Anomaly Detection Using Hidden Markov Models

Term Project (Group 12)

CMPT 318 Spring 2021

Kyle Isaak - 301288868
Colin Kirkby - 301381501

Abstract

The goal of this project was to develop and train Hidden Markov Models (HMMs) to be used in the analysis of power data for the purpose of anomaly detection and protection of power grids. This report outlines the process used for data analysis as well as explaining why HMMs are useful and how they can be used to improve cybersecurity for both power grids and other systems.

Table of Contents

Table of Figures

Project overview:

The goal of this project was to develop and train Hidden Markov Models (HMMs) to be used in the analysis of power data for the purpose of anomaly detection and protection of power grids. This report outlines the process used for data analysis as well as explaining why HMMs are useful and how they can be used to improve cybersecurity for both power grids and other systems.

Problem scope

The scope of the problem includes:

- Performing principal component analysis on energy data to discern valuable information for training models
- Developing and training HMMs for the purpose of anomaly detection within energy grid data
- Properly fitting HMMs to training data
- Testing trained models on anomalous datasets

Technical background

Almost every single North America citizen relies on the proper functioning of our power grids. This makes power grids sensitive targets for malicious attacks and intrusions. Analysis of data created by power providers can help detect and prevent these intrusions. Hidden Markov Models are a useful tool which can be used in this analysis and detection process. HMMs are trained on “normal” datasets so they can later be used to give a user the probability that a certain result or observation will occur.

Project contributions:

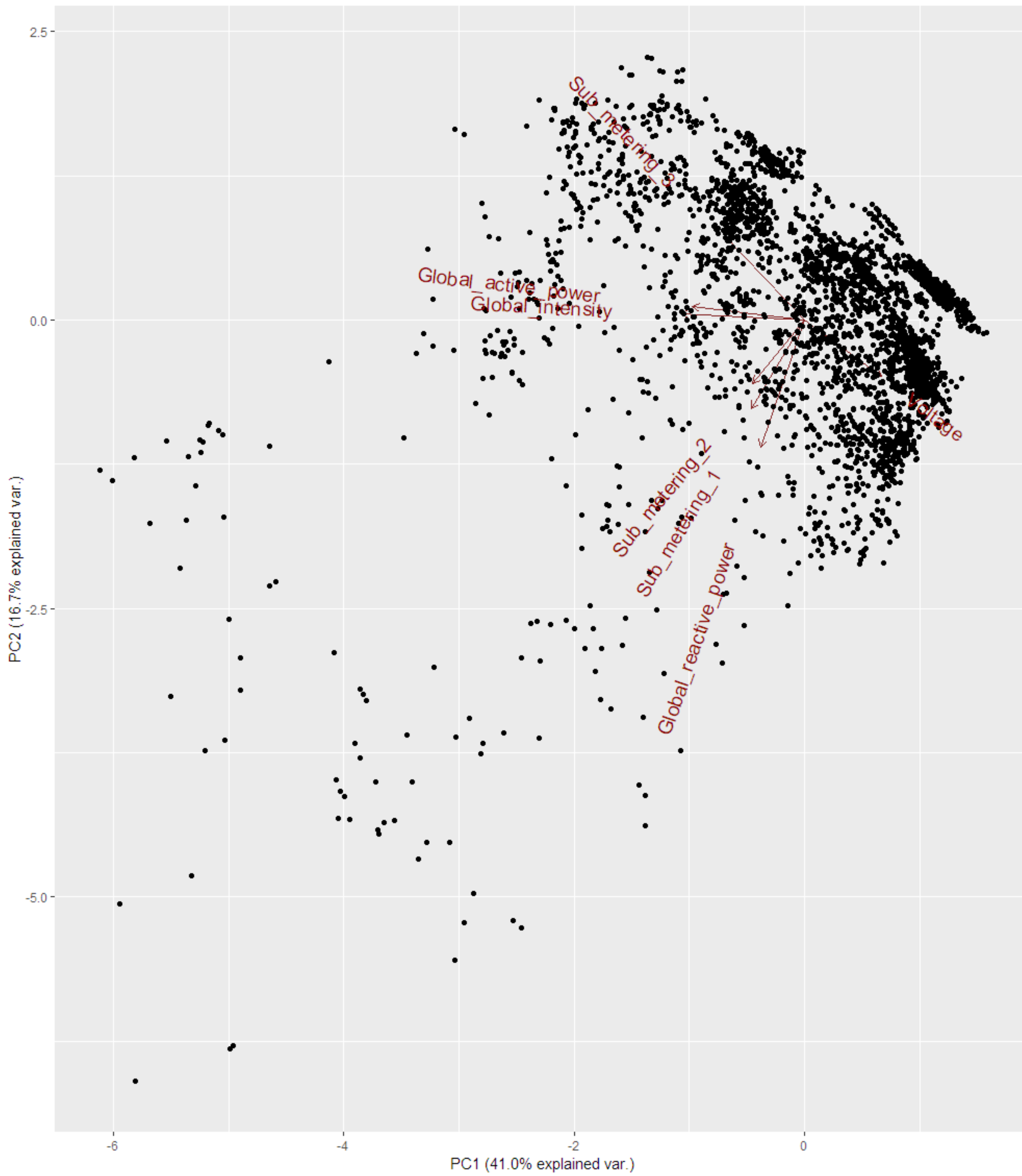
Kyle Isaak:

- Coding of HMMs
- Creating report / proposal
- Presenting

Colin Kirkby:

- Coding of HMMs
- Creating report / proposal / presentation slides

PCA Results:

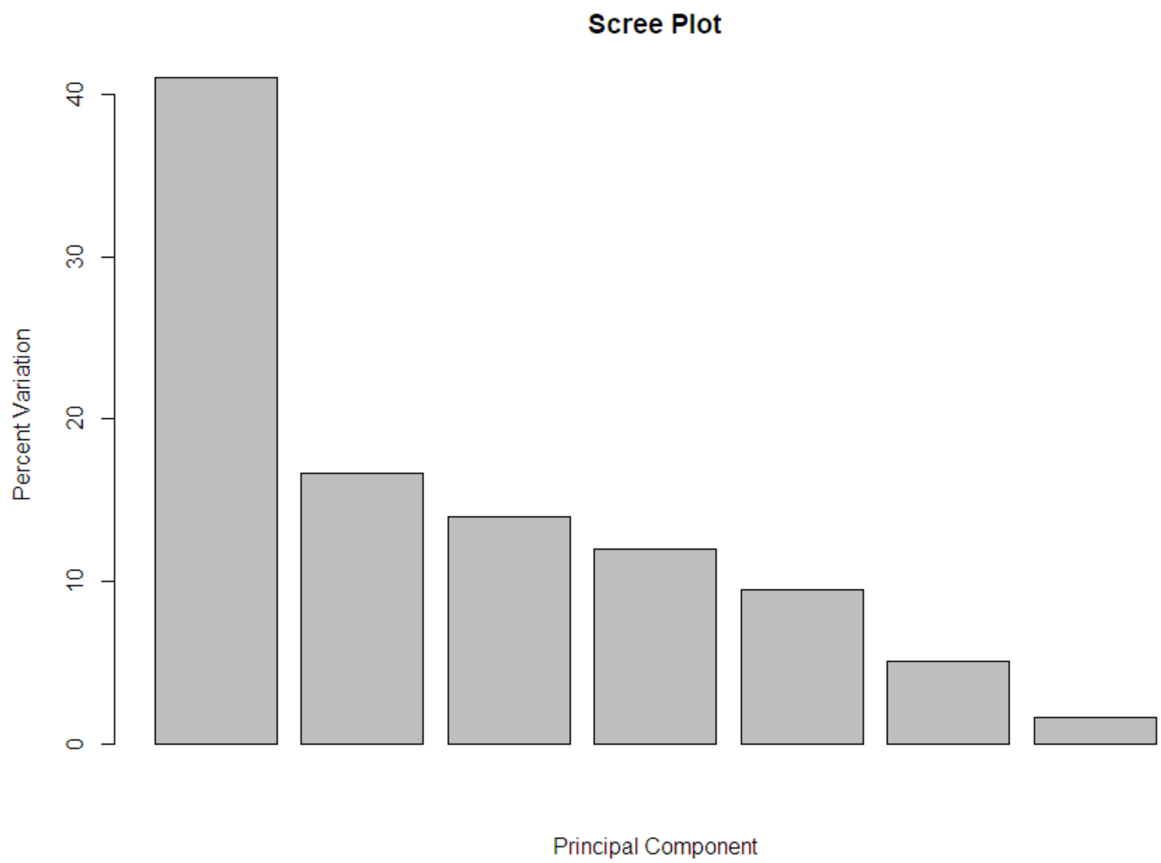


Ranked variable scores

```
Global_intensity    Global_active_power         Sub_metering_3               Voltage
      -0.5664544             -0.5128988             -0.3671846             0.3549207
  Sub_metering_1         Sub_metering_2 Global_reactive_power
      -0.2430386             -0.2367884             -0.2003361
```

PCA conclusions:

From the results of our PCA we can conclude that Global Intensity (GI) and Global Active Power (GAP) are responsible for most of the variance in our data. Sub metering 3 and voltage are also somewhat significant factors but as is visible from the bar plot, PC1 (active power) makes up 40% of the total variation of the dataset and PC2 (intensity) makes up 20% of the total variation of the dataset. This means that global intensity and active power together make up almost 60% of the total variation in the dataset.

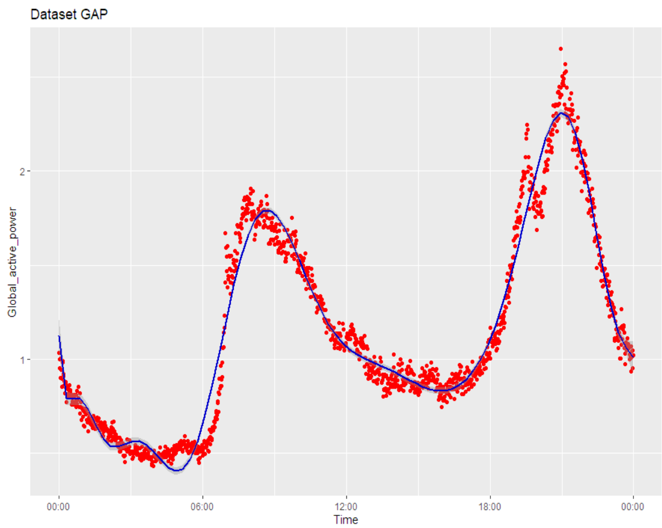


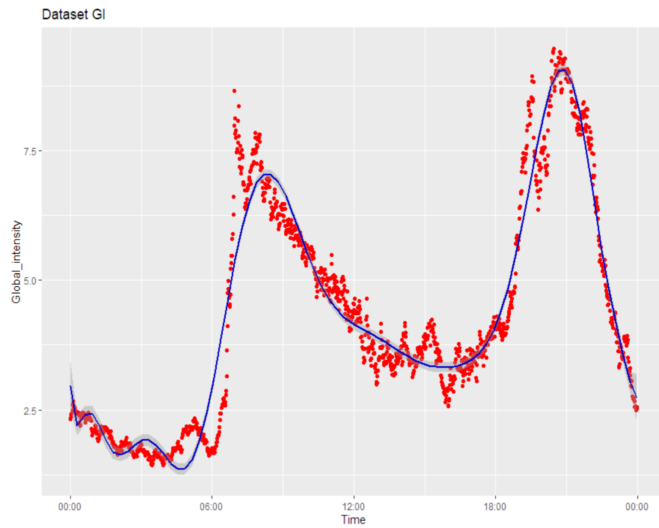
Selection of response variables:

Based on the information found in the PCA, Global Active Power and Global Intensity showed the largest variation. These variables would then be the best indicators to use for creating accurate Hidden Markov Models in addition to being ideal for anomaly detection.

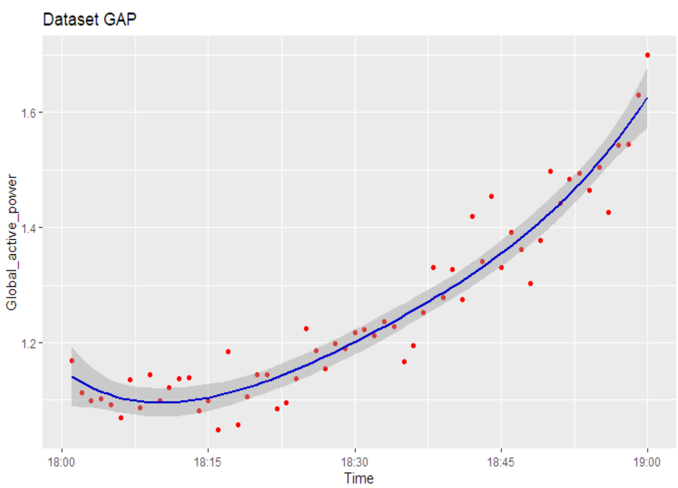
Time window selection:

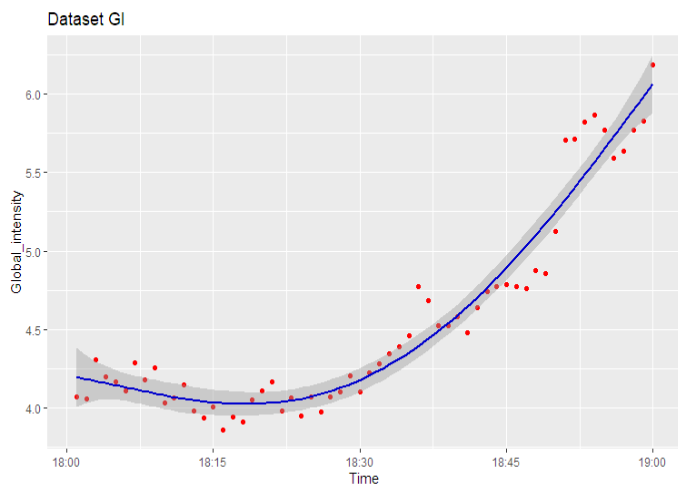
To find our optimal time window we graphed the mean GAP and GI of a Monday.





Using the graphs, the time window from 18:00:00 to 19:00:00 hours was selected. When examining and graphing the data set, we found that there was a severe upwards spike in the global intensity and active power at that time.





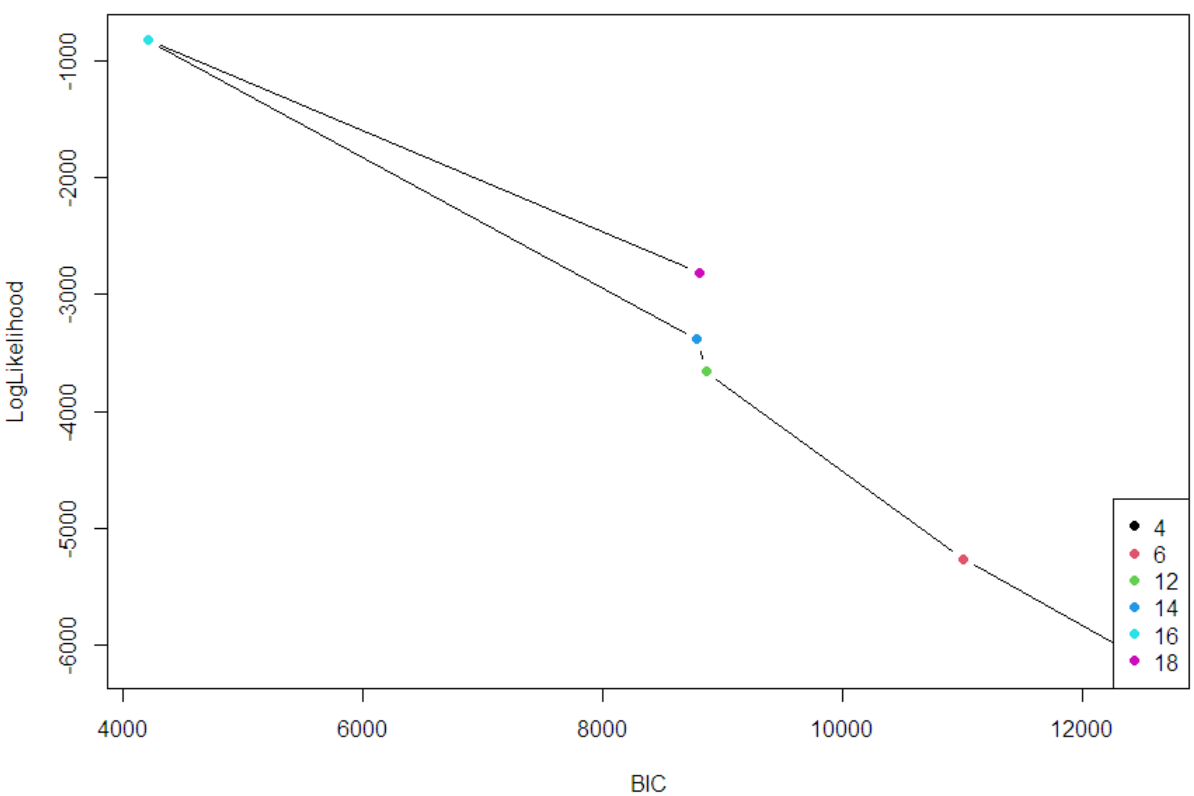
This is an important time period. If there were to be an attack on the power systems attacking at a point where power consumption was already spiking would be more likely to damage critical systems and could potentially have cascading effects. Therefore, we considered that this time period would be a good choice for training our Markov Models so that they have a useful anomaly detection model.

Partitioning of training and test data:

We partitioned our three years of data into a segment of one year for our training data and a segment of two years for our testing data. Following the idea that the testing dataset should be larger than the training set. We feel that a year of data gives a good amount to build an effective model and two years of testing data is sufficient to be able to determine a well fit model.

Model training results:

We trained models using 4 to 18 states to give good coverage. We chose to check every two states as a difference of one state did not make a considerable difference in the outcome of the model.



There was a definitive best model at 16 states that severely outperformed the others in both its log-likelihood and bic values. We chose to take states 12 through 18 to test against the test data as they were all in the range that would create effective models.

4-State model

```
Initial state probabilities model
  pr1   pr2   pr3   pr4
0.175 0.398 0.212 0.216

Transition matrix
        toS1  toS2  toS3  toS4
fromS1 0.988 0.000 0.007 0.004
fromS2 0.001 0.970 0.001 0.028
fromS3 0.015 0.000 0.968 0.017
fromS4 0.010 0.028 0.026 0.936

Response parameters
Resp 1 : gaussian
Resp 2 : gaussian
    Re1.(Intercept) Re1.sd Re2.(Intercept) Re2.sd
St1           2.748  0.888          10.497  4.962
St2           0.395  0.149           1.225  0.498
St3           1.423  0.320           5.984  0.737
St4           0.961  0.517           2.424  1.065
> print(fm1)
Convergence info: Log likelihood converged to within tol. (relative change)
'log Lik.' -6153.433 (df=31)
AIC:  12368.87
BIC:  12556.28
```

6-State Model

```
Initial state probabilities model
  pr1   pr2   pr3   pr4   pr5   pr6
0.148 0.192 0.149 0.254 0.083 0.173

Transition matrix
        toS1  toS2  toS3  toS4  toS5  toS6
fromS1 0.927 0.033 0.012 0.014 0.013 0.001
fromS2 0.013 0.964 0.000 0.000 0.005 0.019
fromS3 0.004 0.000 0.938 0.048 0.010 0.000
fromS4 0.039 0.000 0.035 0.915 0.009 0.002
fromS5 0.012 0.004 0.009 0.021 0.949 0.005
fromS6 0.000 0.010 0.000 0.000 0.002 0.988

Response parameters
Resp 1 : gaussian
Resp 2 : gaussian
    Re1.(Intercept) Re1.sd Re2.(Intercept) Re2.sd
St1           0.809  0.357           3.516  1.161
St2           1.484  0.246           6.036  0.685
St3           0.380  0.153           0.818  0.269
St4           0.410  0.144           1.617  0.340
St5           1.578  0.974           1.491  0.582
St6           2.732  0.872          11.544  4.185
> print(fm2)
Convergence info: Log likelihood converged to within tol. (relative change)
'log Lik.' -5264.335 (df=59)
AIC:  10646.67
BIC:  11003.36
```

Both the 4 state and 6 state models seem to be underfit with very low likelihood. We do not consider these to be good models for the data and we decided that they were not worth running on the test data.

12-State Model

```
Initial state probabilities model
  pr1   pr2   pr3   pr4   pr5   pr6   pr7   pr8   pr9  pr10  pr11  pr12
0.080 0.057 0.148 0.040 0.140 0.054 0.138 0.019 0.183 0.096 0.020 0.025

Transition matrix
        toS1  toS2  toS3  toS4  toS5  toS6  toS7  toS8  toS9 toS10 toS11 toS12
fromS1  0.873 0.000 0.000 0.005 0.060 0.000 0.000 0.000 0.052 0.001 0.006 0.003
fromS2  0.000 0.951 0.022 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.028 0.000
fromS3  0.000 0.003 0.930 0.000 0.000 0.030 0.004 0.003 0.000 0.000 0.000 0.030
fromS4  0.020 0.000 0.000 0.800 0.000 0.000 0.058 0.031 0.006 0.040 0.044 0.000
fromS5  0.053 0.000 0.000 0.000 0.918 0.000 0.000 0.000 0.011 0.006 0.012 0.000
fromS6  0.000 0.000 0.037 0.003 0.000 0.936 0.020 0.000 0.000 0.000 0.004 0.000
fromS7  0.000 0.000 0.012 0.027 0.000 0.048 0.895 0.004 0.000 0.009 0.005 0.000
fromS8  0.000 0.000 0.042 0.024 0.000 0.000 0.012 0.908 0.000 0.015 0.000 0.000
fromS9  0.073 0.000 0.000 0.019 0.008 0.000 0.000 0.012 0.831 0.047 0.011 0.000
fromS10 0.000 0.004 0.004 0.029 0.015 0.000 0.000 0.016 0.016 0.909 0.007 0.000
fromS11 0.017 0.024 0.000 0.011 0.015 0.000 0.005 0.003 0.016 0.007 0.901 0.000
fromS12 0.000 0.000 0.014 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.986

Response parameters
Resp 1 : gaussian
Resp 2 : gaussian
     Re1.(Intercept) Re1.sd Re2.(Intercept) Re2.sd
St1            0.414  0.153           1.339  0.166
St2            2.499  0.904           1.594  0.806
St3            2.247  0.492           8.602  1.852
St4            0.978  0.386           4.114  0.425
St5            0.374  0.153           0.748  0.229
St6            1.589  0.150           6.191  0.337
St7            1.379  0.165           5.429  0.299
St8            1.006  0.389           6.486  0.904
St9            0.402  0.128           1.909  0.244
St10           0.749  0.325           2.823  0.374
St11           1.021  0.479           1.448  0.399
St12           3.301  0.877          15.022  3.601
> print(fm3)
Convergence info: Log likelihood converged to within tol. (relative change)
'log Lik.' -3661.71 (df=191)
AIC:  7705.419
BIC:  8860.127
```

The 12 state model offers a decent log-likelihood and an acceptable BIC but neither are particularly good.

14-State Model

```
Initial state probabilities model
  pr1   pr2   pr3   pr4   pr5   pr6   pr7   pr8   pr9  pr10  pr11  pr12  pr13  pr14
0.103 0.058 0.137 0.037 0.152 0.136 0.000 0.021 0.058 0.025 0.195 0.058 0.021 0.000

Transition matrix
        toS1  toS2  toS3  toS4  toS5  toS6  toS7  toS8  toS9 toS10 toS11 toS12 toS13 toS14
fromS1  0.897 0.006 0.000 0.000 0.006 0.000 0.032 0.000 0.010 0.000 0.000 0.038 0.011 0.000
fromS2  0.000 0.953 0.000 0.000 0.009 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.029 0.008
fromS3  0.000 0.000 0.919 0.000 0.000 0.000 0.005 0.000 0.000 0.009 0.055 0.000 0.012 0.000
fromS4  0.000 0.000 0.000 0.926 0.010 0.044 0.000 0.000 0.005 0.000 0.000 0.000 0.005 0.010
fromS5  0.000 0.003 0.000 0.009 0.947 0.007 0.000 0.031 0.000 0.000 0.000 0.000 0.000 0.003
fromS6  0.000 0.000 0.000 0.054 0.004 0.906 0.000 0.000 0.011 0.005 0.000 0.000 0.005 0.016
fromS7  0.025 0.000 0.008 0.000 0.000 0.000 0.836 0.000 0.000 0.048 0.083 0.000 0.000 0.000
fromS8  0.000 0.000 0.000 0.000 0.010 0.000 0.000 0.986 0.000 0.000 0.000 0.000 0.000 0.004
fromS9  0.012 0.000 0.000 0.012 0.006 0.000 0.000 0.000 0.876 0.000 0.000 0.018 0.034 0.043
fromS10 0.009 0.000 0.044 0.000 0.000 0.000 0.020 0.000 0.023 0.874 0.019 0.000 0.000 0.011
fromS11 0.001 0.000 0.053 0.000 0.000 0.000 0.050 0.000 0.000 0.008 0.875 0.002 0.008 0.002
fromS12 0.036 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.084 0.000 0.000 0.861 0.000 0.019
fromS13 0.012 0.020 0.014 0.000 0.000 0.000 0.000 0.000 0.021 0.000 0.017 0.002 0.914 0.000
fromS14 0.011 0.000 0.000 0.000 0.044 0.014 0.000 0.006 0.009 0.000 0.000 0.000 0.000 0.917

Response parameters
Resp 1 : gaussian
Resp 2 : gaussian
     Re1.(Intercept) Re1.sd Re2.(Intercept) Re2.sd
St1            0.738  0.362           2.583  0.257
St2            2.518  0.905           1.606  0.830
St3            0.375  0.153           0.750  0.230
St4            1.593  0.122           6.293  0.282
St5            2.318  0.482           8.933  1.902
St6            1.457  0.158           5.637  0.159
St7            0.396  0.134           1.945  0.209
St8            3.379  0.796          15.213  3.590
St9            1.240  0.216           4.656  0.586
St10           0.509  0.217           4.633  1.840
St11           0.414  0.149           1.368  0.184
St12           0.843  0.189           3.218  0.174
St13           1.019  0.494           1.464  0.400
St14           1.460  0.403           6.764  0.808
> print(fm4)
Convergence info: Log likelihood converged to within tol. (relative change)
'log Lik.' -3382.989 (df=251)
AIC:  7267.979
BIC:  8785.422
```

The 14 state model presents a good log-likelihood and a very reasonable BIC but it is not a significant improvement over the 12 state model.

16-State Model

```
Initial state probabilities model
  pr1   pr2   pr3   pr4   pr5   pr6   pr7   pr8   pr9  pr10  pr11  pr12  pr13  pr14  pr15  pr16
0.038 0.034 0.096 0.000 0.140 0.056 0.024 0.148 0.106 0.020 0.124 0.000 0.000 0.130 0.058 0.026

Transition matrix
        toS1  toS2  toS3  toS4  toS5  toS6  toS7  toS8  toS9 toS10 toS11 toS12 toS13 toS14 toS15 toS16
fromS1  0.733 0.069 0.000 0.072 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.042 0.000 0.000 0.024 0.061
fromS2  0.255 0.589 0.000 0.000 0.000 0.000 0.000 0.000 0.032 0.000 0.000 0.032 0.000 0.000 0.028 0.064
fromS3  0.000 0.000 0.875 0.000 0.058 0.000 0.000 0.000 0.000 0.006 0.000 0.003 0.006 0.040 0.004 0.008
fromS4  0.119 0.019 0.000 0.706 0.049 0.000 0.000 0.000 0.000 0.000 0.000 0.020 0.000 0.087 0.000 0.000
fromS5  0.000 0.000 0.055 0.006 0.918 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.009 0.000 0.012
fromS6  0.000 0.000 0.000 0.000 0.000 0.931 0.000 0.017 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.052
fromS7  0.000 0.000 0.000 0.000 0.000 0.000 0.986 0.014 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
fromS8  0.000 0.000 0.000 0.000 0.000 0.003 0.030 0.936 0.000 0.003 0.029 0.000 0.000 0.000 0.000 0.000
fromS9  0.005 0.006 0.000 0.000 0.000 0.000 0.000 0.015 0.893 0.003 0.041 0.000 0.000 0.000 0.017 0.020
fromS10 0.000 0.007 0.000 0.000 0.000 0.000 0.000 0.031 0.016 0.924 0.000 0.000 0.022 0.000 0.000 0.000
fromS11 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.031 0.008 0.000 0.952 0.000 0.002 0.000 0.000 0.006
fromS12 0.072 0.041 0.000 0.000 0.000 0.024 0.000 0.000 0.000 0.000 0.000 0.794 0.000 0.000 0.069 0.000
fromS13 0.000 0.000 0.103 0.068 0.010 0.000 0.000 0.020 0.000 0.094 0.000 0.000 0.698 0.007 0.000 0.000
fromS14 0.000 0.000 0.093 0.038 0.004 0.000 0.000 0.000 0.000 0.005 0.000 0.000 0.035 0.812 0.006 0.009
fromS15 0.000 0.002 0.000 0.000 0.000 0.000 0.000 0.000 0.053 0.039 0.000 0.031 0.000 0.000 0.862 0.012
fromS16 0.023 0.000 0.018 0.000 0.018 0.023 0.000 0.000 0.019 0.001 0.000 0.001 0.000 0.017 0.000 0.880

Response parameters
Resp 1 : gaussian
Resp 2 : gaussian
     Re1.(Intercept) Re1.sd Re2.(Intercept) Re2.sd
St1            0.842  0.532           2.400  0.000
St2            1.133  0.295           2.796  0.444
St3            0.415  0.150           1.364  0.179
St4            0.366  0.151           2.848  0.394
St5            0.375  0.153           0.751  0.228
St6            2.588  0.928           1.507  0.826
St7            3.351  0.799          15.099  3.587
St8            2.234  0.489           8.544  1.865
St9            1.263  0.200           4.835  0.580
St10           0.999  0.399           6.614  1.127
St11           1.540  0.150           6.013  0.399
St12           0.702  0.117           2.755  0.138
St13           0.472  0.174           4.014  0.731
St14           0.404  0.134           1.922  0.189
St15           0.874  0.209           3.357  0.295
St16           1.067  0.538           1.471  0.427
> print(fm5)
Convergence info: likelihood decreased in EM iteration; stopped without convergence.
'log Lik.' -820.9077 (df=319)
AIC:  2279.815
BIC:  4208.358
```

The 16 state model has the largest improvement in both log-likelihood and also BIC values of all the models trained; this model has the best fit to the training data.

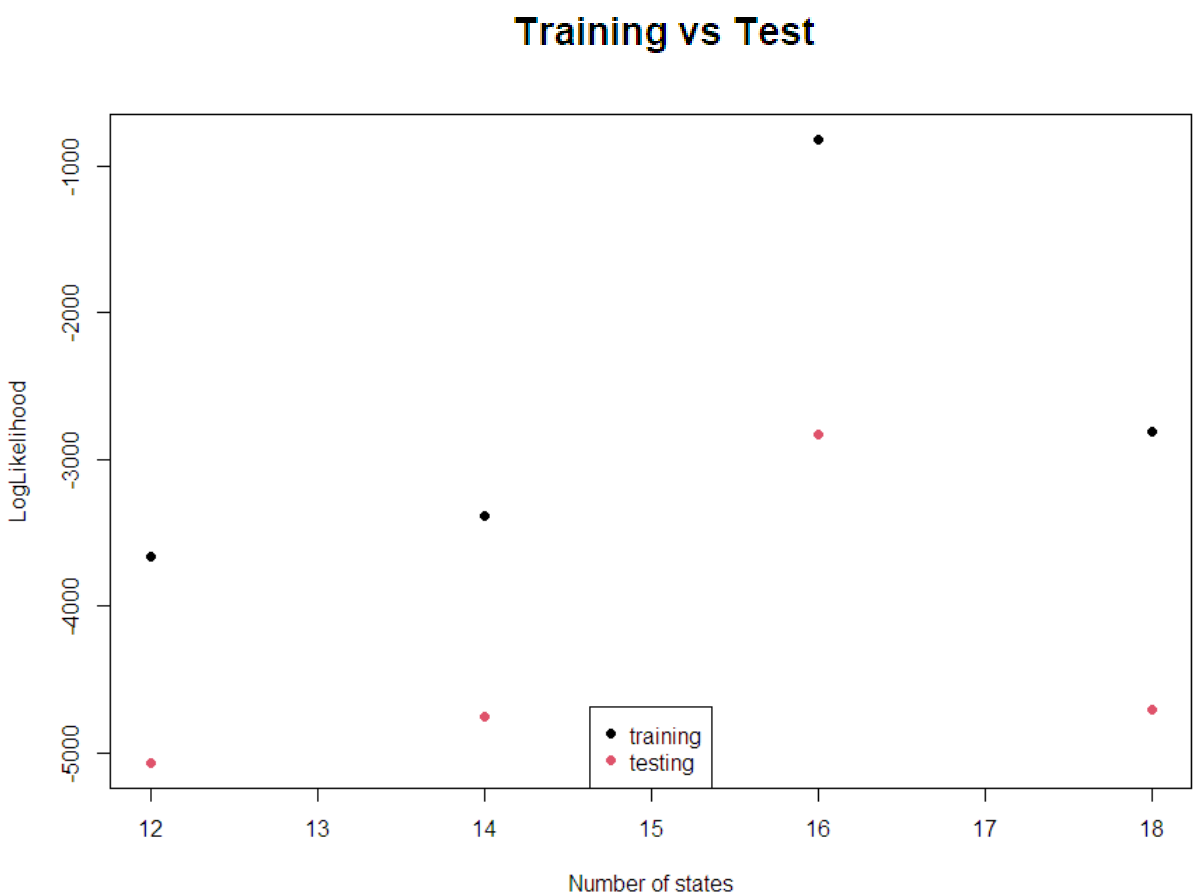
18-State Model

```
Initial state probabilities model
  pr1   pr2   pr3   pr4   pr5   pr6   pr7   pr8   pr9  pr10  pr11  pr12  pr13  pr14  pr15  pr16  pr17  pr18
0.019 0.000 0.058 0.227 0.135 0.138 0.074 0.058 0.057 0.000 0.038 0.000 0.019 0.059 0.019 0.020 0.019 0.060

Transition matrix
        toS1  toS2  toS3  toS4  toS5  toS6  toS7  toS8  toS9 toS10 toS11 toS12 toS13 toS14 toS15 toS16 toS17 toS18
fromS1  0.976 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.019 0.000 0.005 0.000 0.000 0.000 0.000 0.000 0.000 0.000
fromS2  0.000 0.837 0.097 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.011 0.000 0.000 0.000 0.028 0.027
fromS3  0.000 0.058 0.789 0.000 0.012 0.000 0.000 0.021 0.000 0.000 0.000 0.000 0.025 0.025 0.014 0.057 0.000 0.000
fromS4  0.000 0.000 0.000 0.906 0.000 0.048 0.000 0.004 0.000 0.000 0.000 0.032 0.000 0.000 0.002 0.005 0.000 0.002
fromS5  0.000 0.012 0.005 0.000 0.902 0.000 0.000 0.000 0.000 0.011 0.005 0.000 0.060 0.000 0.000 0.005 0.000 0.000
fromS6  0.000 0.000 0.000 0.055 0.000 0.920 0.000 0.000 0.000 0.000 0.000 0.013 0.000 0.000 0.000 0.012 0.000 0.000
fromS7  0.000 0.000 0.000 0.000 0.000 0.000 0.946 0.000 0.000 0.022 0.000 0.000 0.000 0.000 0.000 0.032 0.000 0.000
fromS8  0.000 0.009 0.064 0.000 0.000 0.000 0.000 0.849 0.000 0.000 0.000 0.000 0.000 0.000 0.038 0.000 0.000 0.040
fromS9  0.063 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.903 0.012 0.023 0.000 0.000 0.000 0.000 0.000 0.000 0.000
fromS10 0.002 0.000 0.000 0.000 0.016 0.000 0.000 0.000 0.038 0.891 0.000 0.000 0.028 0.000 0.000 0.000 0.025 0.000
fromS11 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.046 0.000 0.914 0.000 0.000 0.000 0.008 0.000 0.032 0.000
fromS12 0.000 0.000 0.000 0.054 0.000 0.025 0.005 0.000 0.000 0.000 0.000 0.881 0.000 0.010 0.019 0.000 0.000 0.005
fromS13 0.000 0.003 0.008 0.000 0.075 0.000 0.000 0.000 0.000 0.024 0.007 0.008 0.843 0.000 0.000 0.000 0.032 0.000
fromS14 0.016 0.000 0.000 0.000 0.000 0.000 0.014 0.000 0.035 0.000 0.031 0.027 0.000 0.841 0.035 0.000 0.000 0.000
fromS15 0.000 0.018 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.009 0.000 0.057 0.906 0.000 0.000 0.009
fromS16 0.000 0.004 0.020 0.006 0.000 0.013 0.023 0.002 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.920 0.000 0.012
fromS17 0.000 0.000 0.000 0.000 0.021 0.000 0.000 0.000 0.000 0.019 0.031 0.000 0.032 0.000 0.000 0.010 0.887 0.000
fromS18 0.000 0.000 0.016 0.000 0.000 0.000 0.000 0.046 0.000 0.000 0.000 0.023 0.000 0.000 0.000 0.009 0.000 0.906

Response parameters
Resp 1 : gaussian
Resp 2 : gaussian
     Re1.(Intercept) Re1.sd Re2.(Intercept) Re2.sd
St1            3.499  0.816          15.791  3.651
St2            1.228  0.066           5.076  0.236
St3            1.160  0.320           4.155  0.411
St4            0.413  0.144           1.439  0.232
St5            1.463  0.162           5.630  0.156
St6            0.374  0.153           0.753  0.232
St7            2.437  0.915           1.638  0.832
St8            0.844  0.169           3.208  0.167
St9            2.568  0.305          11.106  1.354
St10           2.530  0.663           5.927  0.430
St11           2.008  0.184           8.293  0.575
St12           0.391  0.142           2.556  0.768
St13           1.552  0.104           6.105  0.123
St14           1.406  0.560           8.382  1.255
St15           1.008  0.380           6.200  0.694
St16           0.956  0.484           1.457  0.401
St17           1.642  0.138           6.593  0.228
St18           0.836  0.325           2.539  0.214
> print(fm6)
Convergence info: Log likelihood converged to within tol. (relative change)
'log Lik.' -2814.174 (df=395)
AIC:  6418.347
BIC:  8806.355
```

The 18 state model sees a fairly significant drop in both likelihood and BIC. This brings it close to the quality of the 12 and 14 state models, but it is at a risk of being overfit to the data. We ran into some issues with training the 18 state model where on some runs of the program, it would vary quite significantly in both its log-likelihood and its BIC values. This was another reason that we chose to avoid using it for our anomaly detection because a model that has the potential to vary significantly is not useful or reliable. Excluding this model should prevent false alarms while still minimizing the risk of missing an anomaly.

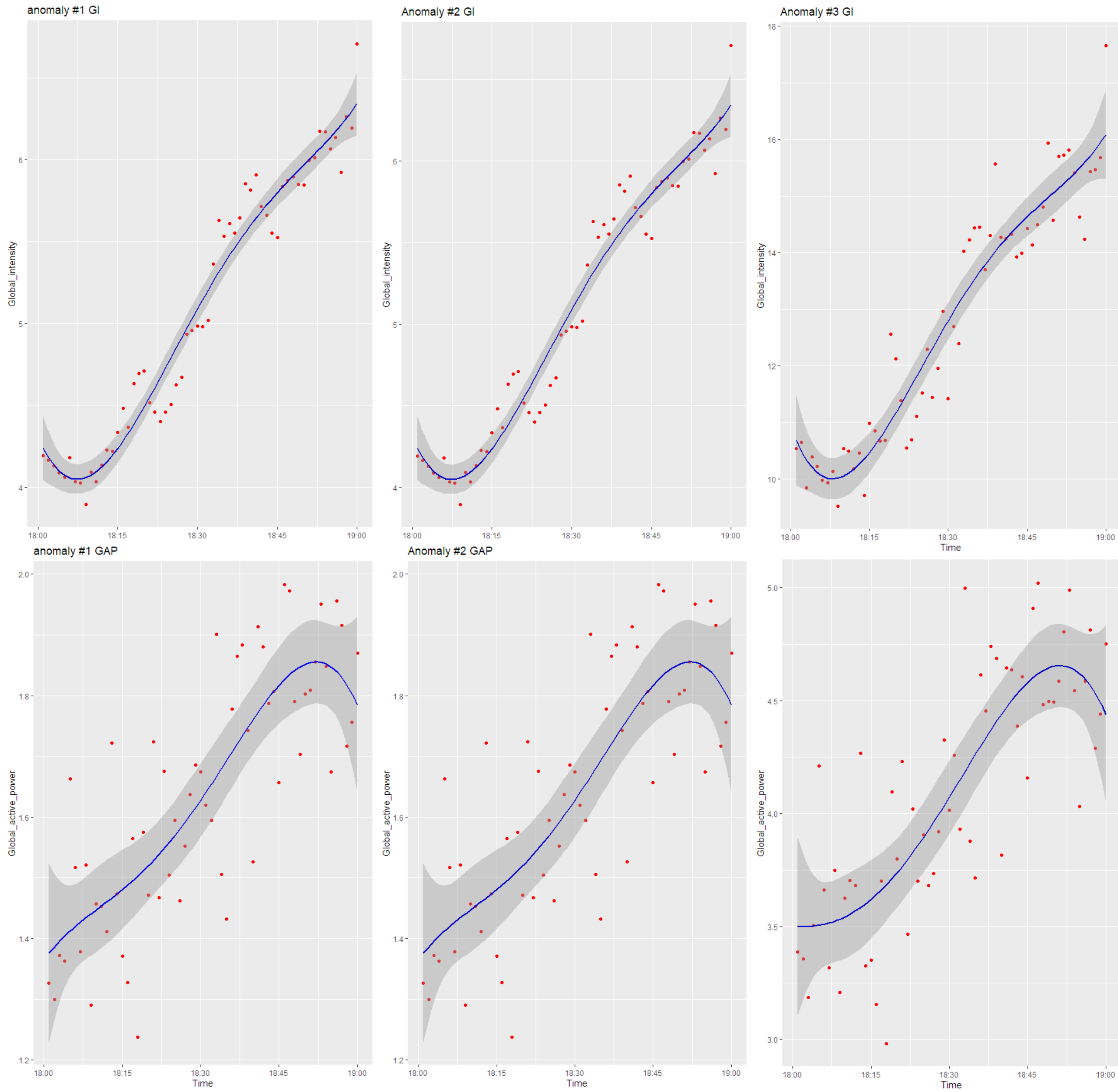
Normalized log-likelihood of training and test data:

When running our models against our testing data, we found that the 16 state model still performed the best and our 18 state model appeared to be a possible case of overfitting to the training data.



We ran all the models through the set pars function of depmix using an unfitted model created from the test data and then used the forward backwards functionality to calculate the log-likelihood of the model. For our results, we scaled our testing results by a factor of 52/100 due to our testing dataset being roughly twice the size of the training set. From these results, we chose to use the 16 state model as our anomaly detection model.

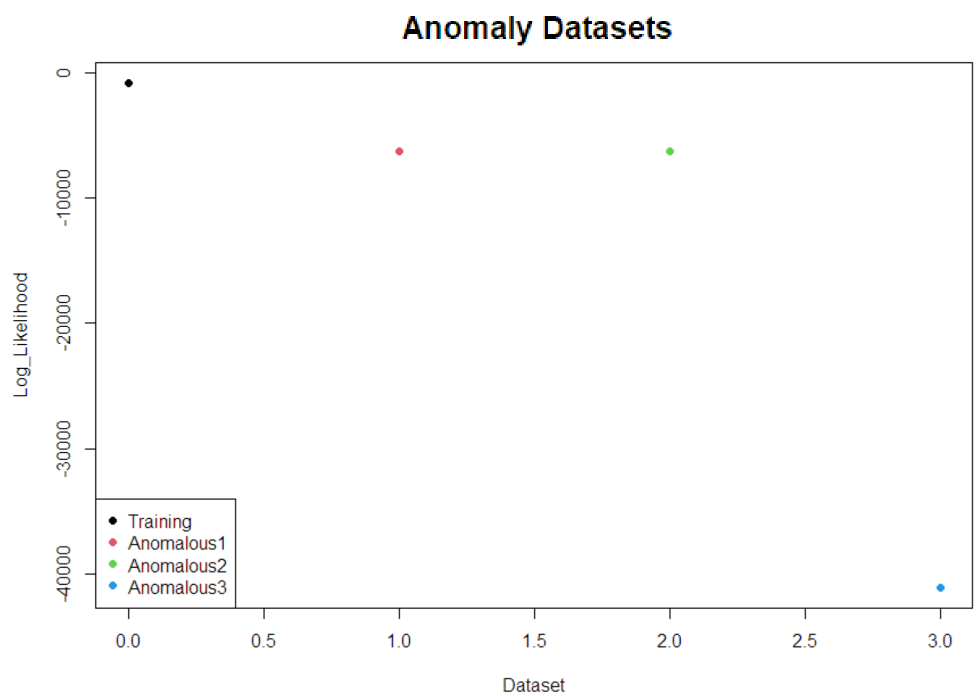
Anomaly detection results:



These graphs (see appendix pages 24-29 for larger versions) are the mean values of the anomaly data sets that we ran our detection algorithm on. From these figures, it is apparent that dataset 1 and 2 are very similar (if not identical), but data set 3 is significantly higher than the other two, as well as our original training dataset.

We used the same methodology when running our models against the anomaly data as we did for running them against the test data.

All three datasets were run against the 16 state model that we created. Data sets 1 and 2 both gave the same log-likelihood (-10,153). After some further analysis, we found that the data given in those two sets was identical, but reordered. This explains the matching results.



For data set 3, our model returned a log-likelihood value of -40,613. This was significantly lower than the previous two and within the range that we would consider largely anomalous. For these reasons, we decided to further investigate the 3rd dataset to find the origin of the anomalies. This was done by running the same model on a day by day basis rather than over an entire year.

We scaled the data by a factor of 52 to account for the smaller data set size. From this, we found that most of the days fell within a standard range between slightly below 0 to approximately -9,000 with only three data points falling outside of that range. If we were doing an actual risk assessment, we would consider those three days to be worth investigating and possible indications of malicious activity or potential risks to the system.

Conclusions:

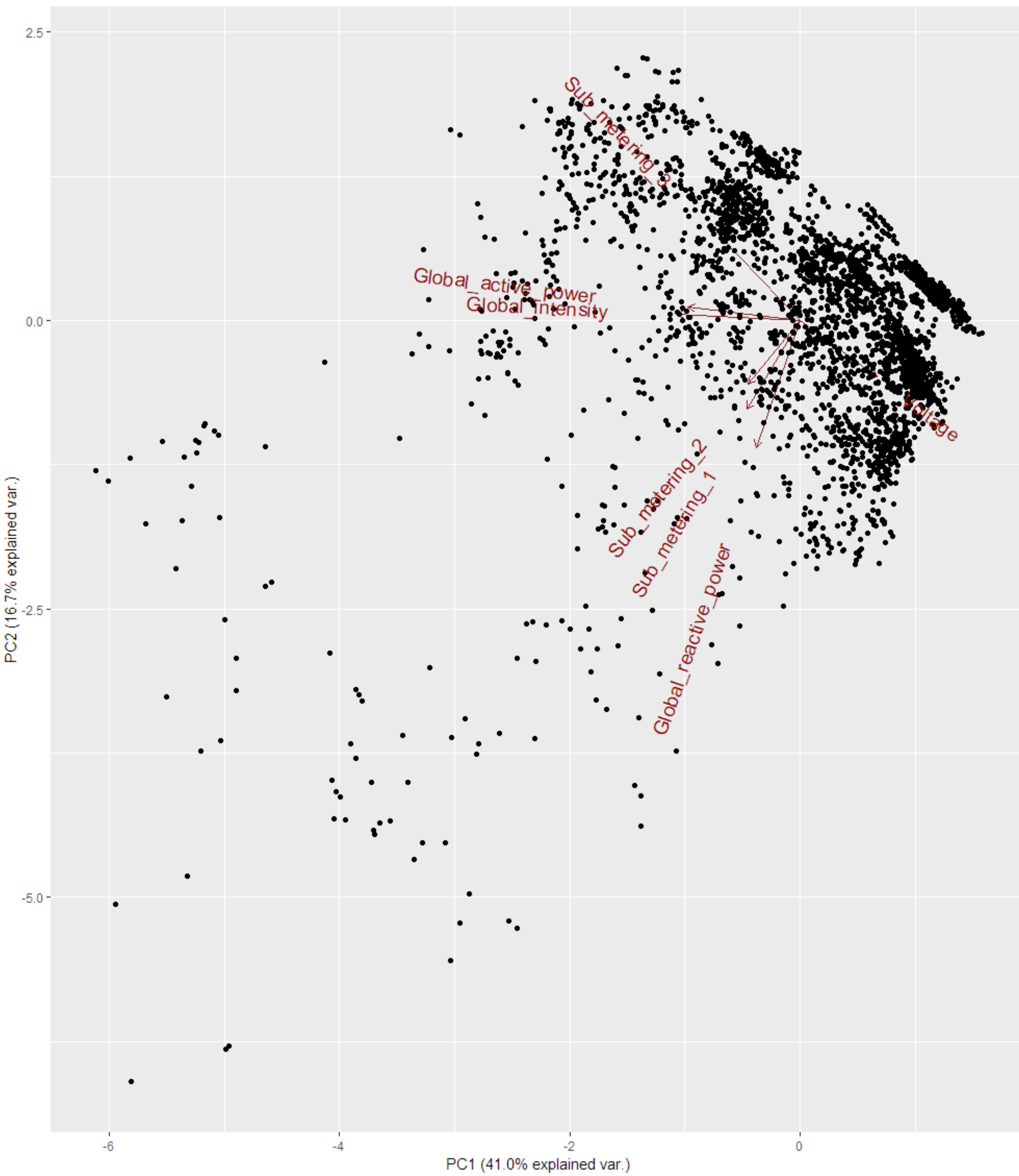
HMMS are a valuable tool that can be used to detect anomalies within data. This can be used for reasons such as: intrusion detection, alarm systems, and system testing. The HMMS that we created performed well with both large and small datasets. This means it could be used for auditing purposes on a yearly basis, or to detect anomalies on a day by day basis. With some further fine tuning, the HMM could take incoming data and detect anomalies in near real time. This would be done by looking at the incoming data along with data received earlier in the day, week, etc. to run the test. Although we designed and trained our HMMs using electrical consumption data with protection of a power grid in mind, they could also be re-trained and used in other systems to improve security. A few examples of this would be monitoring a device or network's incoming connections (most likely the number of connections at any given time) to prevent DDoS attacks, or monitoring banking transactions for potentially fraudulent activity.

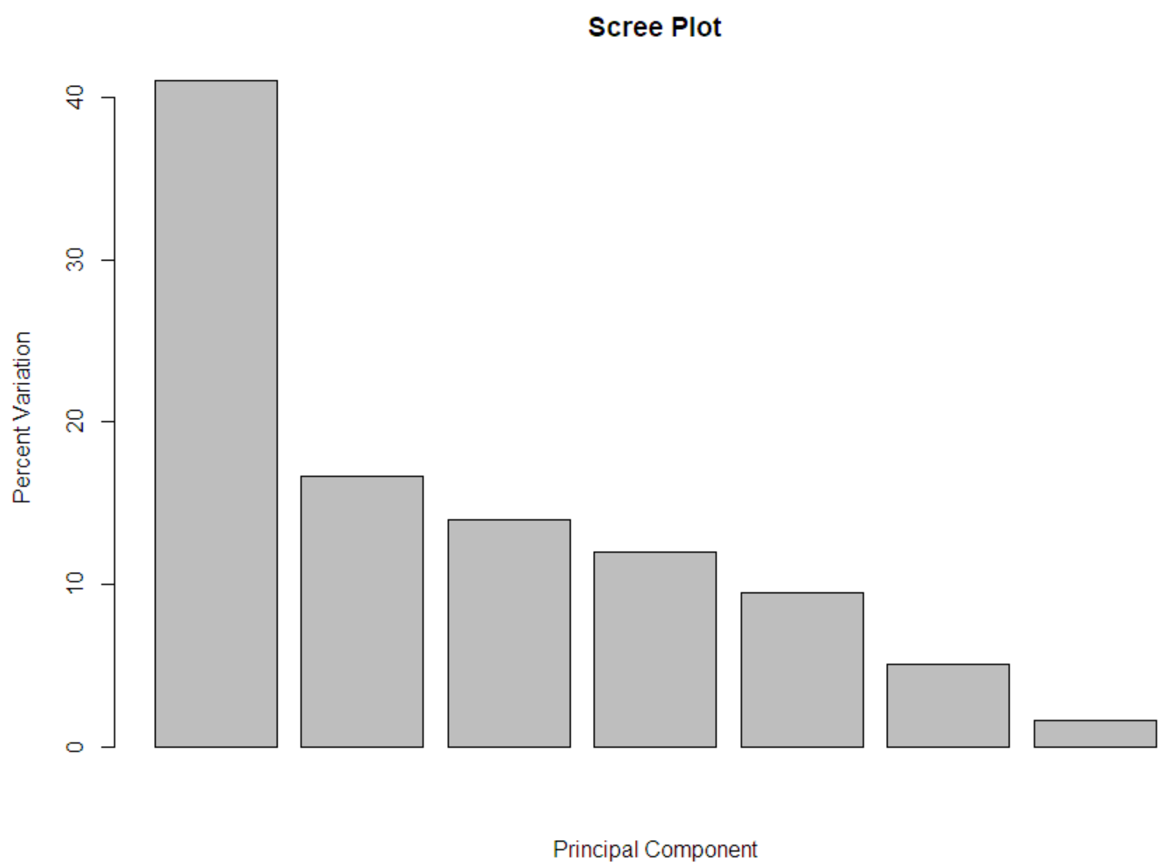
Future work regarding this cybersecurity model would include improving the algorithms to work with real-time data and finding ways to adapt it to new systems and datasets.

Appendix

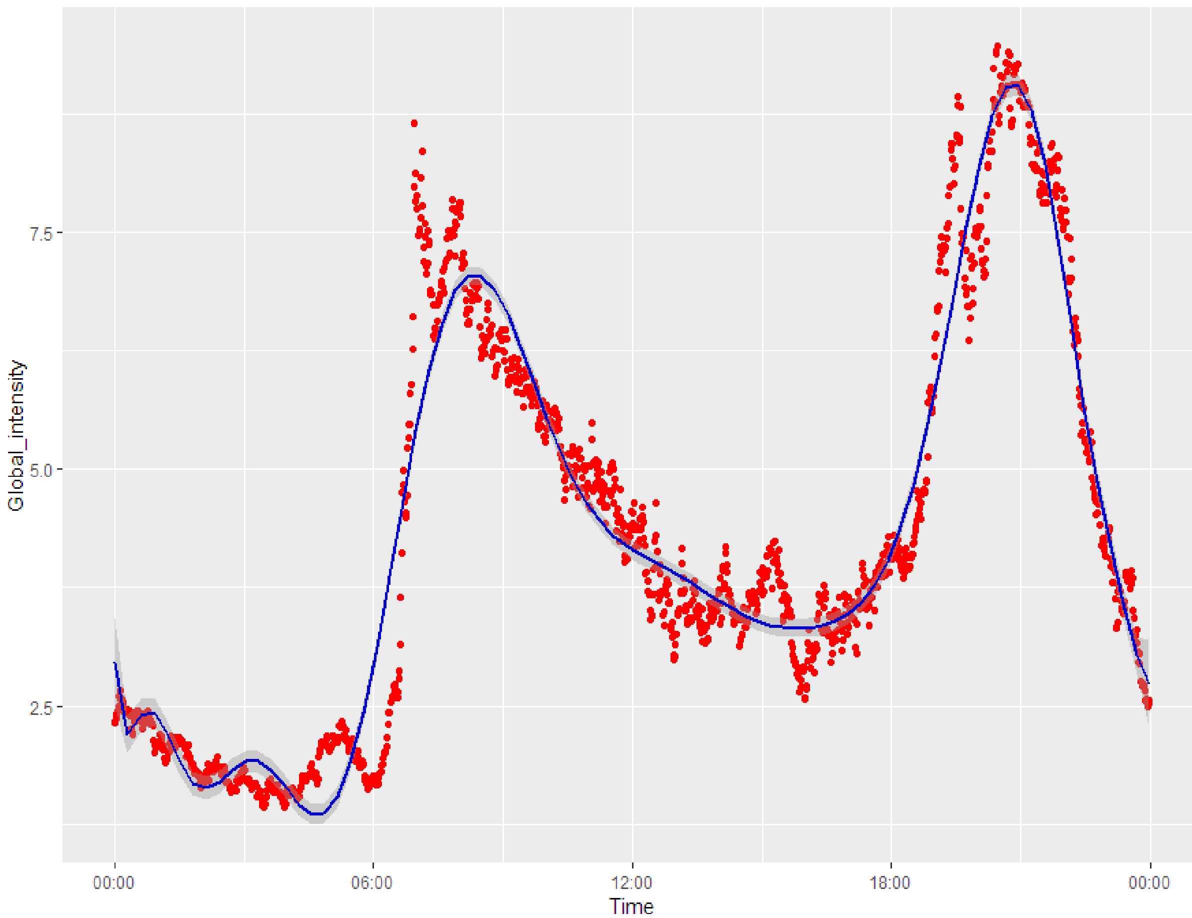
The following pages contain enlarged versions of all figures for improved readability.

The figures appear in the following order:

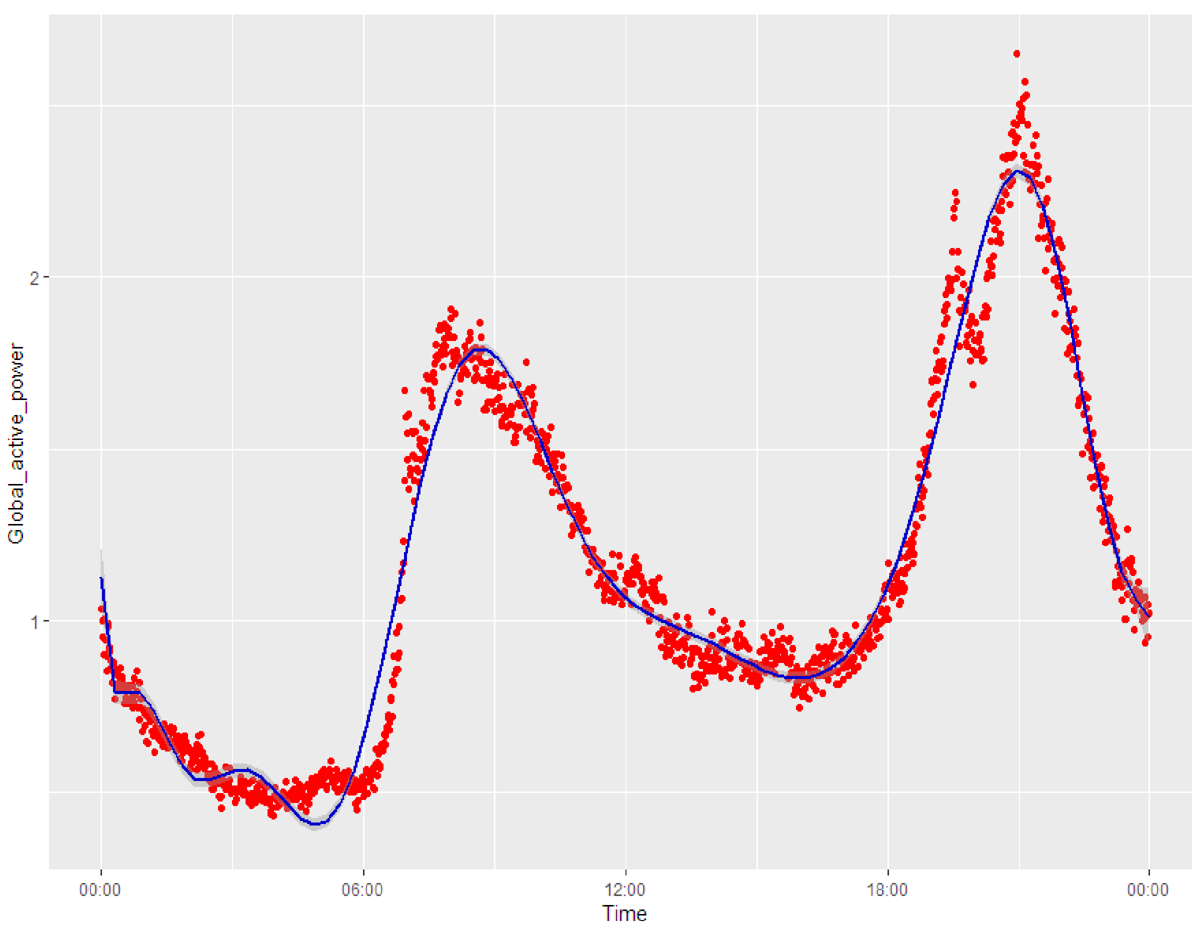

Sub_metering_3
Global_active_power
Global_intensity
Voltage
Sub_metering_2
Sub_metering_1
Global_reactive_power
PC2 (16.7% explained var.)
PC1 (41.0% explained var.)
2.5
0.0
-2.5
-5.0
-6
-4
-2
0

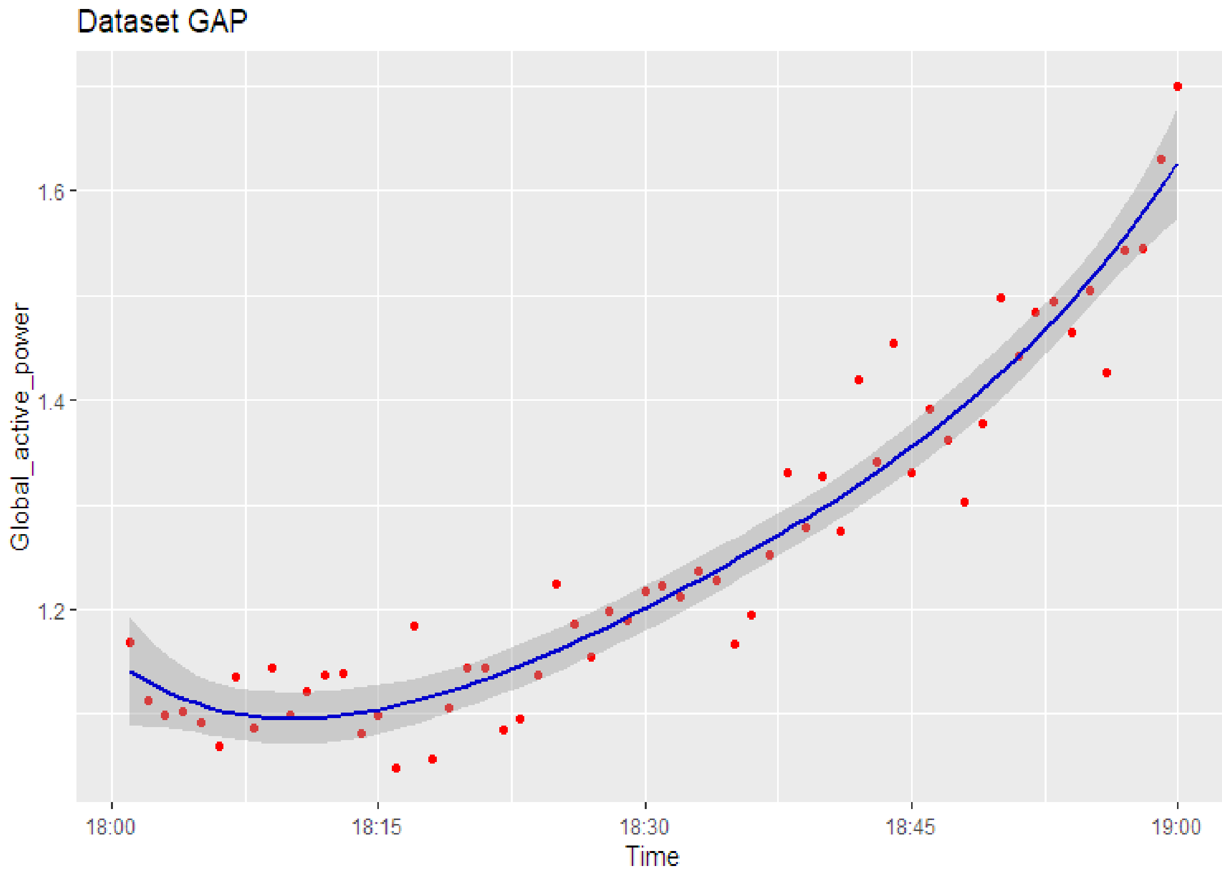

Scree Plot
Percent Variation
0
10
20
30
40
Principal Component

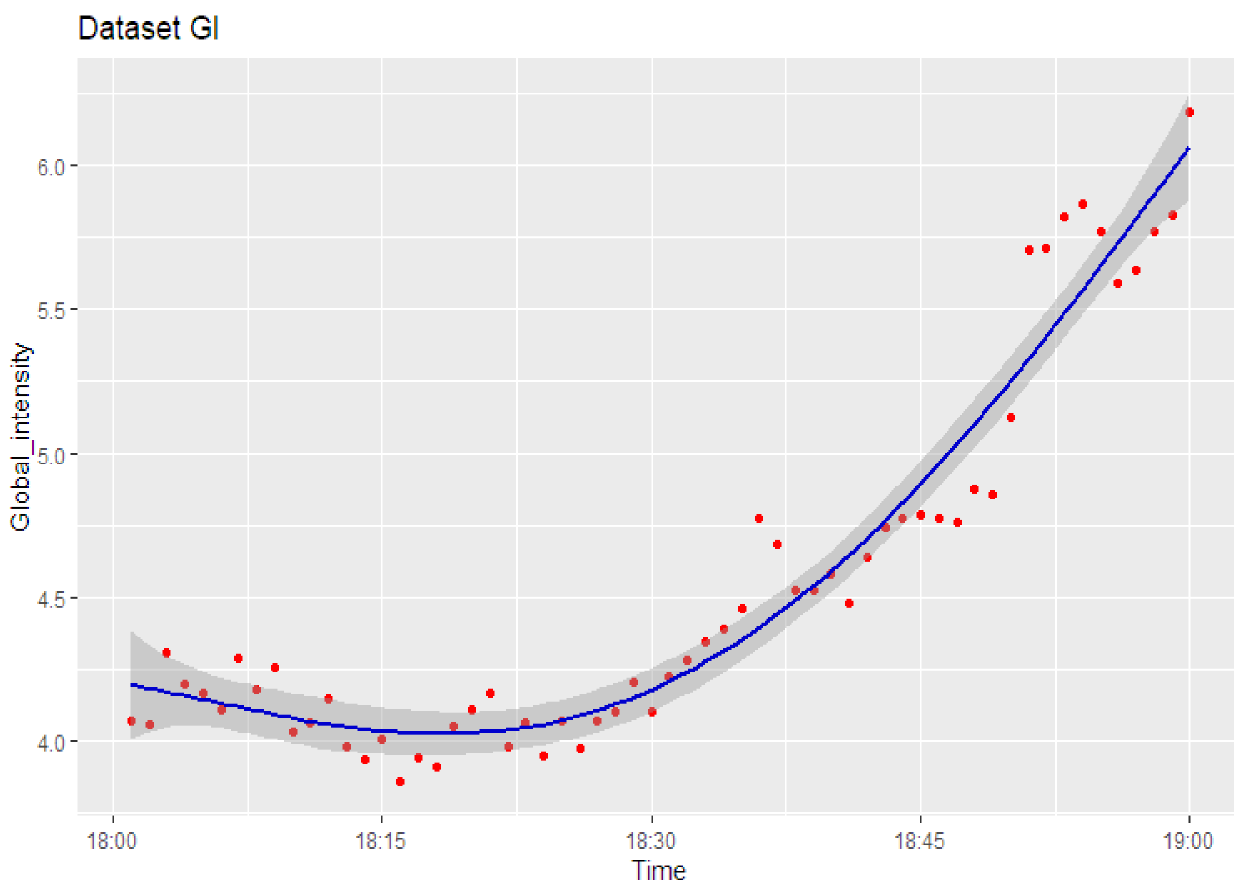
Dataset GI

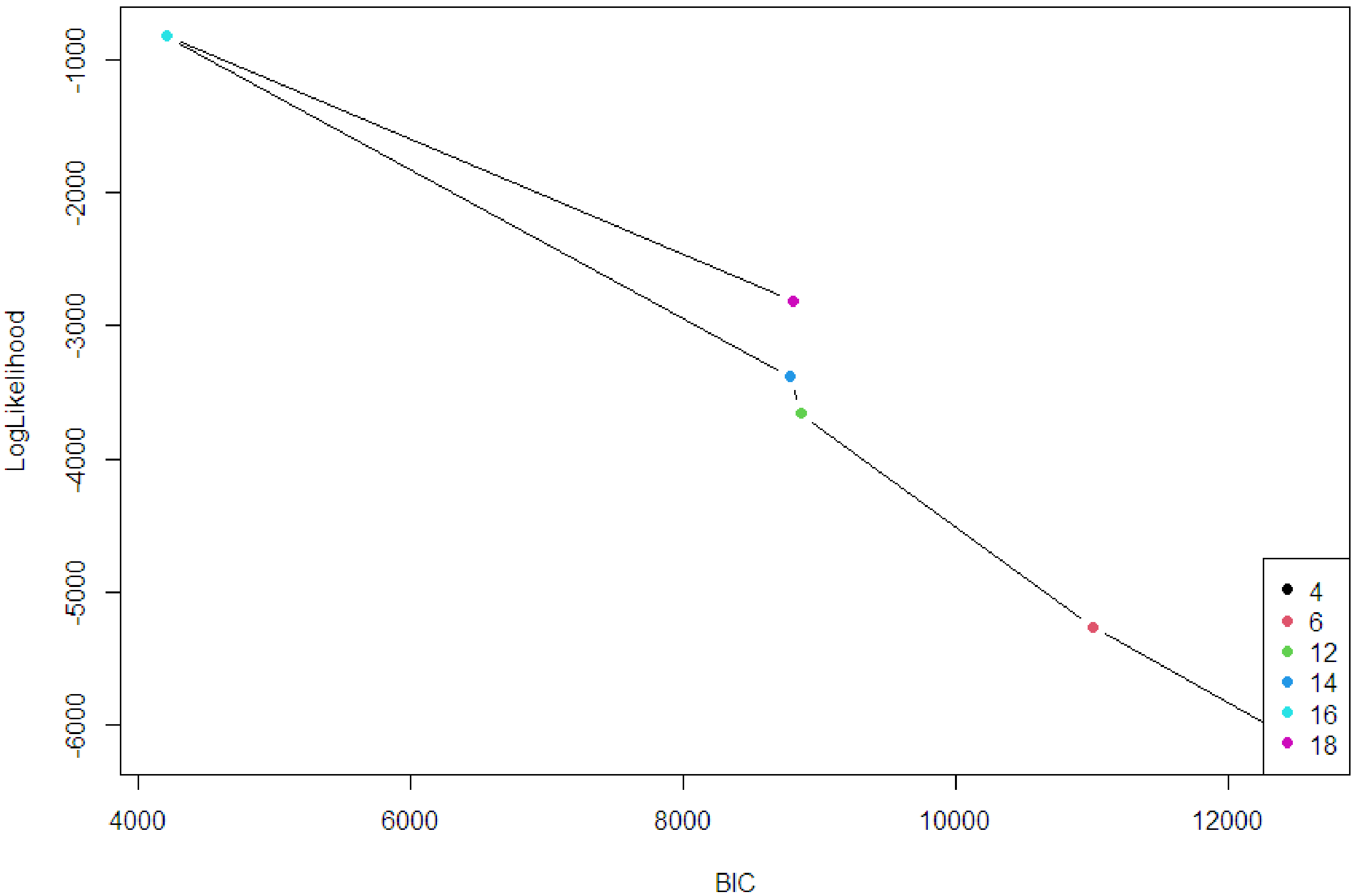


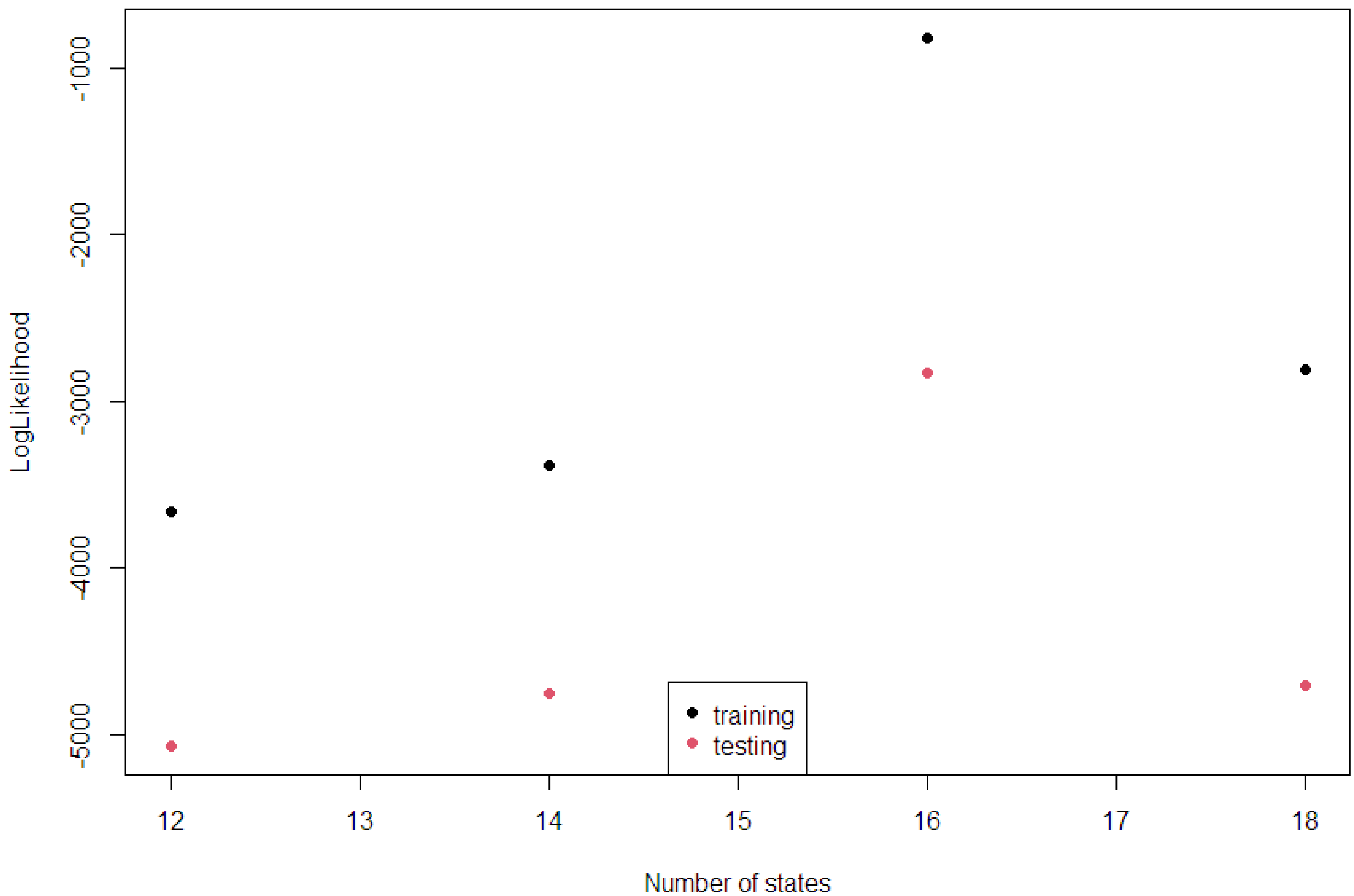
Dataset GAP

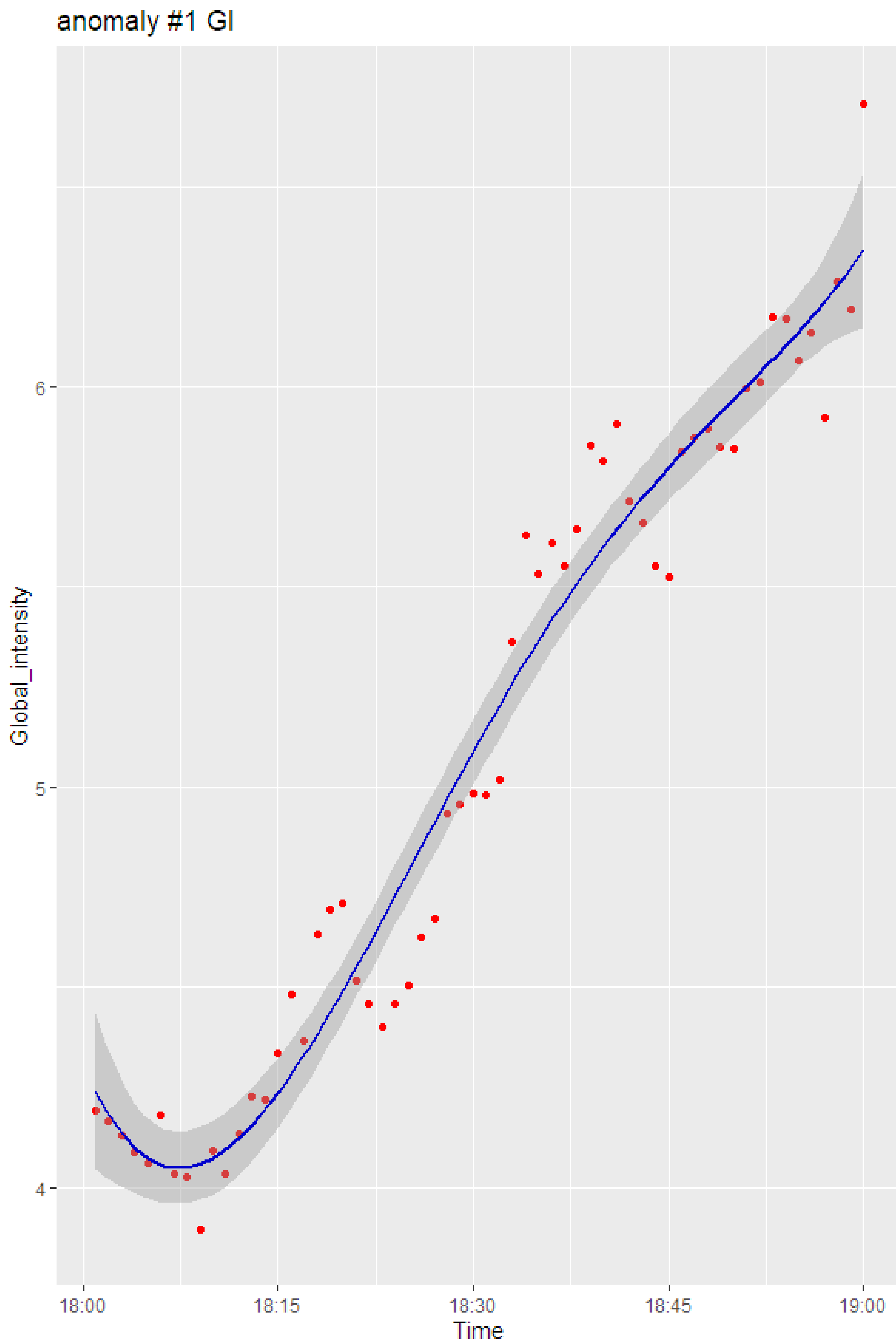



Dataset GAP
Global_active_power
1.6
1.4
1.2
18:00
18:15
18:30
18:45
19:00
Time

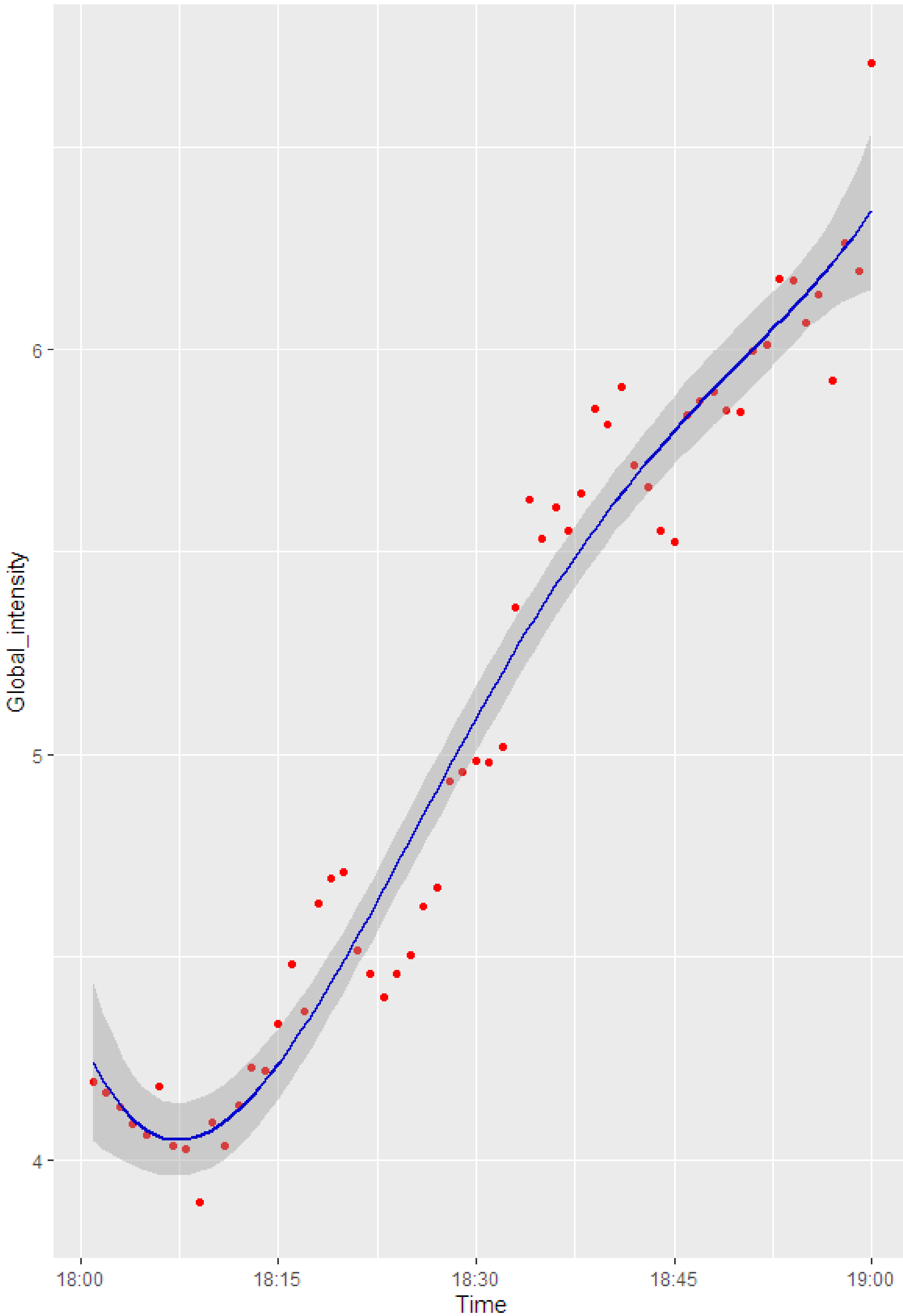

Dataset GI
Global_intensity
6.0
5.5
5.0
4.5
4.0
18:00
18:15
18:30
18:45
19:00
Time

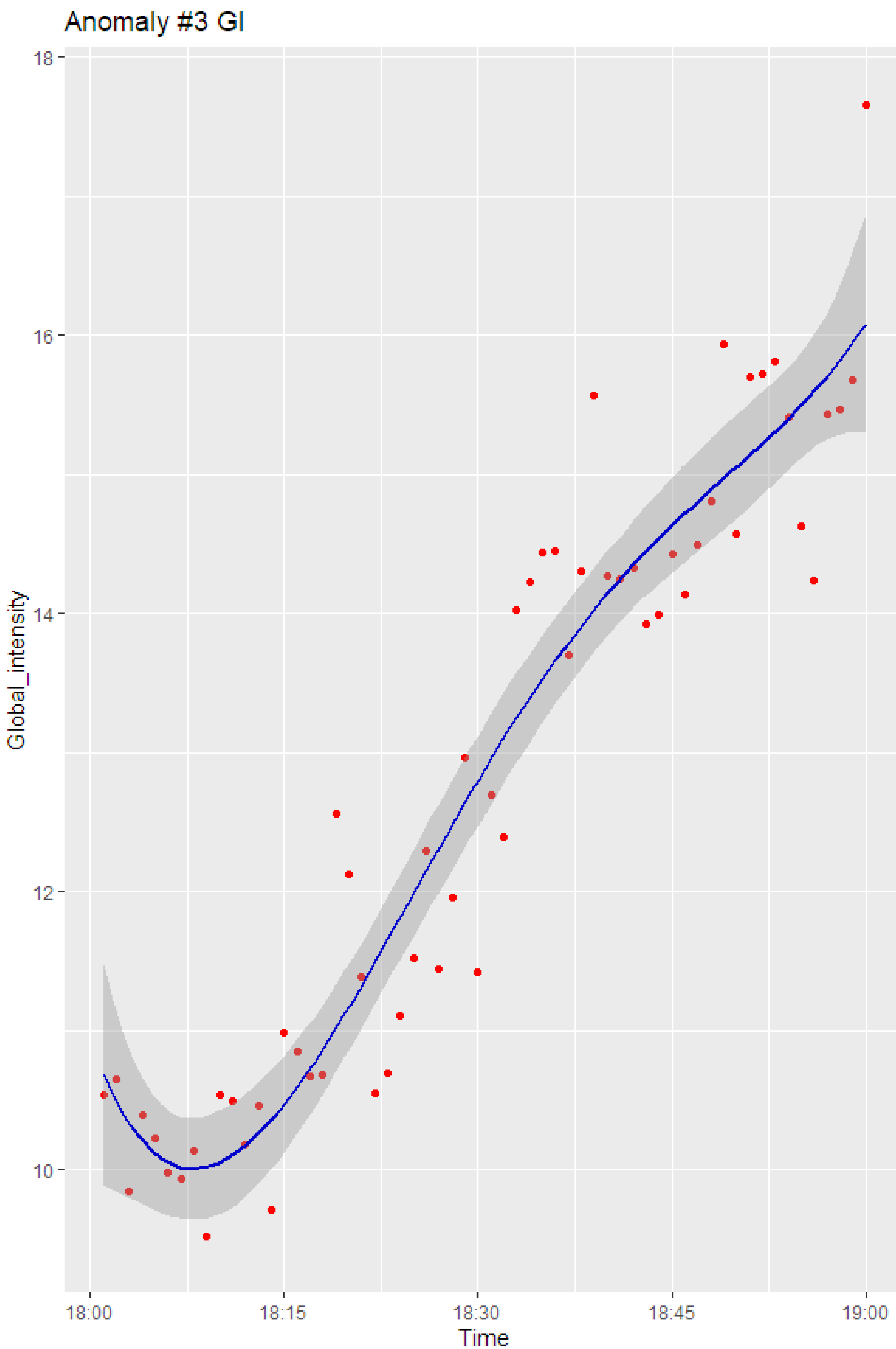

-1000
-2000
-3000
-4000
-5000
-6000
LogLikelihood
4000
6000
8000
10000
12000
BIC
4
6
12
14
16
18

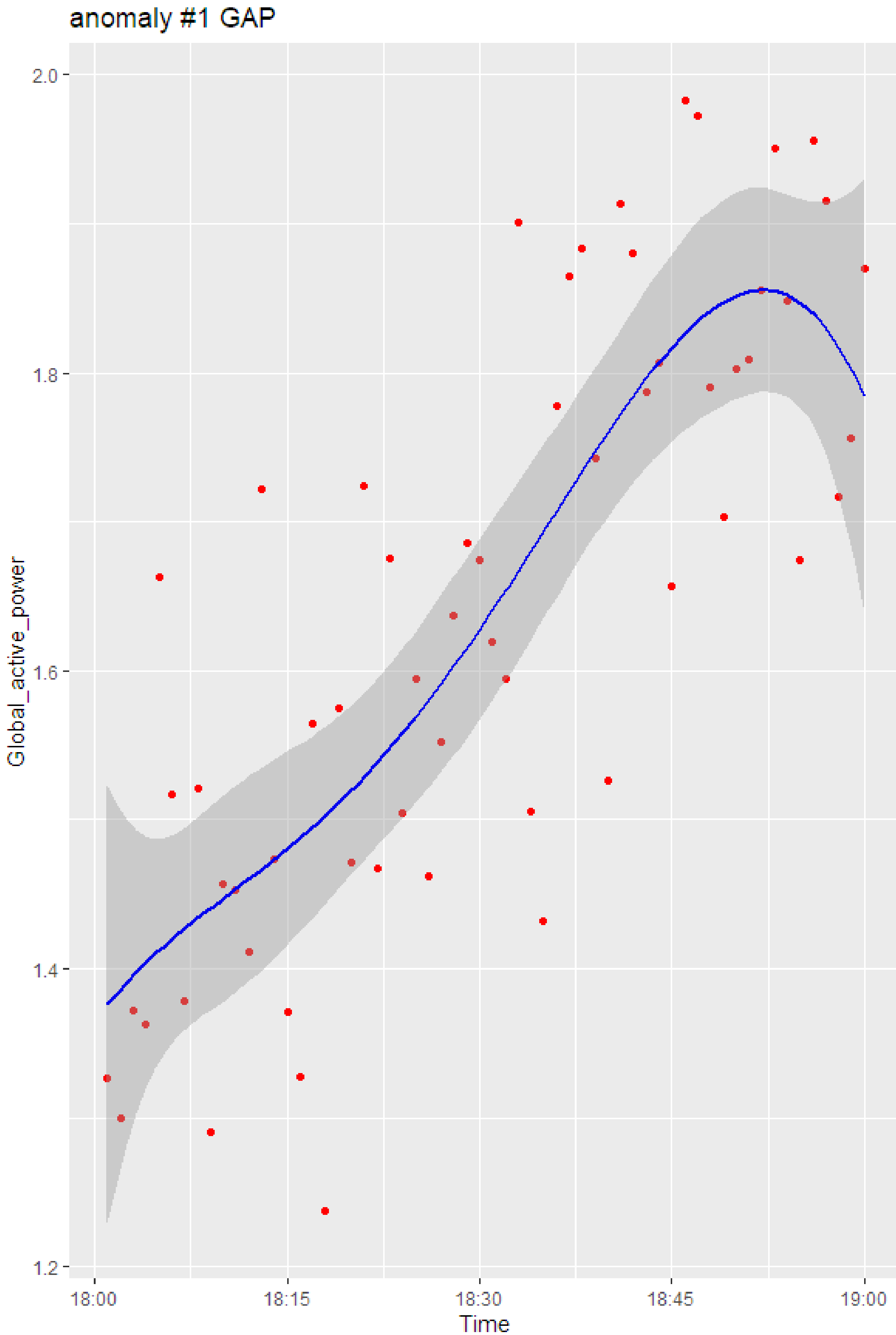

LogLikelihood
-1000
-2000
-3000
-4000
-5000
training
testing
12
13
14
15
16
17
18
Number of states

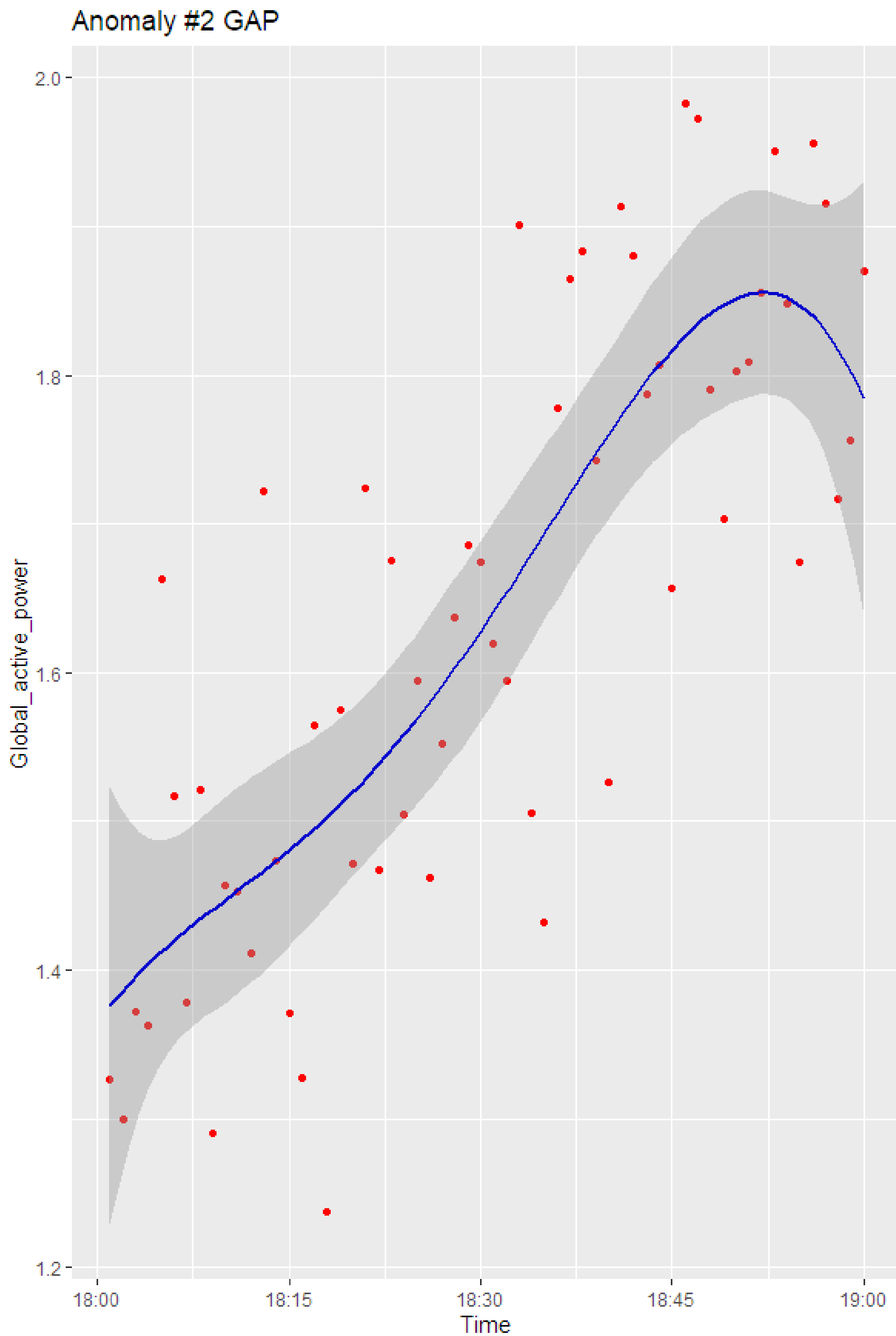

anomaly #1 GI
Global_intensity
6
5
4
18:00
18:15
18:30
18:45
19:00
Time

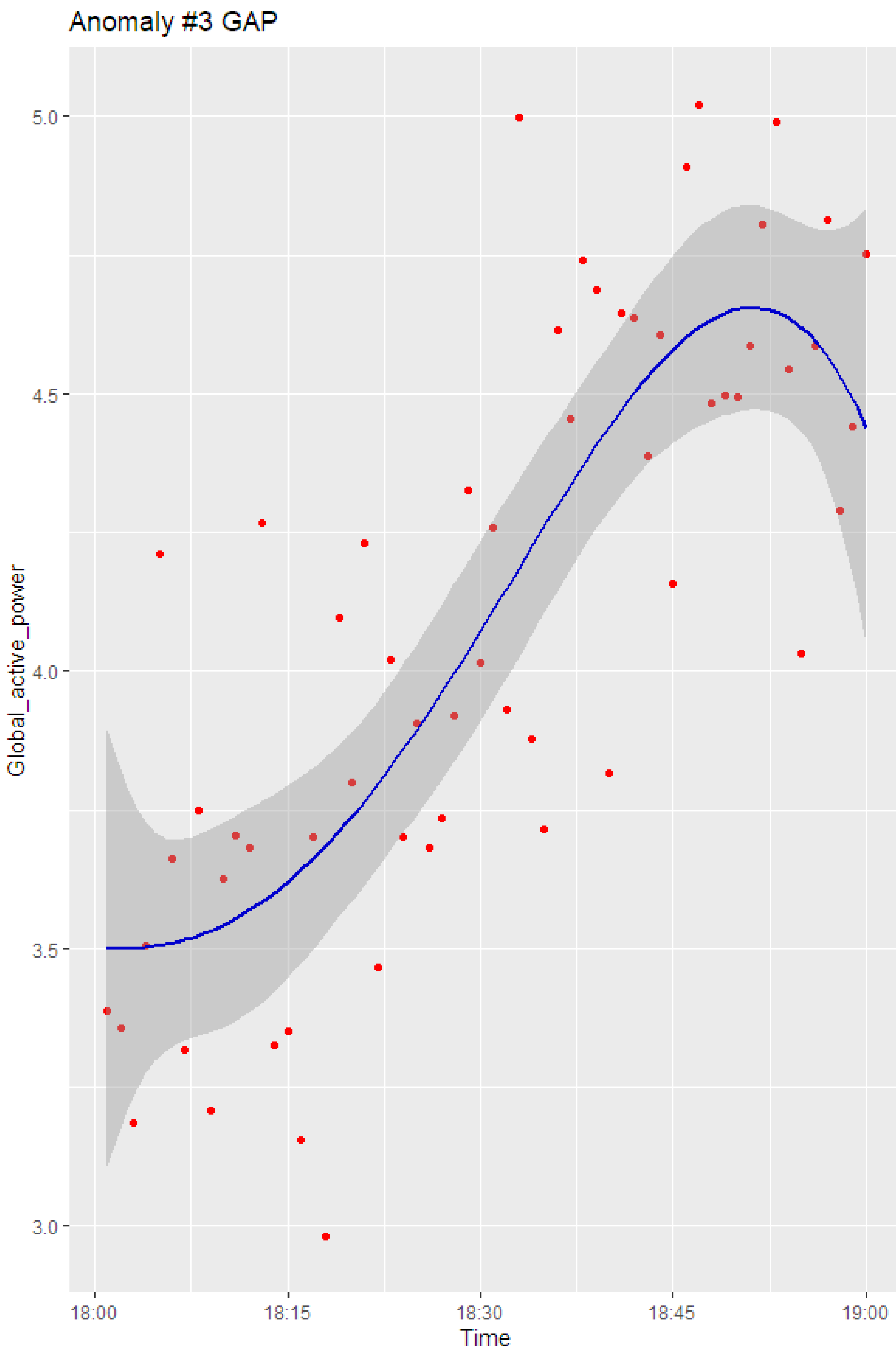
Anomaly #2 GI

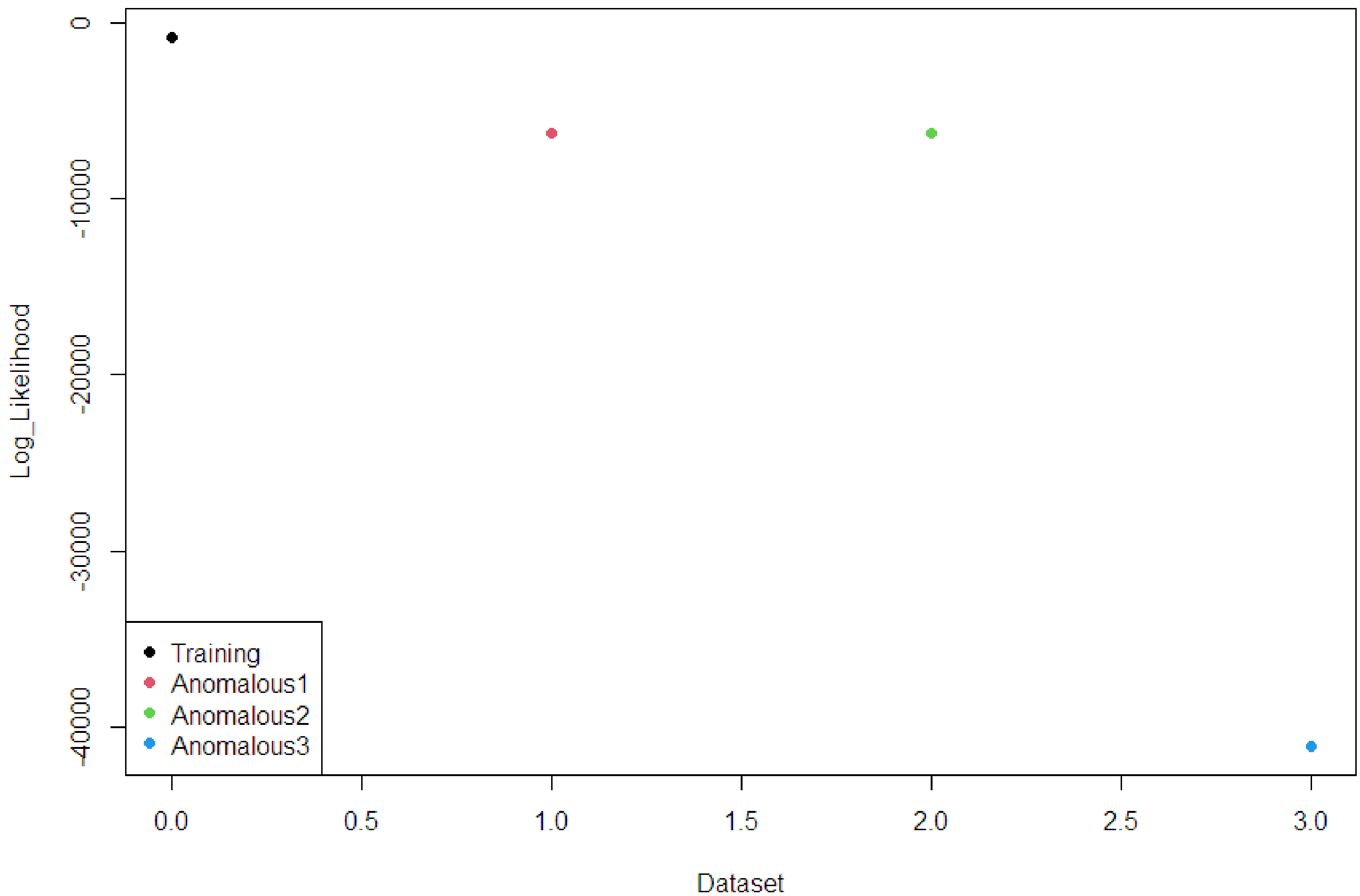



Anomaly #3 GI
18
16
14
12
10
Global_intensity
18:00
18:15
18:30
18:45
19:00
Time


anomaly #1 GAP
Global_active_power
2.0
1.8
1.6
1.4
1.2
18:00
18:15
18:30
18:45
19:00
Time


Anomaly #2 GAP
Global_active_power
2.0
1.8
1.6
1.4
1.2
18:00
18:15
18:30
18:45
19:00
Time


Anomaly #3 GAP
Global_active_power
5.0
4.5
4.0
3.5
3.0
18:00
18:15
18:30
18:45
19:00
Time


Log_Likelihood
0
-10000
-20000
-30000
-40000
Training
Anomalous1
Anomalous2
Anomalous3
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Dataset