the variance for the period 16:00 – 19:59 on Tuesday. Based on this information the period was selected for further analysis. The Biplot of the period Figure 3 showed that two or three variables were sufficient to derive PC1 and PC2. Additionally, it showed several variables defining PC1 were highly correlated and thus any single variable from that group would define PC1 adequately. The variables Global intensity, Global reactive power, Global active power, and Sub metering 1 were closely correlated according to the PCA and Voltage had a high negative correlation with these variables. All these variables had a similar magnitude, as a result, any of these variables were possible candidates for selection to represent PC1.

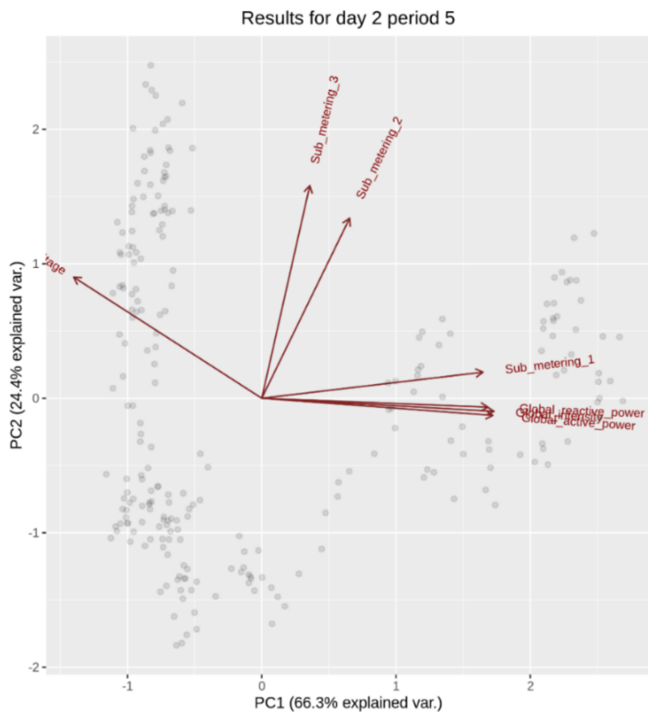


Figure 3 PCA Biplot for Tuesday, 16:00 – 19:59

For the multivariate tests adding Sub_metering_2 and Sub_metering_3 will capture the variance covered by PC2.

5. Hidden Markov Model

5.1. Variable Selection

As noted from the previous section, the flexibility of the time period Tuesday 16:00 – 19:59 placed several variables in contention for selection, namely being Global intensity, Global reactive power, Global active power and Sub metering 1. The variables of Sub metering 2 and Sub metering 3 were also considered due to both being closely correlated with each other.

From Table 3, Global intensity was an appropriate choice as it had a correlation of moderate strength to both Sub metering 2 and Sub metering 3. Global reactive power, Global active power and Sub metering 1 all had a correlation of weak strength to Sub metering 2 and Sub metering 3. Hence, these three variables were not chosen. The chosen variables of Global intensity, Sub metering 2 and Sub metering 3 presents a good balance as each variable has a correlation of moderate strength to each other. This balance is appropriate as it allowed an analysis to be done on three variables that were correlated to each other. The importance of variable selection will become evident in the section on the Multivariate Hidden Markov Model.

	Global active power	**Global reactive power**	**Voltage**	**Global intensity**	**Sub metering 1**	**Sub metering 2**	**Sub metering 3**
Global active power	1.000	0.099	-0.310	0.683	0.167	0.321	0.313
Global reactive power	0.099	1.000	-0.183	0.318	0.230	0.096	0.064
Voltage	-0.310	-0.183	1.000	-0.442	-0.231	-0.216	-0.169
Global intensity	0.683	0.318	-0.442	1.000	0.478	0.464	0.503
Sub metering 1	0.167	0.230	-0.231	0.478	1.000	0.065	0.071
Sub metering 2	0.321	0.096	-0.216	0.464	0.065	1.000	0.052
Sub metering 3	0.313	0.064	-0.169	0.503	0.071	0.052	1.000

Table 3 Correlation Values for Each Disjoint Pair of Response Variables

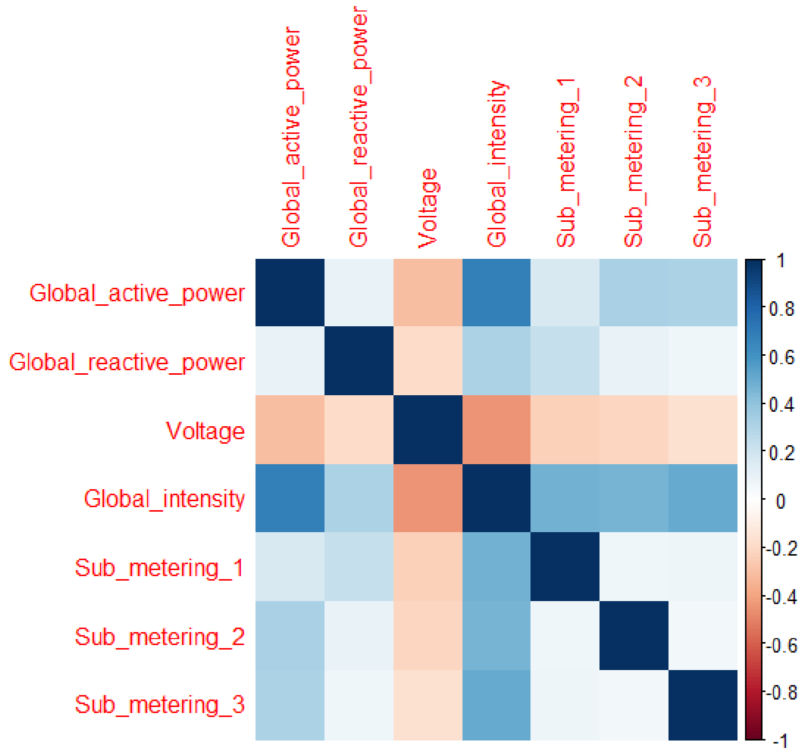


Figure 4 Correlation Matrix of Response Variables

5.2. *Univariate Hidden Markov Model*

5.2.1. Model Training

Global Intensity was chosen to be the feature for a univariate Hidden Markov Model. The dataset was partitioned into two sets: the former being a training dataset and the latter being a testing dataset. Both partitioned datasets only contained the values for Global Intensity during the chosen time window of every Tuesday, from 16:00 – 19:59. The training dataset consisted of all the values for Global Intensity from the years 2006 – 2008 while the testing dataset consisted of all the values for Global Intensity from the year 2009. As a result, an appropriate balance was achieved in partitioning the dataset for training and testing the univariate Hidden Markov Model for Global Intensity.

Prior to training the univariate Hidden Markov Model, the number of different states was chosen to be twenty. This number of different states allowed a wider selection for selecting an optimal model. An optimal model should neither be overfitted nor underfitted on the training dataset. For choosing an optimal model, the following criterion was used:

Determine a state with ...

1. A state with a negative Log-Likelihood that has a proximity to 0 and
2. A state with a large BIC

States	BICs	logLik	States	BICs	logLik
0	0.00	0.000	11	34845.01	-16701.610
2	104344.72	-52136.825	12	41195.84	-19750.106
3	92746.37	-46302.108	13	-91738.29	46854.029
4	75558.37	-37662.422	14	-44029.70	23146.962
5	70418.02	-35036.402	15	-133708.37	68143.676
6	63257.04	-31389.912	16	-276532.39	139723.218
7	59202.13	-29286.307	17	-278844.19	141056.805
8	53939.41	-26568.645	18	- 409486.82	206565.958
9	50761.91	-24883.433	19	-10698.25	7369.664
10	-110999.53	-56103.896	20	-333830.25	169143.813

Table 4 BICs and LogLik by number of States

The results from

Table 4 indicates that the univariate Hidden Markov Model with 11 states is the optimal model according to the criterion defined above. The univariate Hidden Markov Model with 11 states has a small negative Log-Likelihood and a relatively small BIC.

A graph depicting the relationship between BIC's and Log-Likelihood is shown below:

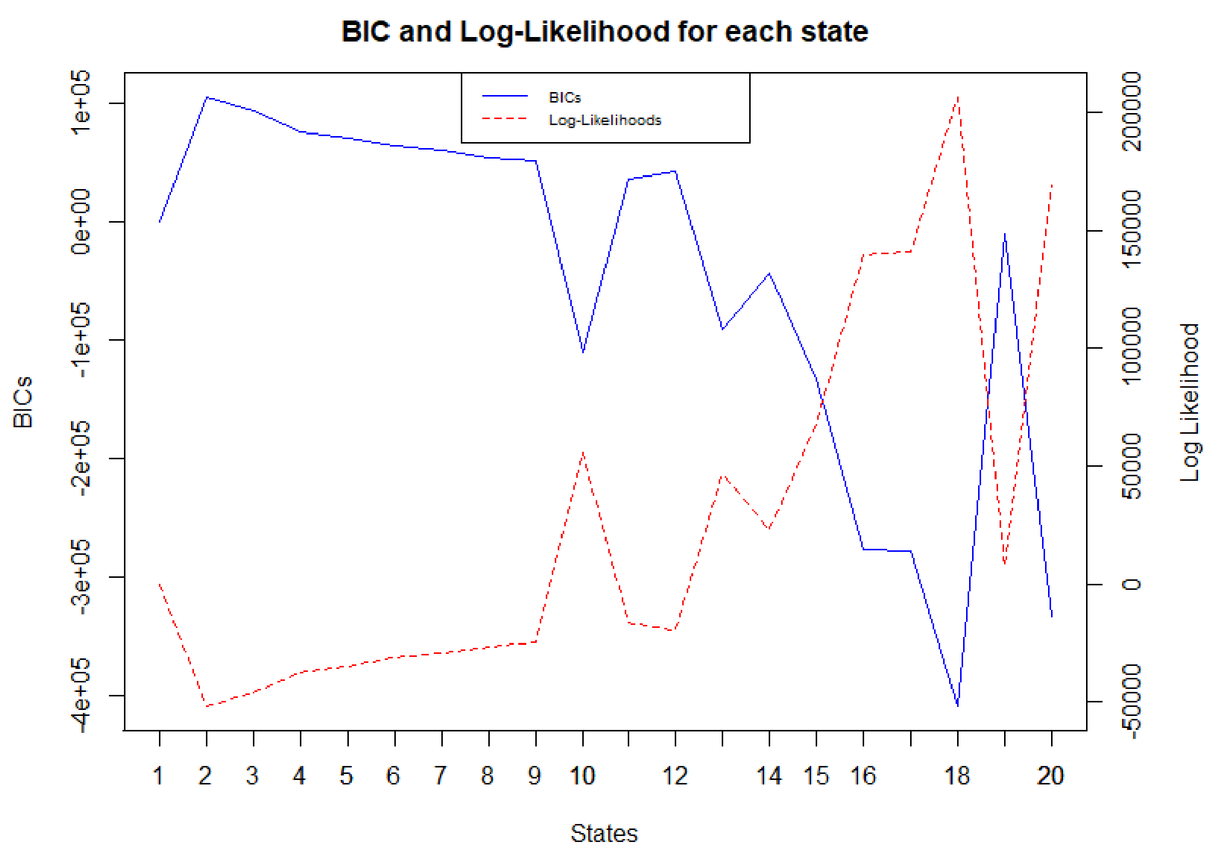


Figure 5 Plot of BIC and Log-likelihood by States for Univariate Model

Similarly, to the results in

Table 4, Figure 5 also indicates that the univariate Hidden Markov Model with 11 states is the optimal model. The univariate Hidden Markov Model with twelve states is also a sufficient candidate as it also had a negative Log-Likelihood that had a proximity to zero and had a small BIC. However, the univariate Hidden Markov Model with 11 states remains to be the optimal state as it has the best results according to the criterion in comparison with all other nineteen states.

5.2.2. Model Testing

The univariate Hidden Markov Model with state 11 was evaluated with the testing dataset. In order to validate the optimal model, the test model had to yield similar results. The model testing process required the parameters of the optimal model to be fed to the test dataset. This allowed the optimal model to compute the Log-Likelihood for the test model with state eleven. The results were normalized in order to compare the Log-Likelihood values of the train and test models.

Univariate HMM For Global Intensity: State = 11
Train Model logLik: **-21598**
Test Model logLik: **-22155**
Difference in logLik value: **557**

Table 5 Comparison of Trained and Tested Models for Univariate Model

The results from Table 5 were consistent for a fitted model. Under the assumption of a fitted model, it should neither be overfitted nor underfitted. The Log-Likelihood for state eleven for both the trained and test models were nearly equal. Thus, the univariate Hidden Markov Model with state eleven has been validated as an optimal model.

5.3. *Multivariate Hidden Markov Model*

5.3.1. *Model Training*

To train multivariate models on normal electricity consumption data, the design of HMMs relies heavily on PCA-based feature selection. Accordingly, Global Intensity, Sub metering 2 and Sub metering 3 were chosen as the important variables for modeling multivariate HMMs using the following time window: Tuesday 16:00 - 19:59.

Data splitting was sequential and straight forward. The given dataset was divided into training set and test set with the last year's data (2009) to be designated the test set and the data from prior years (2006 - 2008) assigned as the training set.

The training set was repeatedly fitted (from states 2 to 20 as recommended), in order to obtain the optimal state thus the optimal model for the HMM.

A table summarizing trained multivariate HMMs:

States	**BICs**	**logLik**	**States**	**BICs**	**logLik**
0	0.00	0.000	11	180596.6	-89465.7
2	402762	-201325	12	189155.9	-93608.3
3	274050.3	-136924	13	182776	-90271.1
4	235530.1	-117608	14	184835.2	-91143.4
5	222614.6	-111084	15	174629.8	-85873.1
6	219574	-109487	16	172518.4	-84639.7
7	206744.8	-102987	17	93897.53	-45141.4

8	202457.4	-100746	18	171227.7	-83608.5
9	200886.6	-99854.4	19	173858	-84715.5
10	193576.7	-96082.7	20	86444.6	-40790.5

Table 6 BICs and LogLik by number of States for Multivariate Models

Table 6 shows the evaluation of trained multivariate HMMs from state 2 to state 20 through their respective BICs and loglikelihood values.

A graphical representation of trained models is plotted below (BICs and loglikelihood vs. States) to compare their performance against one another.

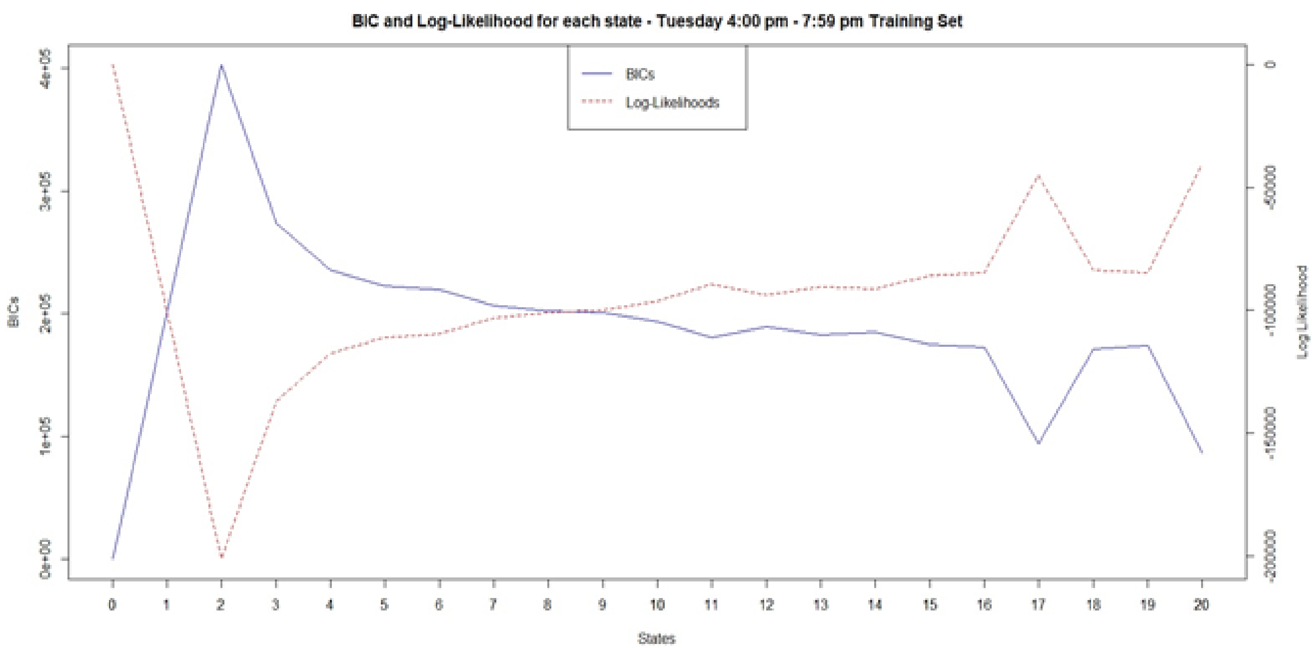


Figure 6 BIC and Log-likelihood by State

The criterion to select the best HMM (i.e., the optimal state) are: negative loglikelihood value close to 0 while maintaining a low BIC value. Based on Figure 6, HMM state = 8 appears to be optimal.

5.3.2. Model Testing

The multivariate HMM with state = 8 is evaluated with the test dataset. This was done by comparing normalized loglikelihood of trained model with normalized loglikelihood of test model. To validate the optimal training model, both loglikelihood values have to yield similar results.

Multivariate HMM: State = 8
Train Model logLik: **-99725**
Test Model logLik: **-98729**
Difference in logLik value: **996**

Table 7 Comparison of Log-Likelihood of Trained Model and Tested Data

As shown in Table 7, the difference in loglikelihood between the trained model and test model is small (996) given the magnitude of their individual values. Consequently, the selection of HMM with state = 8 seems to be consistent for a fitted model and can be subsequently deployed for anomaly detection.

5.4. Anomaly Detection

5.4.1. Anomaly Detection with Univariate HMM

The determined univariate optimal model for Global Intensity was further evaluated by using three separate datasets that contained anomalies. The optimal model was used to detect any instances of anomalies in all the three datasets. Each of the three datasets were filtered to only contain the values of Global Intensity from the chosen time window. Each dataset was used to create an individual model. In order to determine anomalies, the process used was similar to the one in Model Testing. The process required the parameters of the optimal model to be fed to each of the datasets containing anomalies. This allowed the optimal model to compute the Log-Likelihood for each of the models containing anomalies. To maintain consistency, the results of each model were normalized in order to compare their Log-Likelihood values to the optimal model.

Dataset Containing Anomalies 1	**Dataset Containing Anomalies 2**	**Dataset Containing Anomalies 3**
Train Model logLik: **-21598**	Train Model logLik: **-21598**	Train Model logLik: **-21598**
Test Model logLik: **-29228**	Test Model logLik: **-48799**	Test Model logLik: **-49609**
Difference in logLik value: **7630**	Difference in logLik value: **27201**	Difference in logLik value: **28011**

Table 8 Comparison of the Log-Likelihood of Trained and Tested Models Containing Anomalies

The results from Table 8 indicated that anomalies were detected in datasets 2 and 3. Likewise, anomalies were detected in the dataset 1 but was not as evident. Anomaly detection was deduced by comparing the Log-Likelihood values for the optimal model to each of the models containing anomalies. The difference in the Log-Likelihood values between the optimal model and the models containing anomalies for datasets 2 and 3 were significant. Thus, anomalies exist in both datasets 2 and 3. The difference in the Log-Likelihood values between the optimal model and the model containing anomalies for dataset 1 is smaller in magnitude in comparison to datasets 2 and 3. This demonstrates that there are fewer anomalies in dataset 1.

5.4.2. Anomaly Detection with Multivariate HMM

The optimal multivariate HMM was further evaluated using three separate datasets (of the same chosen time window - Tuesday 16:00 - 19:59), all contain anomalies. Specifically, the selected HMM was used to detect any instances of anomalies in all three datasets. Anomaly is detected when the difference between normalized loglikelihood of trained model with normalized loglikelihood of anomaly data is significantly larger than the difference between normalized loglikelihood of trained model vs. normalized loglikelihood of test model.

Dataset Containing Anomalies 1	**Dataset Containing Anomalies 2**	**Dataset Containing Anomalies 3**
Train Model logLik: **-99725**	Train Model logLik: **-99725**	Train Model logLik: **-99725**
Anomalies 1 logLik: **-123153**	Anomalies 2 logLik: **-118665**	Anomalies 3 logLik: **-120200**
Difference in logLik value: **23428**	Difference in logLik value: **18940**	Difference in logLik value: **20475**

Table 9 Comparisons Between Log-Likelihood of multivariate HMM and Anomalies datasets

The results from Table 9 confirms that anomalies do exist in all three datasets. Compared with the difference in loglikelihood between trained multivariate HMM vs test model (996), the difference in loglikelihood between trained multivariate HMM vs anomalies data are significantly larger.

5.5. *Anomaly detection in Adjacent and Longer Time Periods*

Prior testing has focused on periods where the training and the test data covered the same relatively short period. If the use of multiple short test periods is a requirement a system able to provide24 /7 monitoring would require a significant level of effort and high degree of complexity to create and maintain the 42 HMM sets considered earlier.

The ability of a multivariate HMM trained for one period to detect anomalies in adjacent periods was evaluated to determine if further investigation of this approach to reducing complexity was warranted. The utility of using longer training periods was also evaluated.

The HMM was trained using data from Tuesday 16:00 to 19:59 and then used to search for anomalies in data from 12:00 to 15:59 and 20:00 to 23:59. The results were compared to the those obtained by training and testing for the same period

Test	Train 16:00 – 19:59 Test 12:00 – 15:59	Train 12:00 – 15:59 Test 12:00 – 15:59
Reference Test - difference	2675	7127
Anomaly 1 - difference	1276	12356
Anomaly 2 - difference	490	10097
Anomaly 3 - difference	2971	11570

Table 10 Anomaly Detection with Prior Period Test Data

The results documented in Table 10 suggest that there are minimal anomalous data present in the period Tuesday 12:00 – 15:39 and that the model generated specifically for the period as well as the one for the following period were equally effective at identifying the absence of anomalies in the period. The other alternative is that given the differences in PCA results between periods, the variables selected for period 16:00 – 23:59 are inappropriate for the period 12:00 – 15:59.

Test	Train 16:00 – 19:59 Test 20:00 – 23:59	Train 20:00 – 23:59 Test 20:00 – 23:59
Reference Test - difference	3679	12631
Anomaly 1 - difference	16195	10317
Anomaly 2 - difference	48071	15376
Anomaly 3 - difference	37257	13291

Table 11 Anomaly detection with Post Period Test Data

The results documented in Table 11 are quite interesting in that the HMM trained for the prior period clearly detected anomalies while the one trained for the period did not. Clearly this an interesting result but without knowledge of whether anomalies are present in the period 20:00 – 23:59 it is difficult to draw any conclusions. Again, given the differences in PCA results between periods, the difference in the results may be an artifact of the variable selection of the multivariate HMM.

To evaluate the potential of training a multivariate HMM for longer periods and thus reducing the number of models necessary to provide 24 / 7 monitoring additional tests were undertaken. The HMM was trained using data from Tuesday 16:00 to 23:59 and then used to search for anomalies in data from 16:00 to 23:59 and 18:00 to 21:59.

Test	Train 16:00 – 23:59 Test 16:00 – 23:59	Train 16:00 – 23:59 Test 18:00 – 21:59
Reference Test - difference	17315	4643
Anomaly 1 - difference	6532	10050
Anomaly 2 - difference	34922	11568
Anomaly 3 - difference	26980	13291

Table 12 Anomaly Detection using an 8-hour Training Period

The results obtained from an 8-hour training period, as shown in Table 12 do not appear to have any significant ability to detect anomalies. This maybe an inherent limit in the methods used or again may reflect the need to use PCA for the explicit period used for HMM analysis. It would be useful to determine if Univariate HMM analysis is less subject to the data variance.

5.6. *Conclusions*

Both Univariate and Multivariate HMM analysis provide a robust ability to detect anomalies when properly structured and optimized for the data of interest with PCA.

There may be opportunities to reduce the number of models necessary to continuously monitor a system. To determine the feasibility of doing this, additional work with more information about the distribution of anomalies in the test data is needed.

6. Reinforcement Learning

Reinforcement Learning (RL) has a long history, with the underlying theory developed more 100 years ago. Since the 1980's it has been an area of active research and has recently demonstrated the ability to best human experts in solving specific problems such as playing "Go". The core concept of RL is that a system can be constructed that can "learn" from its mistakes and successes. The four key elements of an RL system are: an environment, an agent that interacts with the environment, a subsystem that feeds back the changes to the environment caused by the agent's actions and a reward subsystem that provides positive or negative rewards to the agent based on what its actions did to the environment. Preliminary analysis suggests it is possible to develop an RL system capable of detecting and flagging, in near real-time, malicious actions of an APT threatening the electrical grid.

6.1. *History*

Unlike the canonical search problems, an entity (or agent) of Reinforcement Learning learns from trial-and-error of its own actions rather than being explicitly taught [1]. The consequences of its actions from past experiences and new choices are inextricably tied to a reward system which represents the success or failure of an action's outcome. As such, the objective of an agent is to acquire the optimal strategies that maximize long-term rewards and minimize the amount of penalties [2].

The earliest work attributed to reinforcement learning came from the early 1900s. RL's primary components were explored independently before converged in the 1980s as a defined sub-specialty of machine learning. The two most notable areas that have contributed to modern RL are optimal control problems and learning by trial-and-error [3].

1. At a high-level overview, optimal control is deeply rooted in dynamic programming, dealing with the optimization of cost – a minimization problem. It focuses on minimizing the long-term penalty of the current state through minimizing the sum of its current cost and the cost of its recursive subproblems [4]. In the context of reinforcement learning, this minimization cost problem is transformed into a maximization of reward problem.
2. The notion that centers on learning by trail-and-error is called the Law of Effect. It combines search among many actions in each situation and remembering what actions work the best. In the words of Andrew G. Barto: "The essence of RL is about creating a system that caches (or remembers) the results of many search results."

These two ideas were subsequently unified and paved way to the first RL computation method: temporal difference learning (TD). TD's underline assumption dictates that the future predictions and thus rewards are confirmed or disconfirmed at the current state and in a bit-by-bit manner as new states are made available [5]. In other words, partial information of new states is perceived to be relevant to the final predictions and thus become available to the current state, leading up to the final predictions [6].

Since TD's development, many more RL algorithms such as Sarsa, Q-learning, Policy Gradient, and Dyna have emerged to make up the modern paradigm of reinforcement learning. The most well-known RL utility came from DeepMind's development of AlphaGo. AlphaGo is a computer program that competed and won against top Go masters without learning from human games and rules prior to matches. Its subsequent iterations have since then outcompeted the previous version to the record of 100-0 victory [7]. DeepMind's second notable achievement was the construction of a DL-RL framework to address resource management in Google's data centers. Neural network powered RL recommendation system learns to efficiently allocate and schedule computer resources to waiting jobs thus minimizing the average job slowdown while saving energy consumption [8]. Lastly, one of the most popular applications in RL is autonomous driving as abundant number of companies race toward commercialization of the first fully self-driving car. Using numerous sensor technologies, real-time data are fed into RL platform to train adaptive responses in handling normal and unexpected events.

6.2. *Principles of Operation*

At a high-level, Reinforcement Learning is learning through interaction [9]. The four necessary elements defined in Reinforcement Learning include the agent, environment, action and reward [10]. The goal of Reinforcement Learning is to have the agent interact with its environment where the consequences of the agent's actions will lead to a reward. Based on the reward, the agent will learn whether an action taken is positive or negative which enables the agent to learn over time how to choose most optimal action to take for any given situation. At first glance, the context of Reinforcement Learning may appear to be alike other machine learning approaches where an agent might be exposed to the correct action to take during training scenarios [11]. However, Reinforcement Learning differs in the aspect that the agent will strictly only learn through trial and error.

An agent is defined to be autonomous, usually being a program that utilizes a machine learning algorithm to interact with an environment. Everything that an agent interacts with is defined as the environment. The goal of the agent is to be able to maximize its predictability capabilities to predict anomalies within the environment. This prediction is known as an action where once an action is taken, the agent will transition into a new state. The agent will transition to a new state whenever it takes an action as once an action is taken, it will be exposed to a new set of uncertainties, requiring the agent to take another action. Since the objective is to maximize an agent's predictability capabilities through rewards, the agent aims to develop a sequence of actions that will enable it to achieve this goal. In order to develop the optimal sequence of actions, the agent utilizes a Markov Decision Process so that the agent can take the best action in each state, known as an optimal policy [12].

The Markov Decision Process (MDP) is defined by two key characteristics: The Markov Property and the Markov Chain [13]. The Markov Property states that the conditional property of states to occur in the future will not depend on the current sequence of states, but instead only the current state. In other words, the next state will only be dependent on the current state. The Markov Chain utilizes the Markov Property in a random process where the Markov Chain contains a set of states where each state contains probabilities describing the likelihood of transitioning to each of the next possible states [14]. The MDP is built upon a Markov Chain but also includes a set of possible actions for the agent, a reward function that determines the immediate reward of an action, and a discounted reward function that determines the cumulative rewards amassed by the agent as additional variables to consider. The reward function is one of the most critical components of the MDP. Its goal is to evaluate the reward of the action taken by the agent while also having to consider balancing the importance of immediate and long-term rewards. The balancing of the importance of immediate and long-term rewards is done by the discount function. Since each environment contains numerous uncertainties, the agent will continuously improve its estimation of achieving the optimal policy as the progression of transitioning through states occurs [15]. The expected result is for an agent to be able to maximize the accumulation of rewards once it has been sufficiently trained.

In hindsight, the key strengths of Reinforcement Learning are that it can solve complex problems autonomously without the reliance on humans, it can create an appropriate model for a specific problem, and that it is self-learning where it will correct its own mistakes [16]. However, it holds the assumption that a system is Markov which is not true as in reality, most events are the result of multiple events preceding it. Reinforcement Learning can also be problematic when models are overtrained (have too many states) as it may make the results less credible or inaccurate [17].

6.3. Application to Electrical System Security

Cyber-security has many facets and different threats can best be countered with different techniques. As noted in the **Error! Reference source not found.** one potential risk for the electrical grid is a quiescent APT that is activated to damage the grid infrastructure and/or connected equipment by issuing incorrect control signals to various elements. This could be done as an opening act of overt hostility, a terrorist attack or even a ransomware exploit. In any event to avoid major damage the malicious control signals would need to be detected and countered within seconds.

While approaches like Hidden Markov Models could detect the incorrect control signals, the process is retrospective and probably would not be able raise an alert before the damage had been done. While this would provide useful information for a postmortem, it would be much better to be able to intercede before the damage occurred. Reinforcement Learning has the potential to provide close to real-time detection of anomalous control signals and could alert operators to intercede quickly enough to prevent damage.

A SCADA System can be modeled as a black box that for a given set inputs states will produce a limited set of responses (output signals) with varying probabilities i.e., a Markov Model. An attractive aspect of this generalization is that there is a wealth of training data available for machine learning. It is safe to assume that if the Electrical Grid is operating normally, the sets of input parameters and the resulting output control signals define how the system should behave. Thus, there are thousands of hours of observations covering virtually every state possible. The observations, the input parameters, define a detailed set of environment states. Comparison of the actual control signals with the ones produced by the model can be used to generate rewards. The maximum positive reward would be awarded for reproducing the actual control signals. Development work will need to be done to define exactly how to penalize non-optimum control outputs but there are no obvious issues with designing and tuning a solution to provide “reward” feedback to the model as it learns to emulate the SCADA system. Figure 7 is a conceptual diagram of how a Reinforcement Learning Model would be trained to reproduce the SCADA control outputs.

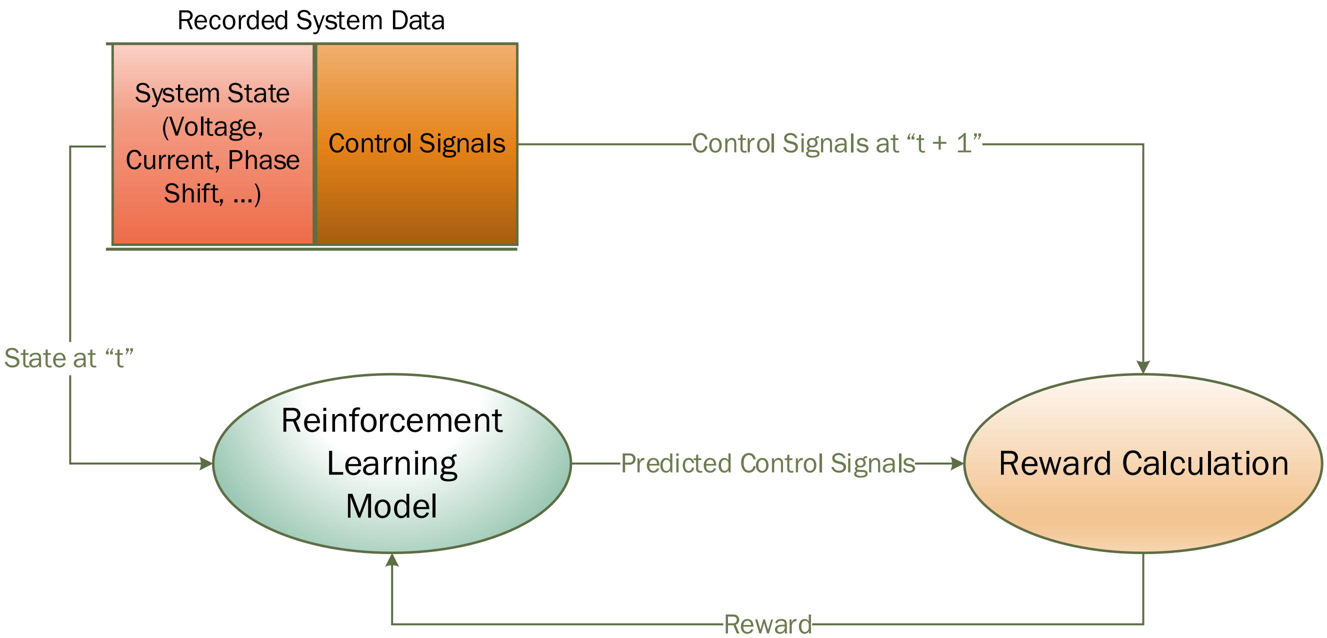


Figure 7 RL Model Training Configuration

In operation the RL security system model would behave much as it does during training except that the calculated "reward" would be used to identify anomalies. The system would be weighing the question "What has the system done in the past under similar circumstances?" and then evaluating "What is the system being commanded to do under these circumstances now?" If there are significant differences, as indicated by a large negative reward, then it is cause for alarm and an indication that the SCADA system has been compromised and is taking inappropriate and quite possibly destructive actions. Figure 8 provides a conceptual diagram of how the trained Reinforcement Learning Model could be used for detecting anomalous control actions being taken by the SCADA System.

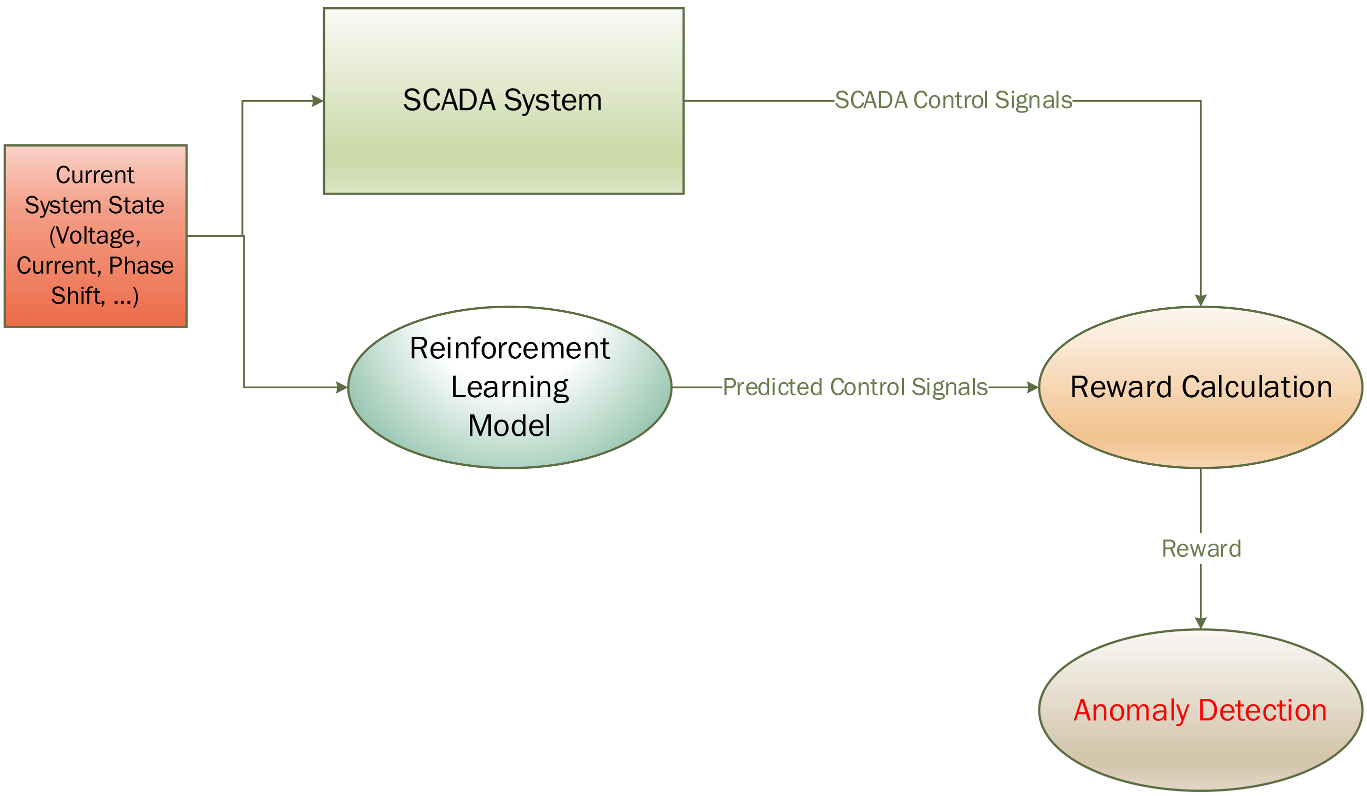


Figure 8 RL Model Anomaly Detection Configuration

Developing the reward calculation subsystem will be the major technical challenge of implementing this type of solution as the outputs of the SCADA system are complex. However, the output is deterministic so it should be possible to assign weights to different responses to provide meaningful feedback to the agent about its performance.

7. References

1. D. Silver, *Deep Reinforcement Learning*, 17-Jun-2016. [Online]. Available: https://deepmind.com/blog/article/deep-reinforcement-learning. [Accessed: 29-Nov-2020].

2. F. Woergoetter and B. Porr, “Reinforcement learning,” *Scholarpedia*, vol. 3, no. 3, p. 1448, 2008.

3. R. S. Sutton and A. G. Barto, “1.6 History of Reinforcement Learning,” in *Reinforcement learning: an introduction*, Cambridge, MA: The MIT Press, 2012, pp. 16–23.

4. Todorov E (2006). In *Bayesian Brain: Probabilistic Approaches to Neural Coding*, Doya Kat al (eds), chap 12, pp 269-298, MIT Press

5. Sutton, R. (1988). Learning to predict by the methods of temporal differences. Machine learning, 3(1):9–34.

6. J. L. McClelland and D. E. Rumelhart, “Chapter 9 Temporal-Difference Learning,” in *Explorations in parallel distributed processing: a handbook of models, programs, and exercises*, Cambridge, MA: MIT Press, 1991, pp. 184–184.
7. “AlphaGo,” *Wikipedia*, 21-Nov-2020. [Online]. Available: https://en.wikipedia.org/wiki/AlphaGo. [Accessed: 29-Nov-2020].
8. C. Gamble and J. Gao, *Safety-first AI for autonomous data centre cooling and industrial control*, 17-Aug-2018. [Online]. Available: https://deepmind.com/blog/article/safety-first-ai-autonomous-data-centre-cooling-and-industrial-control. [Accessed: 29-Nov-2020].
9. *Reinforcement Learning: an Introduction*, by Richard S. Sutton and Andrew G. Barto, The MIT Press, 2018, pp. 37–39.

10. Lee, Dan. “Reinforcement Learning, Part 3: The Markov Decision Process.” *Medium*, AI³ | Theory, Practice, Business, 9 Nov. 2019, medium.com/ai%C2%B3-theory-practice-business/reinforcement-learning-part-3-the-markov-decision-process-9f5066e073a2.

11. Heidenreich, Author Hunter. “Reinforcement Learning Tutorial With Mario.” *SAP Conversational AI Blog*, SAP Conversational AI, 31 Aug. 2020, cai.tools.sap/blog/the-future-with-reinforcement-learning-part-1/.

12. Arulkumaran, Kai, et al. “A Brief Survey of Deep Reinforcement Learning.” *ArXiv.org*, Cornell University, 28 Sept. 2017, arxiv.org/abs/1708.05866.

13. Lee, Dan. “Reinforcement Learning, Part 2: Introducing Markov Process.” *Medium*, AI³ | Theory, Practice, Business, 25 Oct. 2019, medium.com/ai%C2%B3-theory-practice-business/reinforcement-learning-part-2-introducing-markov-process-d3586d4003e0.

14. Ashraf, Mohammad. “Reinforcement Learning Demystified: Markov Decision Processes (Part 1).” *Medium*, Towards Data Science, 11 Apr. 2018, towardsdatascience.com/reinforcement-learning-demystified-markov-decision-processes-part-1-bf00dda41690.

15. Koduvely, Dr Hari. “Anomaly Detection through Reinforcement Learning.” *Zighra*, Zighra, 7 Feb. 2019, zighra.com/blogs/anomaly-detection-through-reinforcement-learning/.

16. Bajaj, Prateek. “Reinforcement Learning.” *GeeksforGeeks*, GeekforGeeks, 17 May 2020, www.geeksforgeeks.org/what-is-reinforcement-learning/.

17. Joy, Ashwin. “Pros And Cons Of Reinforcement Learning.” *Pythonista Planet*, Pythonista Planet, 11 June 2020, www.pythonistaplanet.com/pros-and-cons-of-reinforcement-learning/.

A. Appendix A – PCA Results for Complete Dataset

```
day  1  period  1
Importance of components:
                          PC1    PC2     PC3     PC4     PC5     PC6     PC7
Standard deviation     2.0989 1.1517 0.77503 0.59243 0.49077 0.25187 0.10987
Proportion of Variance 0.6294 0.1895 0.08581 0.05014 0.03441 0.00906 0.00172
Cumulative Proportion  0.6294 0.8189 0.90467 0.95480 0.98921 0.99828 1.00000
                             PC1        PC2         PC3         PC4         PC5         PC6         PC7
Global_intensity      -0.4659942  0.1328735 -0.04290936  0.07331237 -0.17423125 -0.12622527  0.84361647
Global_active_power   -0.4471846  0.1836098 -0.04712297  0.28020488 -0.15628675  0.78341643 -0.21774103
Global_reactive_power -0.3216018 -0.2386486  0.84630642 -0.31615534 -0.11834744 -0.01396779 -0.09606837
Voltage               -0.0231821 -0.8222240 -0.19497103  0.29460691 -0.44434740 -0.03048312 -0.01515175
Sub_metering_1        -0.3816144 -0.1016777 -0.48631614 -0.75672411 -0.07342399 -0.01480710 -0.17113449
Sub_metering_2        -0.3868817 -0.3470020 -0.05645967  0.20286419  0.82594962 -0.05549345 -0.01727183
Sub_metering_3        -0.4284894  0.2913382 -0.04470878  0.33996741 -0.21490310 -0.60489899 -0.44928394

 day  1  period  2
Importance of components:
                         PC1     PC2     PC3     PC4    PC5     PC6     PC7
Standard deviation     2.355 0.82606 0.71623 0.45956 0.1962 0.09347 0.04923
Proportion of Variance 0.792 0.09748 0.07328 0.03017 0.0055 0.00125 0.00035
Cumulative Proportion  0.792 0.88945 0.96273 0.99291 0.9984 0.99965 1.00000
                             PC1          PC2          PC3        PC4         PC5          PC6         PC7
Global_intensity       0.4138348 -0.138012454  0.036694412 -0.3489893  0.44062559  0.628987114 -0.31109104
Global_active_power    0.4202095 -0.001337518  0.034436321 -0.2555795  0.19035614 -0.750330637 -0.39709439
Global_reactive_power  0.3165942 -0.647500990  0.445128655  0.5146218 -0.13188951 -0.003946606 -0.01118367
Voltage               -0.4131870 -0.101013903  0.087680409  0.2677540  0.84172423 -0.152238490  0.08953426
Sub_metering_1         0.2985891  0.741825615  0.434635093  0.3951846  0.08349793  0.087572779 -0.02863376
Sub_metering_2         0.3378798  0.026188233 -0.776346599  0.5199557  0.09453363  0.026567774 -0.04940600
Sub_metering_3         0.4217715 -0.022324106 -0.004310604 -0.2232329  0.16676511 -0.099056418  0.85681743

 day  1  period  3
Importance of components:
                          PC1    PC2     PC3     PC4     PC5     PC6     PC7
Standard deviation     2.1217 1.0571 0.77883 0.66551 0.50169 0.24946 0.13251
Proportion of Variance 0.6431 0.1596 0.08665 0.06327 0.03596 0.00889 0.00251
Cumulative Proportion  0.6431 0.8027 0.88937 0.95265 0.98860 0.99749 1.00000
                              PC1         PC2          PC3         PC4         PC5         PC6          PC7
Global_intensity      -0.45526972  0.08124908 -0.102595293  0.15341549 -0.32616377 -0.32256459  0.735958724
Global_active_power   -0.45630197  0.05271851  0.019914158  0.15615015 -0.33150819 -0.45909811 -0.666003805
Global_reactive_power -0.06151252  0.89841645  0.128699098 -0.39527917 -0.06469075  0.10836789 -0.018069705
Voltage                0.35096335  0.30306356  0.274949688  0.79771012 -0.23704120  0.12959387  0.007439205
Sub_metering_1        -0.37756929  0.24258121 -0.426847459  0.39870863  0.66286498  0.12225232 -0.055613409
Sub_metering_2         0.33845896  0.10375301 -0.845405537 -0.01693545 -0.39102781  0.02785542 -0.077490922
Sub_metering_3        -0.44771184 -0.14835555  0.002751166  0.01543960 -0.36201294  0.80057559 -0.072966691

 day  1  period  4
Importance of components:
                          PC1    PC2    PC3     PC4     PC5     PC6     PC7
Standard deviation     2.0209 1.1680 0.9487 0.72670 0.27909 0.18882 0.10115
Proportion of Variance 0.5834 0.1949 0.1286 0.07544 0.01113 0.00509 0.00146
Cumulative Proportion  0.5834 0.7783 0.9069 0.98232 0.99344 0.99854 1.00000
                             PC1         PC2          PC3        PC4          PC5         PC6         PC7
Global_intensity      -0.4753291 -0.20513537 -0.006153247  0.1009940 -0.009752373  0.52761585 -0.66578562
Global_active_power   -0.4824152  0.01858457 -0.013192063  0.1284724  0.576726393 -0.62894008 -0.14856650
Global_reactive_power -0.2382698  0.06813631 -0.880418013 -0.3395007 -0.201752323 -0.06468290  0.05744961
Voltage                0.4597494 -0.08171815 -0.267135901 -0.1700903  0.754345579  0.32617418 -0.07895234
Sub_metering_1        -0.1059445 -0.71166649  0.236239759 -0.6341801 -0.012391159 -0.11126453  0.10853400
Sub_metering_2         0.1533435 -0.66293524 -0.308347484  0.6423525 -0.075278792 -0.09677982  0.11947614
Sub_metering_3        -0.4883209  0.01735948  0.049022897  0.1195933  0.227442014  0.44015485  0.70644844
```

```
 day  1  period  5
Importance of components:
                          PC1    PC2    PC3     PC4     PC5     PC6    PC7
Standard deviation     2.3093 0.9872 0.6300 0.43845 0.28966 0.12883 0.0528
Proportion of Variance 0.7619 0.1392 0.0567 0.02746 0.01199 0.00237 0.0004
Cumulative Proportion  0.7619 0.9011 0.9578 0.98524 0.99723 0.99960 1.0000
                             PC1         PC2         PC3         PC4         PC5         PC6        PC7
Global_intensity       0.4322289 -0.01214079 -0.05440337  0.01090444  0.03235398 -0.00557077  0.8993744
Global_active_power    0.4273975 -0.04138130 -0.10303363 -0.20265391  0.01718280 -0.84685056 -0.2155999
Global_reactive_power  0.4135520  0.08657003  0.24695875 -0.09018633  0.79900724  0.26535236 -0.2086474
Voltage               -0.3804217 -0.08539867  0.59489394 -0.64411564  0.06122269 -0.16140686  0.2222661
Sub_metering_1         0.3603701 -0.27341979  0.70043167  0.38037504 -0.37752238  0.04322550 -0.1252745
Sub_metering_2         0.1773339  0.91202940  0.16682879 -0.13248073 -0.29126864  0.06324552 -0.0503454
Sub_metering_3         0.3917266 -0.27712249 -0.23043556 -0.61119362 -0.35935583  0.42483009 -0.1829700

 day  1  period  6
Importance of components:
                          PC1    PC2     PC3     PC4     PC5     PC6     PC7
Standard deviation     2.2788 1.0153 0.73699 0.39913 0.19753 0.17722 0.05655
Proportion of Variance 0.7419 0.1472 0.07759 0.02276 0.00557 0.00449 0.00046
Cumulative Proportion  0.7419 0.8891 0.96672 0.98948 0.99506 0.99954 1.00000
                             PC1         PC2          PC3        PC4         PC5         PC6          PC7
Global_intensity      -0.4356095  0.08447896 -0.006742087  0.1533102 -0.03372615  0.17970594 -0.863786178
Global_active_power   -0.4114518  0.26982548  0.017314523  0.4307709  0.51865463  0.40393742  0.373992332
Global_reactive_power -0.2995679 -0.50270938 -0.681463144 -0.3383115  0.27600553  0.02730390  0.042084740
Voltage                0.4198352 -0.17325310  0.181135778 -0.3097080  0.14862123  0.79183250 -0.126117384
Sub_metering_1        -0.3311398  0.57482179  0.079434256 -0.7366627 -0.03573703  0.01717047  0.096812676
Sub_metering_2        -0.3052958 -0.48231206  0.698909598 -0.1832792  0.31290594 -0.23265099  0.008187191
Sub_metering_3        -0.4148376 -0.27254201  0.087567899  0.0655410 -0.72966111  0.34982113  0.294765550

 day  2  period  1
Importance of components:
                          PC1    PC2    PC3     PC4     PC5     PC6    PC7
Standard deviation     2.0619 1.0808 0.9126 0.65555 0.45525 0.30878 0.1242
Proportion of Variance 0.6073 0.1669 0.1190 0.06139 0.02961 0.01362 0.0022
Cumulative Proportion  0.6073 0.7742 0.8932 0.95457 0.98418 0.99780 1.0000
                              PC1         PC2         PC3         PC4          PC5         PC6         PC7
Global_intensity      -0.47716646  0.04850597 -0.03411698  0.18794254 -0.090374961  0.05805803  0.84966735
Global_active_power   -0.46289496  0.12439603 -0.09137832  0.01877552 -0.001163186  0.80778510 -0.33020182
Global_reactive_power -0.23066137  0.51869208  0.67993484 -0.38521603 -0.202835402 -0.15016780 -0.05795271
Voltage                0.03350958  0.76722761 -0.60014648 -0.04952736  0.135736631 -0.17041336 -0.01204154
Sub_metering_1        -0.34172582 -0.33835594 -0.37670143 -0.75105848 -0.171624872 -0.18037040 -0.02751964
Sub_metering_2        -0.44126295 -0.08962657  0.12506646  0.06424685  0.821895134 -0.28450312 -0.14502129
Sub_metering_3        -0.43843282 -0.04444607 -0.10224121  0.49524159 -0.476757456 -0.42313816 -0.37913076

 day  2  period  2
Importance of components:
                         PC1    PC2     PC3     PC4     PC5     PC6     PC7
Standard deviation     2.350 0.9026 0.61570 0.48309 0.19203 0.10100 0.05070
Proportion of Variance 0.789 0.1164 0.05416 0.03334 0.00527 0.00146 0.00037
Cumulative Proportion  0.789 0.9054 0.95957 0.99291 0.99818 0.99963 1.00000
                             PC1        PC2        PC3         PC4         PC5         PC6          PC7
Global_intensity       0.4083444 -0.2370785  0.2094760  0.01120266 -0.58576599 -0.61994488 -0.074763070
Global_active_power    0.4139258 -0.1646454  0.2644683  0.01466416 -0.18116154  0.51723176  0.656544445
Global_reactive_power  0.3282429 -0.3954959 -0.8551358 -0.03008611  0.02298730  0.05610278  0.001156066
Voltage               -0.4145499  0.1106890 -0.2051600 -0.05857840 -0.78374930  0.37112301 -0.135569163
Sub_metering_1         0.3290398  0.5752798 -0.1166691 -0.73822810 -0.02631002 -0.03351393  0.019447216
Sub_metering_2         0.3159020  0.6349597 -0.1990580  0.67092616 -0.08345350 -0.01248274  0.012073892
Sub_metering_3         0.4174494 -0.1178815  0.2441162  0.01415011  0.04013243  0.45385111 -0.737872325

 day  2  period  3
Importance of components:
                          PC1    PC2    PC3     PC4     PC5     PC6     PC7
Standard deviation     2.0314 1.2024 0.8373 0.67836 0.43870 0.22068 0.15923
Proportion of Variance 0.5895 0.2065 0.1001 0.06574 0.02749 0.00696 0.00362
Cumulative Proportion  0.5895 0.7961 0.8962 0.96193 0.98942 0.99638 1.00000
                             PC1         PC2         PC3         PC4         PC5          PC6         PC7
Global_intensity      -0.4484102  0.27083932 -0.22153709  0.04306458  0.22081632 -0.398646432 -0.68334640
Global_active_power   -0.4562763  0.17537340 -0.13812295  0.07317663  0.56736000  0.590391677  0.25722302
Global_reactive_power  0.3203125  0.21945828 -0.72040367  0.55326269 -0.10112717 -0.008849108  0.11769408
Voltage                0.1501388 -0.67434304 -0.47734317 -0.36139041  0.38893301 -0.103490437 -0.04776117
Sub_metering_1        -0.4047106  0.09652161 -0.42535432 -0.49404424 -0.60336455  0.142340221  0.13257986
Sub_metering_2         0.2990337  0.58227420 -0.05347446 -0.48982318  0.32165110 -0.344852811  0.32613871
Sub_metering_3        -0.4608544 -0.21125837  0.03492295  0.26848003  0.01841135 -0.585274220  0.57166172
```

```
 day  2  period  4
Importance of components:
                          PC1    PC2    PC3     PC4     PC5     PC6     PC7
Standard deviation     2.1546 0.9461 0.8530 0.70087 0.40373 0.26359 0.10617
Proportion of Variance 0.6632 0.1279 0.1039 0.07017 0.02329 0.00993 0.00161
Cumulative Proportion  0.6632 0.7911 0.8950 0.96518 0.98846 0.99839 1.00000
                             PC1         PC2         PC3         PC4          PC5         PC6         PC7
Global_intensity      -0.4429465 -0.10841416  0.08723045 -0.33239055  0.245890536 -0.18743252  0.76049921
Global_active_power   -0.4314494 -0.15485247 -0.05677522  0.20686762 -0.653228645 -0.55389031 -0.10174651
Global_reactive_power -0.2909505 -0.03712081 -0.89965992  0.08090903  0.286689675  0.04306812 -0.11827847
Voltage                0.4411499  0.12085379 -0.08749983 -0.18468295  0.320790063 -0.79794900 -0.09689291
Sub_metering_1         0.2061596 -0.91720454 -0.05118706 -0.31018309 -0.037964953  0.05122412 -0.11547791
Sub_metering_2        -0.3847503  0.21768482  0.12322468 -0.72216217  0.009315478  0.04027985 -0.51591548
Sub_metering_3        -0.3845125 -0.24419796  0.39283430  0.43397753  0.571143442 -0.12349387 -0.32925163

 day  2  period  5
Importance of components:
                          PC1    PC2     PC3     PC4     PC5     PC6     PC7
Standard deviation     2.1540 1.3080 0.66971 0.32734 0.25659 0.14365 0.08523
Proportion of Variance 0.6628 0.2444 0.06407 0.01531 0.00941 0.00295 0.00104
Cumulative Proportion  0.6628 0.9072 0.97130 0.98661 0.99601 0.99896 1.00000
                              PC1         PC2         PC3         PC4          PC5         PC6          PC7
Global_intensity       0.46176562 -0.04272871  0.04772022 -0.03848349  0.045838181  0.20365057 -0.858844502
Global_active_power    0.45957245 -0.05601527  0.01587521 -0.03902880  0.132260967  0.75732205  0.439147508
Global_reactive_power  0.44994420 -0.02912348 -0.20142410  0.23307063  0.681667093 -0.46389793  0.148111908
Voltage               -0.37411011  0.39612989 -0.25788937  0.65379115  0.239719069  0.35120972 -0.168402612
Sub_metering_1         0.43990053  0.08528478 -0.28205847  0.46398942 -0.676875422 -0.18039645  0.116908740
Sub_metering_2         0.17468497  0.58852497  0.76696978  0.12719546 -0.001300307 -0.11483101  0.074258652
Sub_metering_3         0.09555209  0.69543824 -0.47179715 -0.53269703 -0.010421763 -0.02666761  0.007550473

 day  2  period  6
Importance of components:
                          PC1    PC2     PC3     PC4     PC5     PC6     PC7
Standard deviation     2.3443 0.8714 0.68391 0.44277 0.23249 0.15056 0.06659
Proportion of Variance 0.7851 0.1085 0.06682 0.02801 0.00772 0.00324 0.00063
Cumulative Proportion  0.7851 0.8936 0.96040 0.98841 0.99613 0.99937 1.00000
                             PC1         PC2         PC3         PC4         PC5          PC6         PC7
Global_intensity      -0.4242865  0.04832109 -0.05892871 -0.13531853  0.05714418  0.001669879 -0.89027754
Global_active_power   -0.4136141  0.18309262 -0.07477454 -0.13434706 -0.60253176 -0.609604454  0.19260844
Global_reactive_power -0.2472877 -0.89138515  0.35482061  0.07675575 -0.08219169 -0.070006420  0.02891103
Voltage                0.4058991  0.02677583  0.12776508  0.57964814 -0.60647312  0.085242756 -0.32731805
Sub_metering_1        -0.3276362  0.40161395  0.75015378  0.33232829  0.22257918 -0.013247921  0.09203787
Sub_metering_2        -0.3728698 -0.06077824 -0.53064763  0.70755594  0.24087707 -0.057117811  0.11733537
Sub_metering_3        -0.4202639  0.06220917 -0.06630982 -0.10333852 -0.38930919  0.782799108  0.20024080

 day  3  period  1
Importance of components:
                          PC1    PC2    PC3     PC4     PC5     PC6     PC7
Standard deviation     1.8773 1.1677 0.9677 0.77049 0.65593 0.34309 0.18469
Proportion of Variance 0.5034 0.1948 0.1338 0.08481 0.06146 0.01682 0.00487
Cumulative Proportion  0.5034 0.6983 0.8320 0.91685 0.97831 0.99513 1.00000
                              PC1         PC2         PC3         PC4          PC5          PC6         PC7
Global_intensity      -0.51684112  0.03846489 -0.08969311  0.20878838 -0.002109258  0.001121639  0.82447108
Global_active_power   -0.48553513  0.12701591  0.04069059  0.38184505 -0.077900146 -0.658097284 -0.40187157
Global_reactive_power -0.38652538  0.22649909  0.25108668 -0.45377748  0.717320459  0.062704555 -0.10889112
Voltage               -0.08841662  0.63213603 -0.52951756 -0.43943746 -0.341209120 -0.040509298 -0.03205848
Sub_metering_1        -0.08194932 -0.53928107 -0.74936567 -0.09309413  0.328461142 -0.132082744 -0.08313985
Sub_metering_2        -0.29095864 -0.48397812  0.27884456 -0.60007666 -0.463223380 -0.167996098  0.02152607
Sub_metering_3        -0.49853676 -0.08019182 -0.08696464  0.20693954 -0.201226717  0.718098393 -0.37213715

 day  3  period  2
Importance of components:
                          PC1    PC2    PC3    PC4     PC5     PC6     PC7
Standard deviation     2.1119 1.1117 0.9002 0.6561 0.22364 0.09723 0.06174
Proportion of Variance 0.6371 0.1766 0.1158 0.0615 0.00715 0.00135 0.00054
Cumulative Proportion  0.6371 0.8137 0.9295 0.9910 0.99810 0.99946 1.00000
                             PC1        PC2         PC3         PC4         PC5         PC6           PC7
Global_intensity       0.4611197  0.1588889 -0.02999043  0.07603646  0.50625430 -0.66911316  0.2267952136
Global_active_power    0.4636430  0.1421934 -0.12534892  0.02125930  0.09016866  0.14185418 -0.8487631630
Global_reactive_power  0.2325706  0.2074189  0.92509517 -0.16101233 -0.13436738  0.05346102  0.0157977975
Voltage               -0.4484145 -0.1132766  0.25616534  0.05942707  0.80150062  0.20443219 -0.1809554072
Sub_metering_1         0.2259792 -0.6749359 -0.01273627 -0.69628255  0.09072164 -0.01371966  0.0021564300
Sub_metering_2         0.2318986 -0.6505086  0.19547780  0.69157290 -0.07971457  0.01491575  0.0001740117
Sub_metering_3         0.4608639  0.1413547 -0.15359910  0.03503551  0.24603563  0.69792812  0.4417758639
```

```
 day  3  period  3
Importance of components:
                          PC1    PC2    PC3     PC4     PC5     PC6     PC7
Standard deviation     1.7816 1.3536 1.0382 0.72110 0.47317 0.37682 0.17322
Proportion of Variance 0.4534 0.2617 0.1540 0.07428 0.03198 0.02028 0.00429
Cumulative Proportion  0.4534 0.7152 0.8692 0.94345 0.97543 0.99571 1.00000
                             PC1          PC2         PC3         PC4         PC5        PC6        PC7
Global_intensity      -0.1974069  0.666185797 -0.09957346  0.25962147 -0.02647732 -0.1917234  0.6343894
Global_active_power   -0.3911676  0.460950724 -0.25559612  0.07959211  0.08843072  0.5422179 -0.5109088
Global_reactive_power  0.4593009  0.323253337 -0.07253214  0.04471267 -0.75166842 -0.1458715 -0.3016718
Voltage                0.2114702  0.008017885 -0.82340367 -0.47878368  0.16439595 -0.1056269  0.0990232
Sub_metering_1         0.2244998  0.482815164  0.45449236 -0.53913376  0.38212625 -0.1946220 -0.1880495
Sub_metering_2         0.4674330  0.064665408 -0.15290150  0.63543472  0.50210089 -0.1955000 -0.2446285
Sub_metering_3        -0.5323980 -0.043264426 -0.10751306  0.02421281 -0.03635591 -0.7488047 -0.3748400

 day  3  period  4
Importance of components:
                          PC1    PC2    PC3    PC4     PC5     PC6     PC7
Standard deviation     2.0012 1.1101 1.0478 0.6985 0.33544 0.23683 0.09221
Proportion of Variance 0.5721 0.1760 0.1568 0.0697 0.01607 0.00801 0.00121
Cumulative Proportion  0.5721 0.7481 0.9050 0.9747 0.99077 0.99879 1.00000
                              PC1         PC2         PC3         PC4         PC5         PC6         PC7
Global_intensity      -0.43176630 -0.42396343  0.03620874 -0.09413355  0.42070897  0.22709717  0.62856306
Global_active_power   -0.45486384 -0.23417326  0.11456151 -0.14731556 -0.83651213  0.03847787  0.04693037
Global_reactive_power -0.28953481  0.31806402 -0.50539674 -0.72629257  0.11805556 -0.03925873 -0.12884000
Voltage                0.47708164 -0.04968267 -0.14996925 -0.20054998 -0.18746658  0.81311013  0.10450776
Sub_metering_1        -0.02632768 -0.43003042 -0.79419541  0.39504161 -0.05437198 -0.05344835 -0.14752434
Sub_metering_2         0.25297582 -0.69020360  0.21707167 -0.45215697  0.13446167 -0.18007197 -0.39692609
Sub_metering_3        -0.48015574  0.02185470  0.17257313  0.20300740  0.23050885  0.49897331 -0.62918253

 day  3  period  5
Importance of components:
                          PC1     PC2     PC3     PC4     PC5     PC6     PC7
Standard deviation     2.3617 0.74209 0.68773 0.46468 0.37537 0.18435 0.08813
Proportion of Variance 0.7968 0.07867 0.06757 0.03085 0.02013 0.00485 0.00111
Cumulative Proportion  0.7968 0.87549 0.94306 0.97391 0.99404 0.99889 1.00000
                             PC1         PC2          PC3         PC4         PC5          PC6         PC7
Global_intensity       0.4107471 -0.28776641 -0.110491498  0.06345084 -0.08709936 -0.024076236  0.85092711
Global_active_power    0.4164366 -0.05488095 -0.088661449 -0.05525413 -0.08818477 -0.858562105 -0.26028697
Global_reactive_power  0.3692677  0.24287489 -0.541067696 -0.10501110  0.67876977  0.181938922 -0.08391310
Voltage               -0.3540649  0.71197020 -0.004380967 -0.27476619  0.03336037 -0.331575399  0.42563564
Sub_metering_1         0.3683351  0.50126615 -0.260496171  0.25692141 -0.63140627  0.257789300 -0.11859771
Sub_metering_2        -0.3439932 -0.30978139 -0.699644221 -0.43208126 -0.32942789 -0.003705583 -0.03116770
Sub_metering_3         0.3770208  0.03250436  0.360273564 -0.80851367 -0.12489682  0.229735448 -0.07021257

 day  3  period  6
Importance of components:
                          PC1    PC2     PC3     PC4     PC5    PC6     PC7
Standard deviation     2.3467 0.9229 0.65411 0.38452 0.20801 0.1212 0.08723
Proportion of Variance 0.7867 0.1217 0.06112 0.02112 0.00618 0.0021 0.00109
Cumulative Proportion  0.7867 0.9084 0.96951 0.99063 0.99682 0.9989 1.00000
                             PC1         PC2         PC3         PC4         PC5         PC6         PC7
Global_intensity      -0.4220922  0.01730724 -0.11002741  0.22117545 -0.05391990 -0.20439671  0.84606660
Global_active_power   -0.4173343  0.08050496 -0.09437109  0.20823688  0.71442692 -0.38870861 -0.32493433
Global_reactive_power -0.2418589 -0.87452449 -0.12458015 -0.35801820  0.12371299  0.13228721  0.01446238
Voltage                0.4149204  0.08333160  0.20652745 -0.17691039  0.68239788  0.34277980  0.40469889
Sub_metering_1        -0.3371952 -0.06623259  0.92864218  0.07470044 -0.06341491  0.08699060 -0.04865540
Sub_metering_2        -0.3661271  0.43560607 -0.03585731 -0.82066642 -0.03168095 -0.01620043  0.01237246
Sub_metering_3        -0.4114195  0.16535650 -0.23906532  0.26300377  0.02626631  0.81504594 -0.10990096

 day  4  period  1
Importance of components:
                          PC1    PC2     PC3     PC4     PC5    PC6     PC7
Standard deviation     2.2033 1.0473 0.67354 0.53929 0.44212 0.3028 0.13063
Proportion of Variance 0.6935 0.1567 0.06481 0.04155 0.02792 0.0131 0.00244
Cumulative Proportion  0.6935 0.8502 0.91499 0.95654 0.98446 0.9976 1.00000
                             PC1        PC2          PC3        PC4         PC5         PC6         PC7
Global_intensity      -0.4381044  0.1898737 -0.094525812  0.1867381 -0.08743415  0.09854040 -0.84312000
Global_active_power   -0.4289257  0.1278088 -0.004123812  0.3112934  0.02565348 -0.80741134  0.22404436
Global_reactive_power -0.3812815 -0.1779442 -0.448737965 -0.7127542 -0.31702376 -0.08791378  0.07309610
Voltage               -0.1428677 -0.8759696  0.199144003  0.2912149 -0.29130306  0.02602790 -0.04761048
Sub_metering_1        -0.3546549  0.2002290  0.827219196 -0.3073627 -0.17569211  0.11975502  0.10077616
Sub_metering_2        -0.4007163 -0.2553682 -0.003111978 -0.1148060  0.85349813  0.17117447  0.05712842
Sub_metering_3        -0.4146648  0.2081382 -0.256363974  0.4095709 -0.21681526  0.53509669  0.46682277
```

```
 day  4  period  2
Importance of components:
                          PC1    PC2     PC3     PC4     PC5     PC6     PC7
Standard deviation     2.3027 1.0351 0.63365 0.41987 0.18908 0.09340 0.06154
Proportion of Variance 0.7575 0.1530 0.05736 0.02518 0.00511 0.00125 0.00054
Cumulative Proportion  0.7575 0.9106 0.96792 0.99311 0.99821 0.99946 1.00000
                             PC1         PC2          PC3        PC4          PC5         PC6          PC7
Global_intensity       0.4178913 -0.21805478  0.0678240862 -0.2294212 -0.488386039 -0.56440436  0.404365627
Global_active_power    0.4240602 -0.17007880  0.0007085467 -0.2467706 -0.243754064  0.10941887 -0.811764384
Global_reactive_power  0.3891402 -0.04850847  0.4510786267  0.7997953 -0.025311933  0.04685442 -0.015374631
Voltage               -0.4253998  0.09923571  0.0373499669  0.1474632 -0.829932221  0.31184543  0.003430785
Sub_metering_1         0.3207615  0.42473345 -0.7794412308  0.3085268 -0.109815495 -0.04251924  0.011348099
Sub_metering_2         0.1665851  0.84595350  0.4272182289 -0.2699785 -0.022735529 -0.02591642 -0.004441022
Sub_metering_3         0.4272074 -0.12364745 -0.0220996928 -0.2386946  0.007648574  0.75336099  0.420868592

 day  4  period  3
Importance of components:
                          PC1    PC2    PC3    PC4     PC5     PC6     PC7
Standard deviation     2.0402 1.0634 0.8820 0.7848 0.43519 0.33918 0.09277
Proportion of Variance 0.5946 0.1615 0.1111 0.0880 0.02706 0.01644 0.00123
Cumulative Proportion  0.5946 0.7562 0.8673 0.9553 0.98234 0.99877 1.00000
                             PC1         PC2         PC3         PC4         PC5          PC6          PC7
Global_intensity      -0.4674357  0.18773480  0.11535643 -0.15585500  0.25607032 -0.264398592 -0.757088481
Global_active_power   -0.4672679  0.07561091  0.19568110 -0.08948353  0.34086535 -0.462320887  0.632230615
Global_reactive_power  0.2111076  0.58187164  0.72553887 -0.15339469 -0.05520993  0.247544812  0.050949271
Voltage                0.3139841 -0.35119802  0.03787296 -0.84140720  0.26151467 -0.014587564 -0.008413792
Sub_metering_1        -0.4326077 -0.05877875  0.02460829 -0.38788175 -0.80900711  0.001419401  0.061994724
Sub_metering_2         0.1546008  0.70077244 -0.61963589 -0.26665139 -0.01834468 -0.149741636  0.084887482
Sub_metering_3        -0.4571229  0.05086284 -0.18975400 -0.12104008  0.30326839  0.795259064  0.115696287

 day  4  period  4
Importance of components:
                          PC1    PC2    PC3     PC4     PC5     PC6     PC7
Standard deviation     2.1318 1.0155 0.8976 0.70739 0.24303 0.21260 0.11853
Proportion of Variance 0.6492 0.1473 0.1151 0.07149 0.00844 0.00646 0.00201
Cumulative Proportion  0.6492 0.7965 0.9116 0.98310 0.99154 0.99799 1.00000
                             PC1         PC2         PC3         PC4         PC5         PC6         PC7
Global_intensity      -0.4485125  0.04458998 -0.25462751  0.02944153 -0.56136173  0.46193475 -0.45015059
Global_active_power   -0.4427434 -0.24860387 -0.11893159  0.05642956 -0.16663043 -0.82170090 -0.14794359
Global_reactive_power -0.2495193  0.65231925  0.29687276  0.64276473  0.06143851 -0.07945521  0.02918822
Voltage                0.4430565  0.22895259  0.11852883 -0.07912650 -0.78386762 -0.21879918  0.26201145
Sub_metering_1         0.2512147  0.17157838 -0.88686482  0.31094083  0.08355336 -0.07403363  0.10851955
Sub_metering_2        -0.2757326  0.61769888 -0.16537495 -0.68855477  0.13457129 -0.10799971  0.10578150
Sub_metering_3        -0.4529001 -0.21743300 -0.07042369  0.07582563 -0.11744605  0.20023186  0.82644357

 day  4  period  5
Importance of components:
                          PC1     PC2     PC3     PC4     PC5     PC6     PC7
Standard deviation     2.4001 0.78834 0.64978 0.32633 0.25825 0.14018 0.05435
Proportion of Variance 0.8229 0.08878 0.06032 0.01521 0.00953 0.00281 0.00042
Cumulative Proportion  0.8229 0.91171 0.97203 0.98724 0.99677 0.99958 1.00000
                             PC1         PC2         PC3         PC4         PC5         PC6          PC7
Global_intensity       0.4145778 -0.08450392  0.01602253 -0.09979902  0.03923097 -0.32942014  0.837084865
Global_active_power    0.4133583 -0.03801173 -0.07534475 -0.07794687 -0.03713683 -0.74585730 -0.508188171
Global_reactive_power  0.4040475  0.05524862  0.06916759 -0.39577755 -0.74132762  0.34248720 -0.073519041
Voltage               -0.3533369  0.61104339 -0.01536960 -0.66500626  0.10294559 -0.20912089  0.070570127
Sub_metering_1         0.3816173 -0.12367850  0.52095902 -0.34629343  0.58116306  0.28639689 -0.167274432
Sub_metering_2         0.3194718  0.76200772  0.23713491  0.50705556  0.01104061  0.06180069 -0.001581508
Sub_metering_3         0.3485181  0.13796964 -0.81327412 -0.08981518  0.31471409  0.29716616 -0.051626376

 day  4  period  6
Importance of components:
                          PC1    PC2     PC3     PC4     PC5     PC6     PC7
Standard deviation     2.3079 0.9519 0.72620 0.32302 0.30977 0.19391 0.04654
Proportion of Variance 0.7609 0.1294 0.07534 0.01491 0.01371 0.00537 0.00031
Cumulative Proportion  0.7609 0.8904 0.96570 0.98061 0.99432 0.99969 1.00000
                             PC1         PC2          PC3         PC4        PC5          PC6         PC7
Global_intensity      -0.4315025  0.03655809  0.003825634 -0.04120122 -0.1804217  0.239438943 -0.84904297
Global_active_power   -0.4226822  0.10303829 -0.028382311  0.37466686 -0.3031821  0.619618253  0.44010891
Global_reactive_power -0.2847808 -0.71507347 -0.389671102  0.15596079  0.4793875  0.037576953  0.01334541
Voltage                0.4195441 -0.08342626  0.139236748 -0.39500100  0.3102914  0.736391864 -0.05528510
Sub_metering_1        -0.3078285  0.63886273 -0.425113679 -0.24090012  0.5009714 -0.005745089  0.08565139
Sub_metering_2        -0.3463287  0.03460307  0.800235506  0.11078528  0.4653008 -0.082692499  0.05353493
Sub_metering_3        -0.4041362 -0.24579189  0.082769945 -0.77927620 -0.2862837 -0.090477362  0.26831605
```

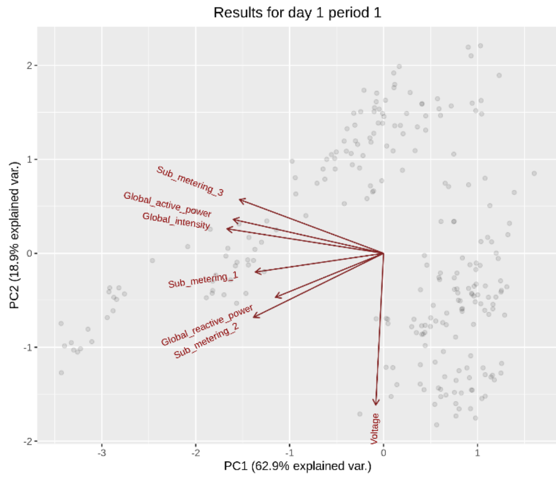
```
 day  5  period  1
Importance of components:
                          PC1    PC2    PC3     PC4     PC5     PC6     PC7
Standard deviation     1.9846 1.0629 0.9350 0.83335 0.52697 0.27020 0.11148
Proportion of Variance 0.5626 0.1614 0.1249 0.09921 0.03967 0.01043 0.00178
Cumulative Proportion  0.5626 0.7240 0.8489 0.94812 0.98779 0.99822 1.00000
                              PC1         PC2         PC3         PC4          PC5          PC6         PC7
Global_intensity      -0.49586061  0.06328034 -0.09826508  0.07116702 -0.150183834 -0.009895182 -0.84424166
Global_active_power   -0.48149553 -0.01374166 -0.20945167  0.05258605 -0.008725991 -0.786126432  0.32135186
Global_reactive_power -0.27693190  0.14560685  0.52599015 -0.78858341 -0.030382242 -0.004500336  0.05132835
Voltage                0.01598594 -0.86493533  0.37426868  0.09274353 -0.297031118 -0.107725191 -0.05586342
Sub_metering_1        -0.25904152  0.35602400  0.62603532  0.59117227 -0.147826292  0.100779774  0.18091531
Sub_metering_2        -0.42495845 -0.30024100  0.02886215  0.08646096  0.793264666  0.290181382  0.08650474
Sub_metering_3        -0.44457316 -0.09795173 -0.37010544 -0.06888147 -0.486925729  0.525281402  0.37151077

 day  5  period  2
Importance of components:
                          PC1    PC2     PC3     PC4     PC5     PC6     PC7
Standard deviation     2.3204 0.9240 0.69438 0.45708 0.25018 0.07196 0.05781
Proportion of Variance 0.7692 0.1220 0.06888 0.02985 0.00894 0.00074 0.00048
Cumulative Proportion  0.7692 0.8911 0.96000 0.98984 0.99878 0.99952 1.00000
                             PC1        PC2         PC3         PC4         PC5         PC6          PC7
Global_intensity       0.4091571 -0.3095566  0.03761499  0.02476684 -0.47077951  0.49411566 -0.518607004
Global_active_power    0.4129103 -0.2925367  0.05419810 -0.07203919 -0.26020486  0.08013548  0.813432242
Global_reactive_power  0.3299739  0.1285083 -0.89772602  0.21557615  0.14861428  0.01071124  0.004106763
Voltage               -0.4100134  0.1612597 -0.20601362  0.23677964 -0.80112097 -0.24666292  0.068852680
Sub_metering_1         0.3547130  0.3982653  0.37519752  0.75520599  0.04171818 -0.04673883  0.023004661
Sub_metering_2         0.2872733  0.7575700  0.05216803 -0.54481169 -0.19348015  0.08099725  0.005026279
Sub_metering_3         0.4204218 -0.2086870  0.06163914 -0.15653567 -0.08720153 -0.82445474 -0.253106159

 day  5  period  3
Importance of components:
                          PC1    PC2    PC3    PC4     PC5    PC6     PC7
Standard deviation     1.8993 1.2646 0.9607 0.6459 0.51516 0.3942 0.18005
Proportion of Variance 0.5153 0.2285 0.1318 0.0596 0.03791 0.0222 0.00463
Cumulative Proportion  0.5153 0.7438 0.8757 0.9353 0.97317 0.9954 1.00000
                               PC1         PC2         PC3         PC4         PC5         PC6         PC7
Global_intensity      -0.390604743 -0.45094856  0.26833509 -0.29591780  0.13384456 -0.11152053  0.67390732
Global_active_power   -0.488040155 -0.07823807  0.19123940 -0.37692073  0.13930869 -0.34593430 -0.66179787
Global_reactive_power  0.383041084 -0.38147570  0.22109718 -0.48429214 -0.54617921  0.30926466 -0.17428892
Voltage               -0.005500565 -0.39946253 -0.88551143 -0.21138088  0.08346713 -0.05717677 -0.03675755
Sub_metering_1        -0.243740348 -0.59311755  0.08594599  0.69335755 -0.22326398  0.08193912 -0.21002378
Sub_metering_2         0.415579733 -0.32583113  0.20683092  0.01258938  0.77355440  0.22537401 -0.17032379
Sub_metering_3        -0.479977231  0.16575835 -0.09079414 -0.10119031  0.09574459  0.84348762 -0.05499623

 day  5  period  4
Importance of components:
                         PC1    PC2    PC3     PC4     PC5     PC6     PC7
Standard deviation     2.090 1.0925 0.9105 0.64084 0.35864 0.22945 0.13156
Proportion of Variance 0.624 0.1705 0.1184 0.05867 0.01837 0.00752 0.00247
Cumulative Proportion  0.624 0.7945 0.9130 0.97163 0.99001 0.99753 1.00000
                             PC1          PC2         PC3         PC4         PC5          PC6         PC7
Global_intensity      -0.4696904  0.095396931 -0.07502141  0.07355805 -0.18291411 -0.220976889 -0.82277758
Global_active_power   -0.4385574 -0.119065523  0.09268925 -0.35618540  0.80421205 -0.096868775  0.04348416
Global_reactive_power -0.1714868  0.637544810  0.62705773  0.37941358  0.08580605 -0.028430778  0.13711981
Voltage                0.4657580  0.089963065  0.00491082 -0.02654839  0.15415949 -0.864192397 -0.06044470
Sub_metering_1        -0.1208508  0.648861051 -0.72145690 -0.03942542  0.08596248  0.001788436  0.18688833
Sub_metering_2        -0.3633033 -0.375667089 -0.23080735  0.70658374  0.03592152 -0.273448386  0.31350908
Sub_metering_3        -0.4405971 -0.004434336  0.13705069 -0.47143053 -0.52912917 -0.345501997  0.40678689
```

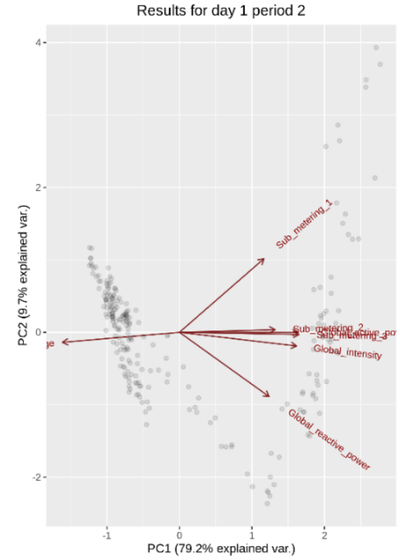
```
 day  5  period  5
Importance of components:
                          PC1    PC2     PC3     PC4     PC5     PC6     PC7
Standard deviation     2.1135 1.0231 0.82765 0.68655 0.52358 0.21753 0.09371
Proportion of Variance 0.6381 0.1495 0.09786 0.06733 0.03916 0.00676 0.00125
Cumulative Proportion  0.6381 0.7876 0.88549 0.95282 0.99199 0.99875 1.00000
                             PC1         PC2         PC3         PC4         PC5         PC6         PC7
Global_intensity       0.4692704 -0.02136106 -0.03592033  0.12894147  0.02876099  0.08348106 -0.86811092
Global_active_power    0.4566951 -0.02235822 -0.15855713  0.20081678 -0.09722987  0.76834488  0.35447768
Global_reactive_power  0.4110051 -0.07543339 -0.09901075 -0.41283527 -0.73469565 -0.30422650  0.11321253
Voltage               -0.4127062 -0.11974694  0.12061378 -0.64189113 -0.14088957  0.54258626 -0.27296997
Sub_metering_1         0.3855246 -0.20987739 -0.20500733 -0.55443964  0.64898333 -0.11959222  0.14969731
Sub_metering_2         0.2874322  0.19970539  0.92241259 -0.09781047  0.08077829  0.01413521  0.10180217
Sub_metering_3        -0.0267727 -0.94608854  0.23737219  0.20885244 -0.04885787 -0.03494778  0.02502721

 day  5  period  6
Importance of components:
                          PC1    PC2     PC3     PC4    PC5     PC6     PC7
Standard deviation     2.3122 0.9488 0.70076 0.39558 0.2593 0.18665 0.06420
Proportion of Variance 0.7637 0.1286 0.07015 0.02235 0.0096 0.00498 0.00059
Cumulative Proportion  0.7637 0.8923 0.96248 0.98483 0.9944 0.99941 1.00000
                             PC1         PC2         PC3         PC4         PC5          PC6         PC7
Global_intensity      -0.4274958 -0.11707987  0.05149572  0.18430040 -0.06276043 -0.141720579 -0.86191506
Global_active_power   -0.4227530  0.02015073  0.10461052  0.34975807 -0.17573243 -0.695933580  0.41520401
Global_reactive_power -0.2555242  0.75384545 -0.51145424 -0.21219019 -0.23946039  0.030612626 -0.03919045
Voltage                0.4166367 -0.01332891 -0.21318342 -0.41642331  0.28399389 -0.693731124 -0.21322580
Sub_metering_1        -0.3753103  0.05431246  0.57593986 -0.72096486 -0.03011075 -0.001896649  0.06152294
Sub_metering_2        -0.2911384 -0.64332739 -0.55949183 -0.32178542 -0.24967025  0.057910141  0.13821186
Sub_metering_3        -0.4171680  0.02461022 -0.18610319  0.04397071  0.87403864  0.100172569  0.12173467

 day  6  period  1
Importance of components:
                          PC1    PC2    PC3     PC4     PC5     PC6     PC7
Standard deviation     1.9800 1.0591 0.9136 0.76724 0.57531 0.43192 0.13003
Proportion of Variance 0.5601 0.1603 0.1192 0.08409 0.04728 0.02665 0.00242
Cumulative Proportion  0.5601 0.7203 0.8396 0.92365 0.97093 0.99758 1.00000
                              PC1          PC2         PC3         PC4         PC5         PC6         PC7
Global_intensity      -0.49263721  0.059324126  0.17054515 -0.02082695 -0.07927484 -0.18530367  0.82683001
Global_active_power   -0.46361337  0.044398448  0.26626318 -0.06846117  0.07155092 -0.68509431 -0.48273663
Global_reactive_power -0.30479753 -0.017146765 -0.78054124  0.35450215 -0.38924929 -0.12145095 -0.07498476
Voltage                0.03765806  0.912258036  0.05967052 -0.16278878 -0.35095057  0.08810962 -0.07332653
Sub_metering_1        -0.39518624  0.276313957  0.05875550  0.53063157  0.56175298  0.39282955 -0.11213783
Sub_metering_2        -0.35770886 -0.002620461 -0.38737657 -0.74826747  0.31172047  0.24581771 -0.06690814
Sub_metering_3        -0.40487499 -0.292650679  0.36565046 -0.03511809 -0.54879741  0.50897279 -0.23508805

 day  6  period  2
Importance of components:
                          PC1    PC2     PC3     PC4     PC5     PC6     PC7
Standard deviation     2.2985 0.9730 0.74015 0.36139 0.25225 0.15186 0.07178
Proportion of Variance 0.7547 0.1352 0.07826 0.01866 0.00909 0.00329 0.00074
Cumulative Proportion  0.7547 0.8900 0.96822 0.98688 0.99597 0.99926 1.00000
                             PC1         PC2         PC3         PC4         PC5          PC6         PC7
Global_intensity       0.4252505 -0.12506514  0.13161364  0.11555463 -0.49645078  0.101083806  0.71844584
Global_active_power    0.4272900 -0.07432259  0.11340781  0.13403925  0.29887939  0.803518137 -0.21471214
Global_reactive_power  0.2984100  0.30752834 -0.89439022  0.10818694 -0.06536431 -0.003442938 -0.02133439
Voltage               -0.4203644  0.14633423 -0.07025041 -0.26615574 -0.66981681  0.486728271 -0.20136318
Sub_metering_1         0.1683256  0.90212555  0.37931374  0.08349176 -0.02479403 -0.072960407 -0.03237639
Sub_metering_2         0.4111065 -0.01238935  0.02282490 -0.90224772  0.03893413 -0.103803571 -0.06304739
Sub_metering_3         0.4165196 -0.22105081  0.14334447  0.25517738 -0.45732928 -0.301870837 -0.62586642

 day  6  period  3
Importance of components:
                          PC1    PC2    PC3    PC4     PC5     PC6     PC7
Standard deviation     1.8779 1.0949 1.0693 0.6995 0.53320 0.48009 0.35659
Proportion of Variance 0.5038 0.1713 0.1633 0.0699 0.04061 0.03293 0.01816
Cumulative Proportion  0.5038 0.6751 0.8384 0.9083 0.94891 0.98184 1.00000
                             PC1         PC2           PC3         PC4          PC5         PC6         PC7
Global_intensity      -0.2020576  0.32852395 -0.7522383823  0.09699607  0.460710536  0.25238107  0.00484936
Global_active_power   -0.4271252  0.25521524 -0.2423978327 -0.37951557 -0.671973941 -0.01813104 -0.31267031
Global_reactive_power -0.4282253 -0.40264460  0.1187018673 -0.23603724 -0.007756143  0.63167596  0.43083879
Voltage                0.3603559 -0.30520475 -0.5369860585  0.11532643 -0.446049830 -0.15446884  0.50251206
Sub_metering_1        -0.3928743  0.27834795  0.1749276407  0.78551152 -0.259042262  0.01106374  0.23090875
Sub_metering_2        -0.4811918 -0.05631003 -0.0004395515 -0.23614983  0.260660809 -0.70560648  0.37906752
Sub_metering_3         0.2785816  0.70072547  0.2057587760 -0.32365235 -0.046143960  0.12286993  0.51678071
```

```
 day  6  period  4
Importance of components:
                          PC1    PC2    PC3     PC4     PC5     PC6     PC7
Standard deviation     1.5724 1.5456 1.0856 0.73739 0.54906 0.28823 0.17817
Proportion of Variance 0.3532 0.3413 0.1684 0.07768 0.04307 0.01187 0.00453
Cumulative Proportion  0.3532 0.6945 0.8629 0.94053 0.98360 0.99547 1.00000
                             PC1         PC2         PC3         PC4         PC5         PC6         PC7
Global_intensity       0.40416531  0.38452670  0.43663661  0.01562796 -0.05501288  0.16778799 -0.68316516
Global_active_power    0.52241549 -0.07121939 -0.22703786  0.49219446  0.62425547 -0.19262217  0.03755115
Global_reactive_power -0.26777265 -0.23133844  0.62257945  0.65626825 -0.05423899  0.15933686  0.16780063
Voltage                0.18176535 -0.58038356 -0.23820486 -0.03673265 -0.03899061  0.73115670 -0.18951266
Sub_metering_1         0.35756920 -0.24571843  0.55056475 -0.53586510  0.26696663 -0.01689021  0.38721740
Sub_metering_2         0.56800775 -0.07816772 -0.06621467  0.18368735 -0.72635806 -0.18514397  0.26694054
Sub_metering_3         0.09221219  0.62467329 -0.07713392  0.06760647  0.06284375  0.58326835  0.49659703

 day  6  period  5
Importance of components:
                          PC1    PC2     PC3     PC4     PC5     PC6     PC7
Standard deviation     2.2207 1.1365 0.59733 0.54959 0.25737 0.21536 0.07333
Proportion of Variance 0.7045 0.1845 0.05097 0.04315 0.00946 0.00663 0.00077
Cumulative Proportion  0.7045 0.8890 0.93999 0.98314 0.99261 0.99923 1.00000
                            PC1         PC2         PC3          PC4         PC5         PC6          PC7
Global_intensity       0.4439465 -0.08034386  0.17845574 -0.006791573  0.25087026  0.09761029  0.831925543
Global_active_power    0.4410658 -0.06636928  0.08033339  0.068373261  0.38742329  0.65949180 -0.452660425
Global_reactive_power  0.3650975  0.24241161 -0.77084142  0.418510302  0.04468958 -0.19119656  0.006307541
Voltage               -0.3943500  0.37362247 -0.26796861 -0.027409213 -0.25411322  0.69151852  0.299273283
Sub_metering_1         0.3563505  0.34201412 -0.16450899 -0.837265847 -0.11710130 -0.08547241 -0.083337264
Sub_metering_2         0.1354560  0.78210043  0.50337590  0.305216808 -0.05003856 -0.12518540 -0.072462121
Sub_metering_3         0.4174091 -0.24925510  0.12347437  0.158762199 -0.83915437  0.13346118 -0.034616309

 day  6  period  6
Importance of components:
                          PC1     PC2    PC3    PC4     PC5     PC6     PC7
Standard deviation     2.4159 0.83563 0.5325 0.3005 0.21377 0.18603 0.10465
Proportion of Variance 0.8338 0.09975 0.0405 0.0129 0.00653 0.00494 0.00156
Cumulative Proportion  0.8338 0.93356 0.9741 0.9870 0.99349 0.99844 1.00000
                            PC1         PC2        PC3         PC4         PC5        PC6         PC7
Global_intensity      -0.4027986 -0.16043230  0.3017384 -0.02637566  0.15987676 -0.1298807 -0.82331249
Global_active_power   -0.4019233  0.04904604 -0.2032610  0.41252338 -0.78238517 -0.1055745 -0.03590347
Global_reactive_power -0.3865467  0.31472740 -0.2826648  0.43616855  0.44980392  0.5280710  0.01425941
Voltage                0.3889876  0.15046406 -0.5193000  0.35022434  0.21665841 -0.5488288 -0.29251613
Sub_metering_1        -0.2930335 -0.78835845 -0.4681601 -0.09714615  0.17259873 -0.0666757  0.17255476
Sub_metering_2        -0.3611365  0.46470147 -0.3995418 -0.67851177 -0.01411644 -0.1823864 -0.01253125
Sub_metering_3        -0.3989900  0.11272397  0.3731750  0.21543699  0.28819931 -0.5951491  0.45295286

 day  7  period  1
Importance of components:
                          PC1    PC2    PC3     PC4     PC5     PC6     PC7
Standard deviation     1.6257 1.5356 0.9193 0.78005 0.54377 0.45956 0.19633
Proportion of Variance 0.3776 0.3368 0.1207 0.08693 0.04224 0.03017 0.00551
Cumulative Proportion  0.3776 0.7144 0.8352 0.92208 0.96432 0.99449 1.00000
                            PC1        PC2         PC3         PC4        PC5         PC6         PC7
Global_intensity      -0.1240037  0.6265127  0.06342437 -0.03204164  0.1558270 -0.07359086  0.74656405
Global_active_power   -0.2472140  0.5134124  0.22540772 -0.46662641  0.2228632  0.27504332 -0.53049718
Global_reactive_power -0.3865819 -0.1597204  0.62307527  0.53178522  0.3851126 -0.05956473 -0.04653848
Voltage                0.4567909  0.2410319  0.49239166 -0.04562052 -0.2904603 -0.61274816 -0.16996300
Sub_metering_1         0.1133753  0.4746900 -0.46082023  0.63267224  0.1236837 -0.10780540 -0.34966599
Sub_metering_2        -0.5244672  0.1509384  0.06335495  0.18349907 -0.8085635  0.09803874 -0.03285531
Sub_metering_3        -0.5267685 -0.1103152 -0.31325268 -0.24982052  0.1558797 -0.72020393 -0.08255807

 day  7  period  2
Importance of components:
                          PC1    PC2    PC3     PC4     PC5     PC6     PC7
Standard deviation     1.9512 1.1528 0.9842 0.75538 0.47009 0.27633 0.16536
Proportion of Variance 0.5439 0.1899 0.1384 0.08151 0.03157 0.01091 0.00391
Cumulative Proportion  0.5439 0.7337 0.8721 0.95362 0.98519 0.99609 1.00000
                             PC1        PC2         PC3           PC4         PC5          PC6           PC7
Global_intensity       0.48833459 -0.1423477  0.12325591 -0.0547410982  0.26254727 -0.565232593 -0.5784971024
Global_active_power    0.45949967  0.1408591 -0.01640620  0.0002103741 -0.87580181 -0.007052957 -0.0408781198
Global_reactive_power -0.04750302 -0.7491828  0.10809569  0.6338418471 -0.14793058  0.032865082  0.0080521156
Voltage               -0.47993537  0.1704838 -0.09648807  0.1680067304 -0.22475022 -0.786761437  0.1831798684
Sub_metering_1         0.25658218  0.5425623 -0.18584922  0.7406921491  0.22493402  0.077662774 -0.0003968365
Sub_metering_2         0.07908188 -0.2355606 -0.96195500 -0.1014004758  0.02367226 -0.019285405 -0.0410548616
Sub_metering_3         0.49568793 -0.1403340  0.06038912 -0.0900816465  0.20121709 -0.232347296  0.7926945957
```

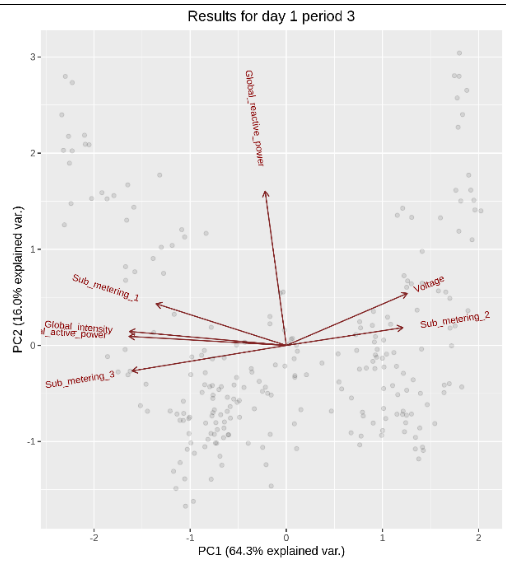
```
 day  7  period  3
Importance of components:
                          PC1    PC2     PC3     PC4     PC5    PC6     PC7
Standard deviation     2.1480 1.0161 0.76649 0.55670 0.53849 0.3108 0.26373
Proportion of Variance 0.6591 0.1475 0.08393 0.04427 0.04142 0.0138 0.00994
Cumulative Proportion  0.6591 0.8066 0.89056 0.93484 0.97626 0.9901 1.00000
                             PC1         PC2         PC3         PC4          PC5         PC6        PC7
Global_intensity       0.2652762  0.75679379 -0.06007755  0.33986783 -0.356134229  0.23799972 -0.2330167
Global_active_power    0.4406157 -0.05519860  0.19386231 -0.18980899  0.251405962 -0.34397048 -0.7400542
Global_reactive_power  0.3663478  0.11284620 -0.63041727 -0.64949028 -0.065430106  0.06064037  0.1607268
Voltage               -0.3995355  0.15088529 -0.55182532  0.23820230  0.000130967 -0.65903285 -0.1484223
Sub_metering_1         0.4079492  0.16166278 -0.09822578  0.37070456  0.708958976 -0.08147535  0.3887304
Sub_metering_2         0.4247719 -0.09505757  0.26181364  0.06671335 -0.506683770 -0.56501493  0.4019521
Sub_metering_3        -0.3078851  0.59459637  0.42270118 -0.47750592  0.215246521 -0.24730303  0.1936075

 day  7  period  4
Importance of components:
                          PC1    PC2    PC3     PC4    PC5     PC6     PC7
Standard deviation     1.9198 1.1466 1.0505 0.70167 0.4583 0.41427 0.14837
Proportion of Variance 0.5265 0.1878 0.1577 0.07033 0.0300 0.02452 0.00314
Cumulative Proportion  0.5265 0.7144 0.8720 0.94233 0.9723 0.99686 1.00000
                             PC1         PC2         PC3         PC4         PC5        PC6         PC7
Global_intensity      -0.4851802  0.05495648  0.28133806 -0.03882008  0.28947924 -0.2532385  0.73006431
Global_active_power   -0.4314623 -0.32630340  0.17702893 -0.03097322 -0.78307658  0.2408614  0.06200548
Global_reactive_power -0.3824118  0.27902051 -0.20938324  0.77248666 -0.08791866 -0.2837826 -0.21695513
Voltage                0.4778497 -0.08550377  0.12991436  0.03669336 -0.41181636 -0.7367608  0.18361784
Sub_metering_1        -0.2884767  0.62009663  0.14975968 -0.53978635 -0.15309422 -0.2624796 -0.35514885
Sub_metering_2         0.1298386  0.02707159  0.89793916  0.28462647  0.12560169  0.1198533 -0.25487550
Sub_metering_3        -0.3221561 -0.64814337  0.00823882 -0.16452989  0.29403538 -0.4138196 -0.43736044

 day  7  period  5
Importance of components:
                          PC1    PC2     PC3     PC4     PC5    PC6     PC7
Standard deviation     2.3128 1.0134 0.59244 0.39660 0.30543 0.1295 0.07508
Proportion of Variance 0.7642 0.1467 0.05014 0.02247 0.01333 0.0024 0.00081
Cumulative Proportion  0.7642 0.9109 0.96100 0.98347 0.99680 0.9992 1.00000
                              PC1          PC2           PC3         PC4         PC5         PC6         PC7
Global_intensity       0.43006558 -0.001823191  0.0199791192 -0.16822962  0.12713148  0.13541522 -0.86708720
Global_active_power    0.42952428 -0.005244332 -0.0007986953 -0.01979098  0.10391517  0.82123673  0.36036224
Global_reactive_power  0.40704842  0.104092580 -0.0756827870  0.62716579 -0.64266289 -0.10187738 -0.03188911
Voltage               -0.38918153  0.154264367 -0.5575888762  0.53199257  0.33123139  0.27080629 -0.21856008
Sub_metering_1         0.37796668  0.042225799 -0.7797616193 -0.34531507  0.03848240 -0.29720979  0.19563470
Sub_metering_2         0.02786573  0.980078247  0.1424936439 -0.11333348  0.05257998 -0.03480886  0.03930526
Sub_metering_3         0.41202730 -0.054772485  0.2337398902  0.40356798  0.66788012 -0.36602266  0.17232408

 day  7  period  6
Importance of components:
                         PC1     PC2     PC3     PC4     PC5     PC6     PC7
Standard deviation     2.475 0.72767 0.41448 0.33952 0.19863 0.12918 0.04819
Proportion of Variance 0.875 0.07564 0.02454 0.01647 0.00564 0.00238 0.00033
Cumulative Proportion  0.875 0.95064 0.97518 0.99165 0.99728 0.99967 1.00000
                             PC1         PC2         PC3         PC4         PC5         PC6         PC7
Global_intensity      -0.4024206  0.08503313 -0.03333377  0.06413076 -0.17420127  0.19928387  0.86921976
Global_active_power   -0.3963821  0.17868028 -0.03329838  0.30254093 -0.15734944  0.71862031 -0.42088102
Global_reactive_power -0.3262376 -0.78202469  0.28590066 -0.20215194  0.35593434  0.17991036 -0.01857002
Voltage                0.3859011 -0.08034675 -0.48665431 -0.58500841 -0.01559382  0.50686076  0.09168690
Sub_metering_1        -0.3569154  0.53768542  0.31719653 -0.63659606  0.25707390 -0.04791116 -0.09620345
Sub_metering_2        -0.3825447  0.04183449 -0.72423842  0.11997063  0.50512283 -0.23200555 -0.06340005
Sub_metering_3        -0.3896912 -0.22796661 -0.23253718 -0.31876087 -0.70478832 -0.31380435 -0.21281489
```

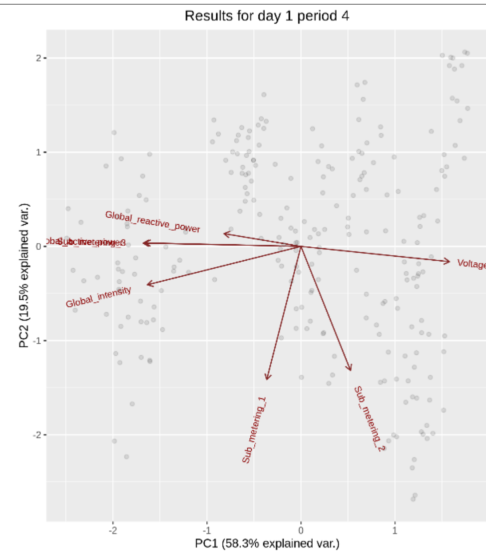
B. Appendix B – PCA Plots for Complete Dataset

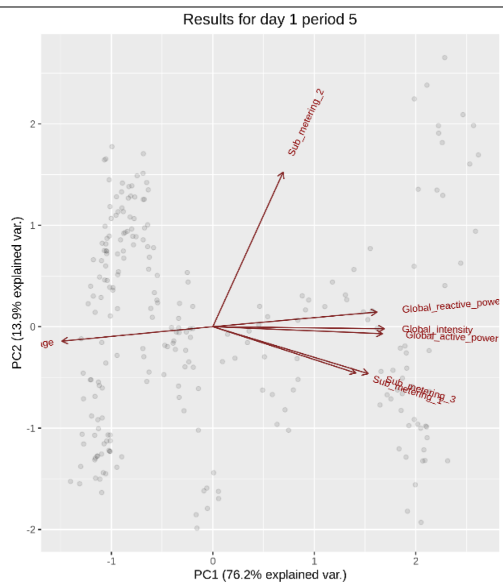


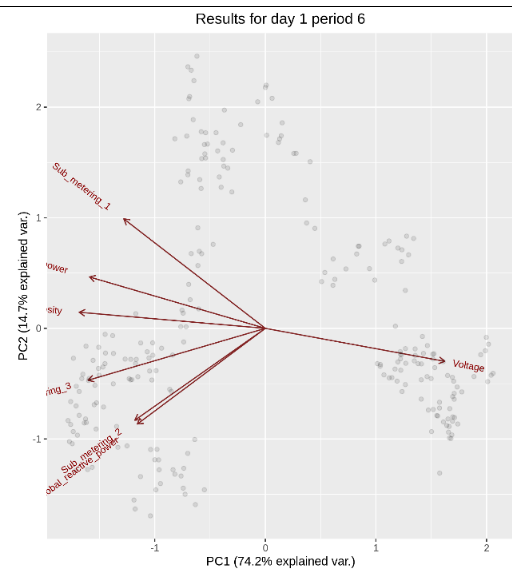


C.







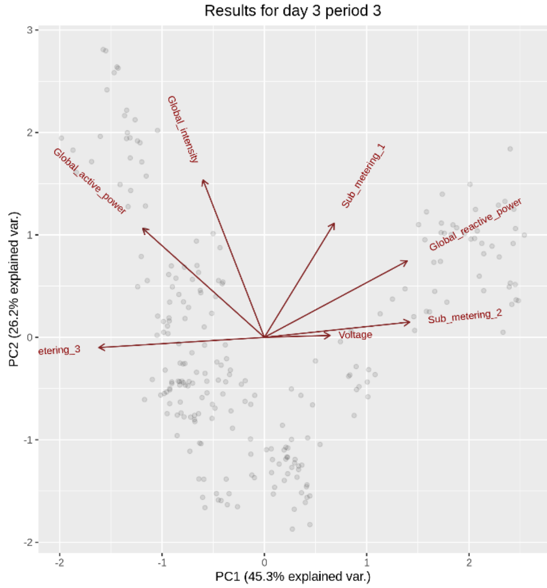



Results for day 2 period 1
PC1 (60.7% explained var.)
PC2 (16.7% explained var.)
Global_reactive_power
Voltage
Global_active_power
Global_intensity
Sub_metering_3
Sub_metering_2
Sub_metering_1
Results for day 2 period 2
PC1 (78.9% explained var.)
PC2 (11.6% explained var.)
Sub_metering_2
Sub_metering_1
Sub_metering_3
Global_active_power
Global_intensity
Global_reactive_power
Results for day 2 period 3
PC1 (59.0% explained var.)
PC2 (20.7% explained var.)
Sub_metering_2
Global_intensity
Global_active_power
Sub_metering_1
Sub_metering_3
Voltage
Results for day 2 period 4
PC1 (66.3% explained var.)
PC2 (12.8% explained var.)
Voltage
Sub_metering_1
Results for day 2 period 5
PC1 (66.3% explained var.)
PC2 (24.4% explained var.)
Sub_metering_3
Sub_metering_2
Sub_metering_1
Global_active_power
Results for day 2 period 6
PC1 (78.5% explained var.)
PC2 (10.8% explained var.)
Sub_metering_1
Sub_metering_2
Voltage
Global_reactive_power

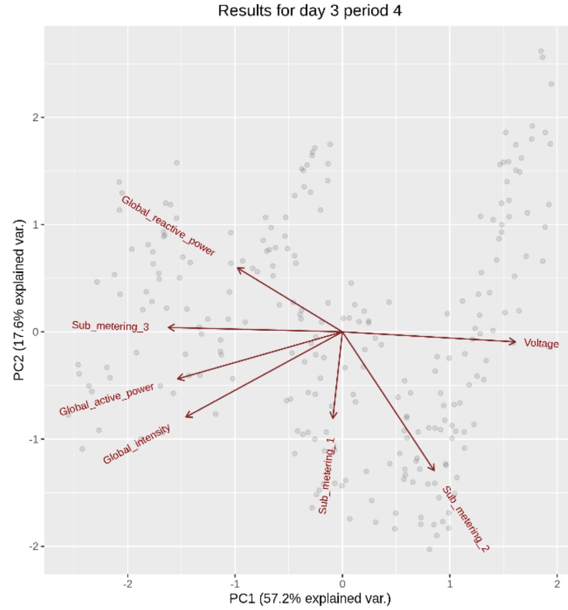
Results for day 3 period 1

Results for day 3 period 2

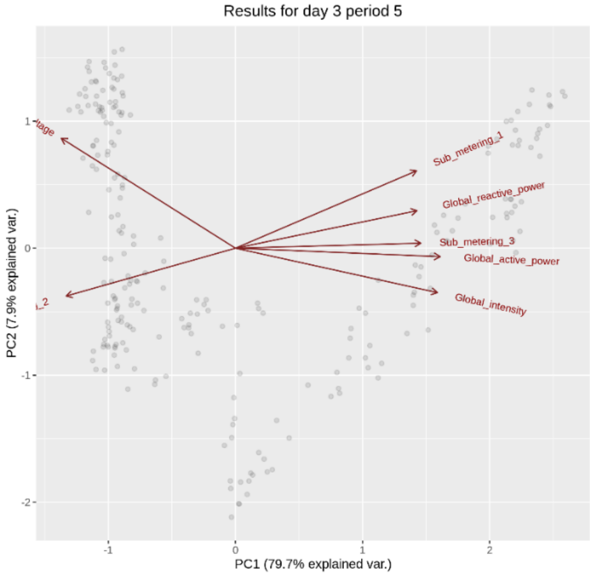
Results for day 3 period 3



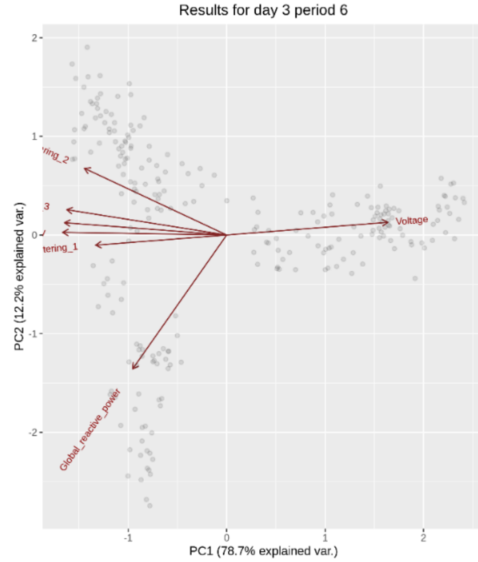
Results for day 3 period 4

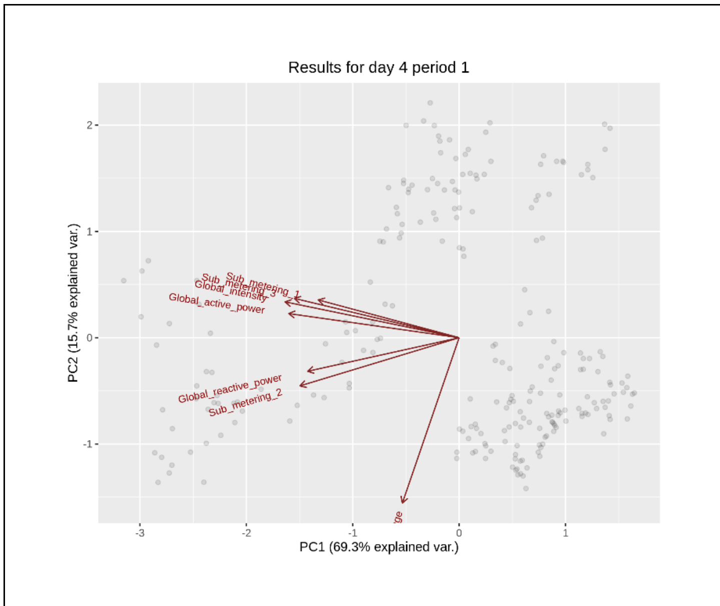


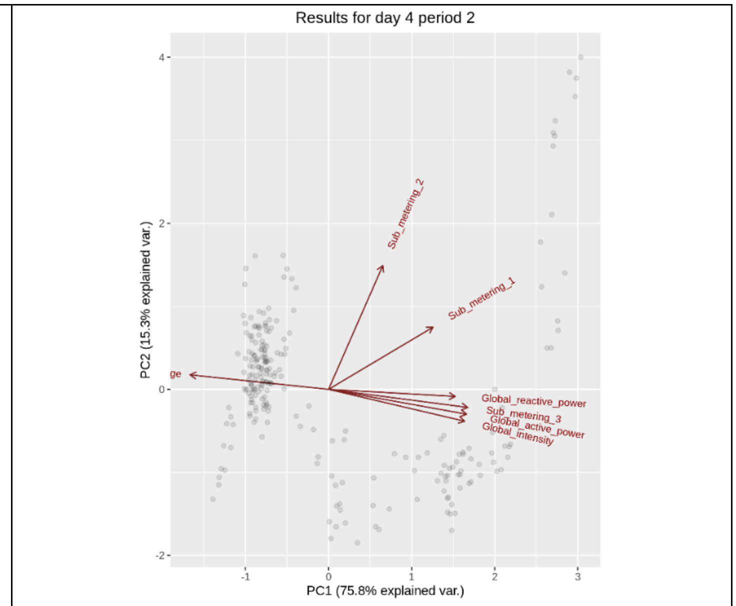
Results for day 3 period 5

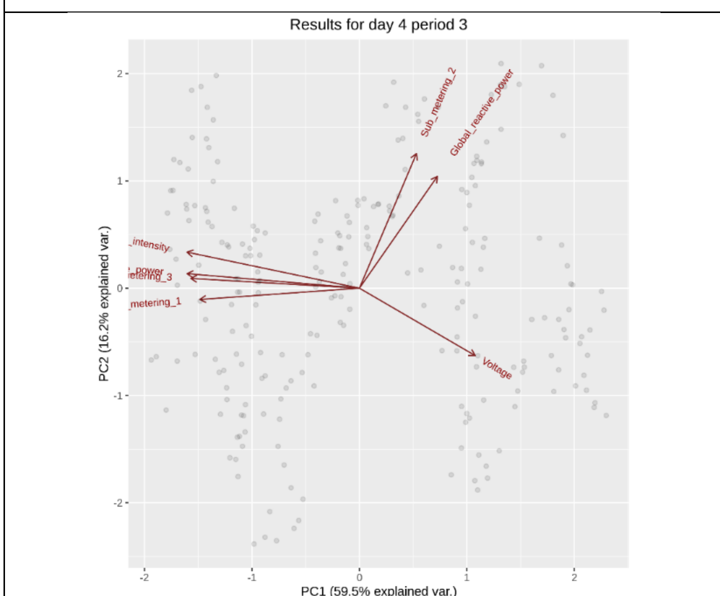


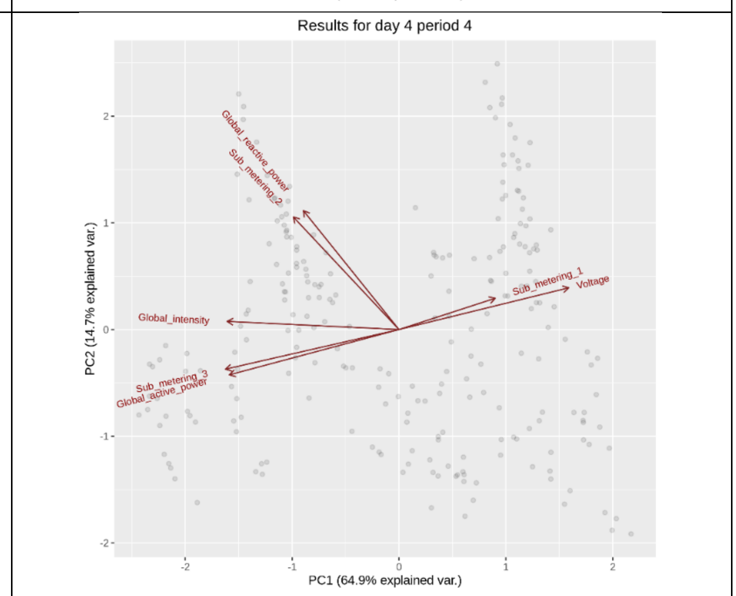
Results for day 3 period 6

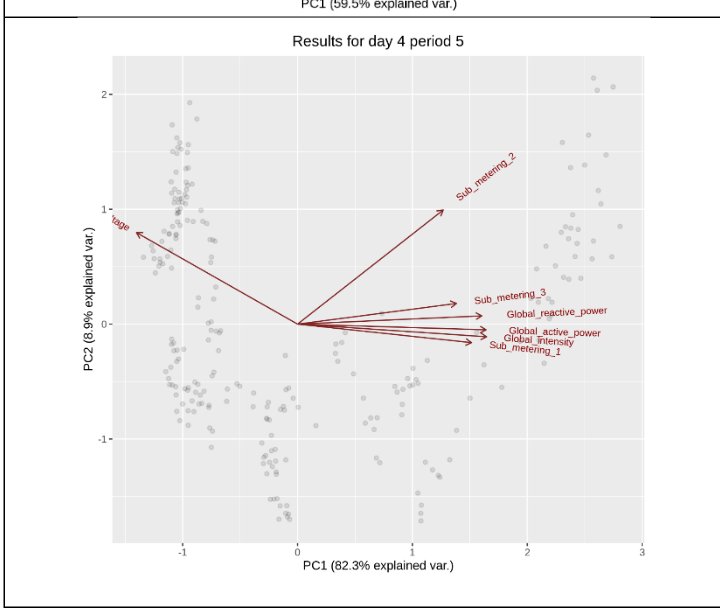


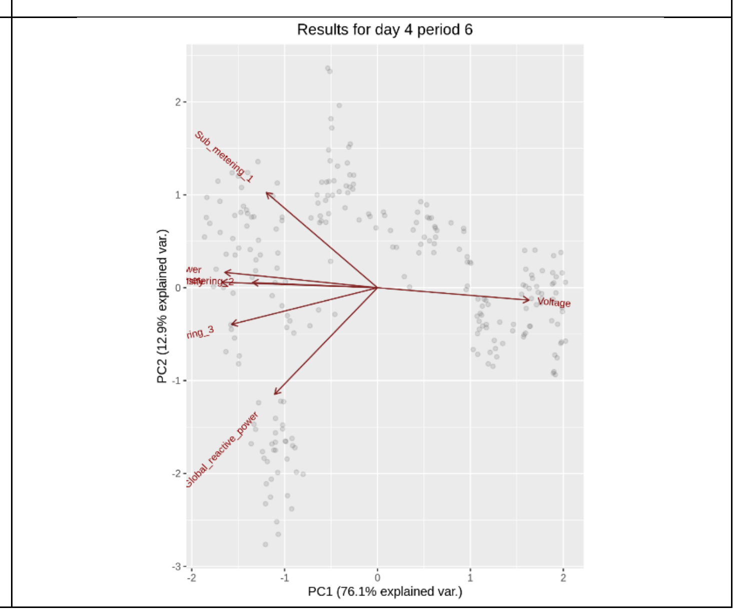

Results for day 4 period 1
PC1 (69.3% explained var.)
PC2 (15.7% explained var.)
Sub_metering_1
Global_intensity
Global_active_power
Global_reactive_power
Sub_metering_2

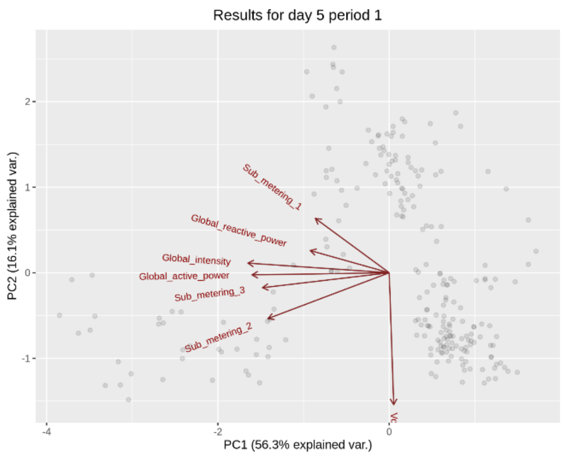

Results for day 4 period 2
PC1 (75.8% explained var.)
PC2 (15.3% explained var.)
Sub_metering_2
Sub_metering_1
Global_reactive_power
Sub_metering_3
Global_active_power
Global_intensity

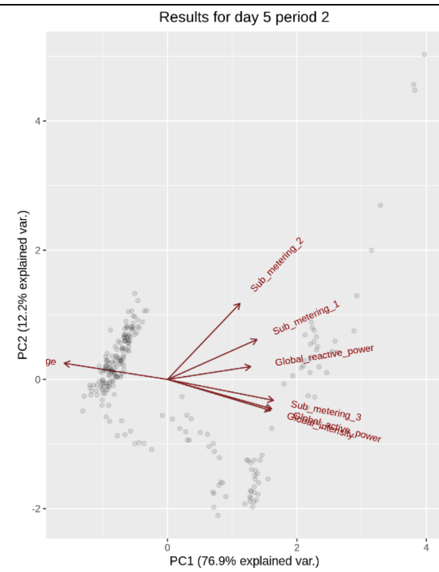

Results for day 4 period 3
PC1 (59.5% explained var.)
PC2 (16.2% explained var.)
Sub_metering_2
Global_reactive_power
Voltage

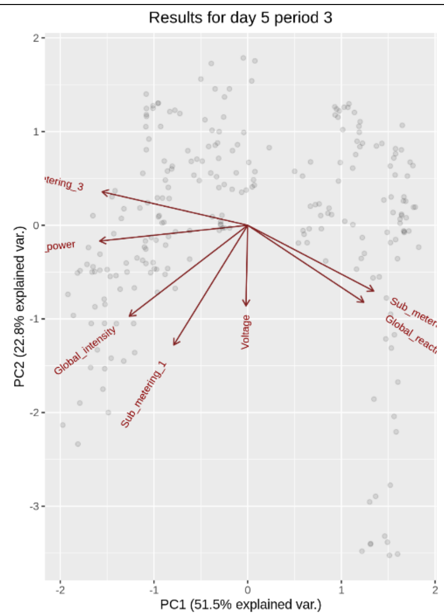

Results for day 4 period 4
PC1 (64.9% explained var.)
PC2 (14.7% explained var.)
Global_reactive_power
Sub_metering_2
Global_intensity
Sub_metering_1
Voltage
Sub_metering_3
Global_active_power

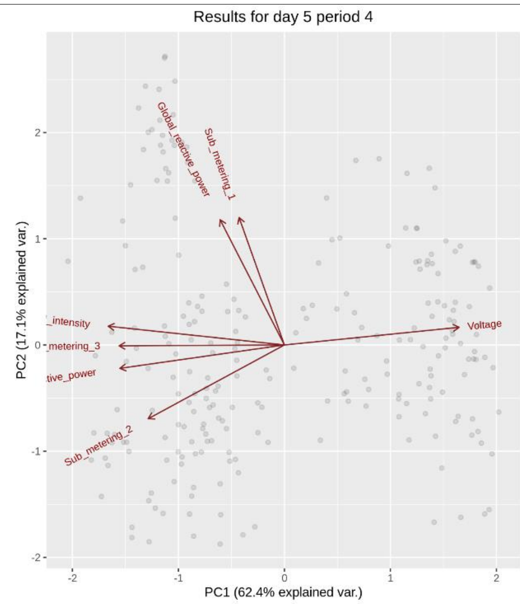

Results for day 4 period 5
PC1 (82.3% explained var.)
PC2 (8.9% explained var.)
Sub_metering_2
Sub_metering_3
Global_reactive_power
Global_active_power
Global_intensity
Sub_metering_1

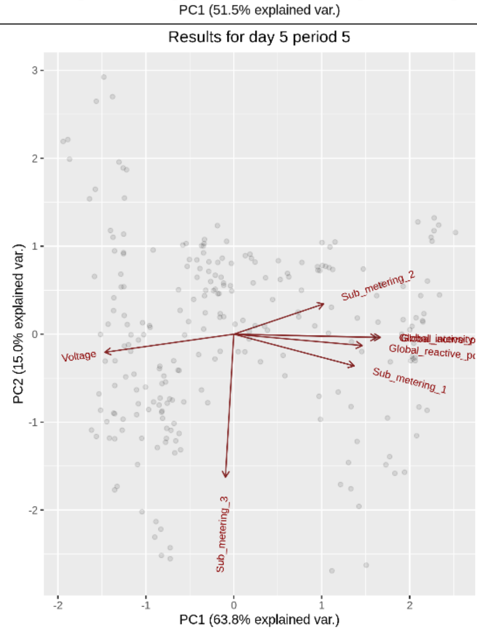

Results for day 4 period 6
PC1 (76.1% explained var.)
PC2 (12.9% explained var.)
Sub_metering_1
Voltage
Global_reactive_power

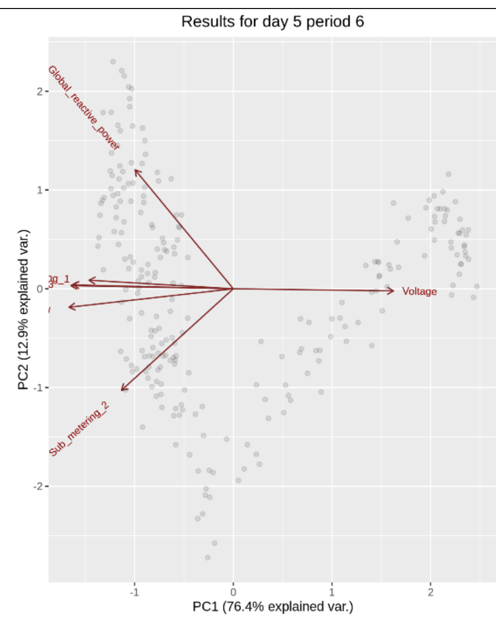

Results for day 5 period 1
Sub_metering_1
Global_reactive_power
Global_intensity
Global_active_power
Sub_metering_3
Sub_metering_2
PC2 (16.1% explained var.)
PC1 (56.3% explained var.)

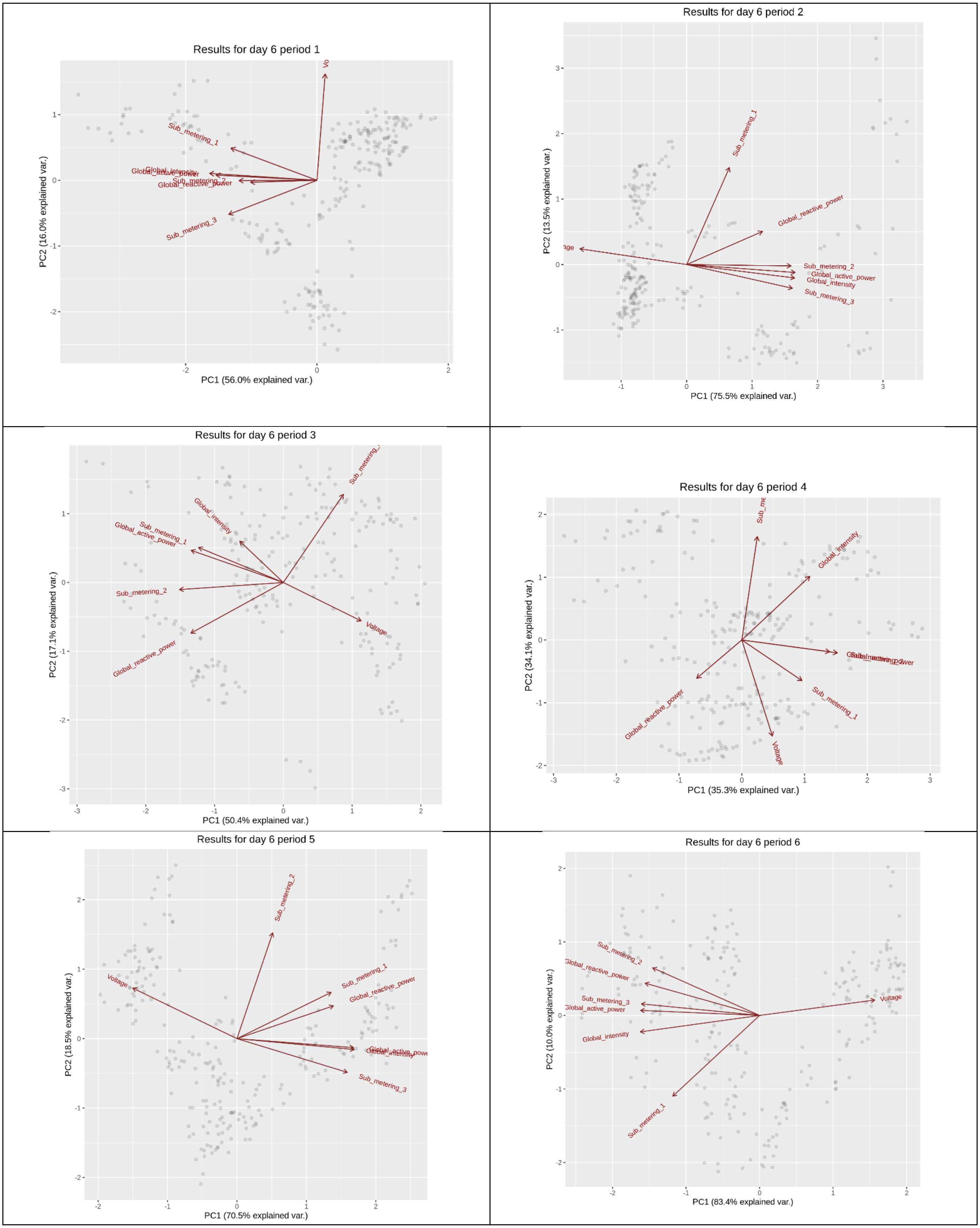

Results for day 5 period 2
Sub_metering_2
Sub_metering_1
Global_reactive_power
Sub_metering_3
PC2 (12.2% explained var.)
PC1 (76.9% explained var.)


Results for day 5 period 3
Global_intensity
Sub_metering_1
Voltage
PC2 (22.8% explained var.)
PC1 (51.5% explained var.)

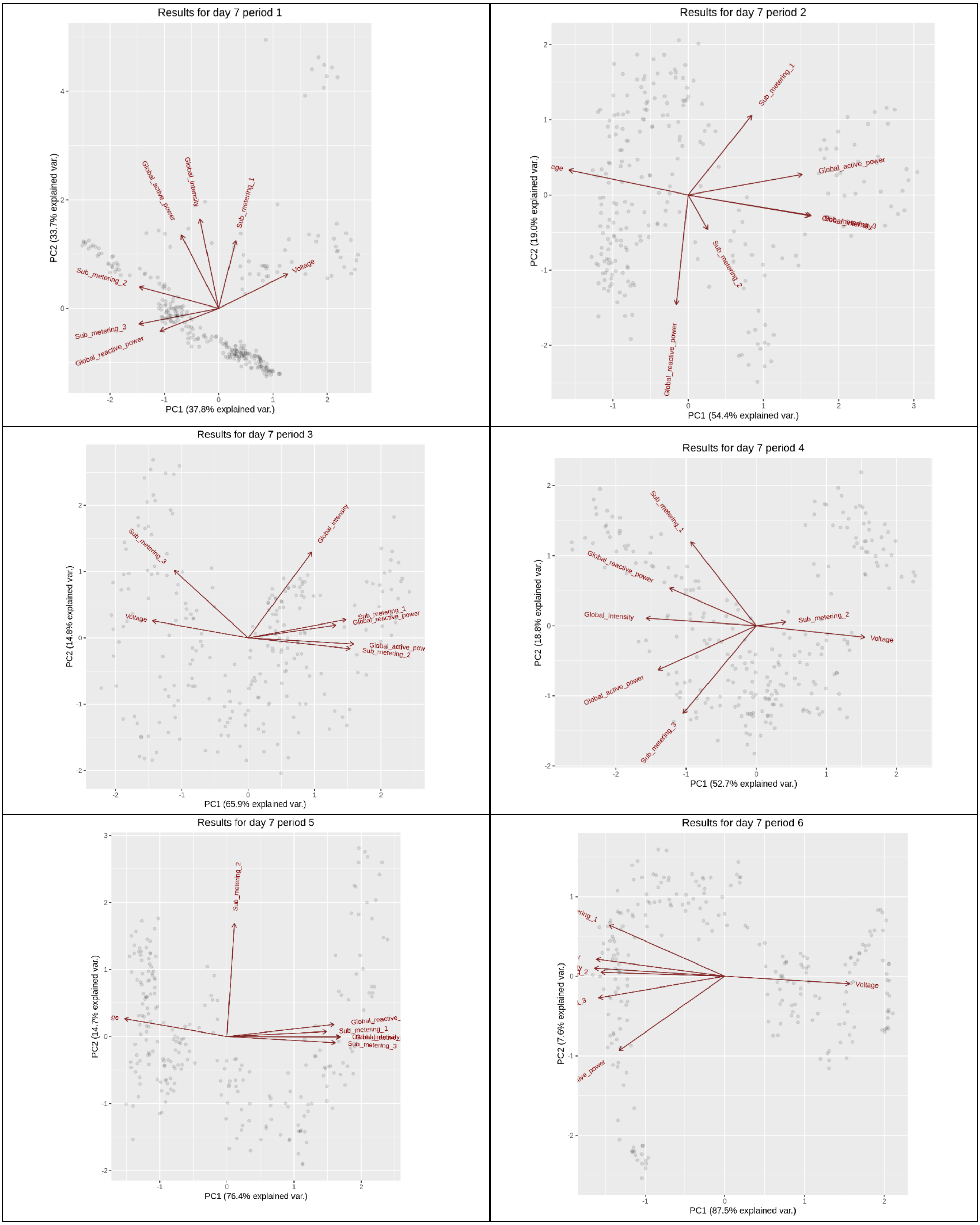

Results for day 5 period 4
Global_reactive_power
Sub_metering_1
Voltage
Sub_metering_2
PC2 (17.1% explained var.)
PC1 (62.4% explained var.)


Results for day 5 period 5
Sub_metering_2
Voltage
Global_reactive_power
Sub_metering_1
Sub_metering_3
PC2 (15.0% explained var.)
PC1 (63.8% explained var.)


Results for day 5 period 6
Global_reactive_power
Voltage
Sub_metering_2
PC2 (12.9% explained var.)
PC1 (76.4% explained var.)



Results for day 6 period 1
PC1 (56.0% explained var.)
PC2 (16.0% explained var.)
Sub_metering_1
Global_intensity
Global_active_power
Sub_metering_2
Global_reactive_power
Sub_metering_3
Results for day 6 period 2
PC1 (75.5% explained var.)
PC2 (13.5% explained var.)
Sub_metering_1
Global_reactive_power
Sub_metering_2
Global_active_power
Global_intensity
Sub_metering_3
Results for day 6 period 3
PC1 (50.4% explained var.)
PC2 (17.1% explained var.)
Global_intensity
Sub_metering_1
Global_active_power
Sub_metering_2
Global_reactive_power
Voltage
Results for day 6 period 4
PC1 (35.3% explained var.)
PC2 (34.1% explained var.)
Global_intensity
Global_reactive_power
Sub_metering_1
Voltage
Results for day 6 period 5
PC1 (70.5% explained var.)
PC2 (18.5% explained var.)
Sub_metering_2
Sub_metering_1
Global_reactive_power
Voltage
Sub_metering_3
Results for day 6 period 6
PC1 (83.4% explained var.)
PC2 (10.0% explained var.)
Sub_metering_2
Global_reactive_power
Sub_metering_3
Global_active_power
Global_intensity
Voltage
Sub_metering_1


Results for day 7 period 1
PC1 (37.8% explained var.)
PC2 (33.7% explained var.)
Global_active_power
Global_intensity
Sub_metering_1
Sub_metering_2
Voltage
Sub_metering_3
Global_reactive_power
Results for day 7 period 2
PC1 (54.4% explained var.)
PC2 (19.0% explained var.)
Sub_metering_1
Global_active_power
Sub_metering_2
Global_reactive_power
Results for day 7 period 3
PC1 (65.9% explained var.)
PC2 (14.8% explained var.)
Global_intensity
Sub_metering_3
Voltage
Sub_metering_1
Global_reactive_power
Sub_metering_2
Results for day 7 period 4
PC1 (52.7% explained var.)
PC2 (18.8% explained var.)
Sub_metering_1
Global_reactive_power
Global_intensity
Sub_metering_2
Voltage
Global_active_power
Sub_metering_3
Results for day 7 period 5
PC1 (76.4% explained var.)
PC2 (14.7% explained var.)
Sub_metering_2
Sub_metering_1
Sub_metering_3
Results for day 7 period 6
PC1 (87.5% explained var.)
PC2 (7.6% explained var.)
Voltage