4th Asia Pacific Conference of the Prognostics and Health Management,
Tokyo, Japan, September 11 – 14, 2023 DC01-04

Anomaly Detection in Spacecraft Propulsion System using Time Series Classification based on K-NN

Yoshiki Kato[1], Taku Kato[2], and Tsubasa Tanaka[3]

[1,2,3] *Mitsubishi Electric Corporation, Kamakura, Kanagawa, 2478501, Japan*
Kato.Yoshiki@dc.MitsubishiElectric.co.jp
Kato.Taku@aj.MitsubishiElectric.co.jp
Tanaka.Tsubasa@ak.MitsubishiElectric.co.jp

ABSTRACT

In this paper, we propose an anomaly detection method developed by the team called “Team Tsubasa” in the PHMAP2023 Data Challenge. This is an anomaly detection competition for spacecraft propulsion systems (PHM Society, 2023). We joined the Data Challenge with the aim of deepening our knowledge of anomaly detection technology through the competition. In spacecraft propulsion systems, solenoid valve faults and bubble anomalies can occur, and it is considered important to detect them. Also, when other unknown anomalies occur, it is necessary to detect them without confusing them with known anomalies. In this paper, we propose time series classification by k-NN algorithm (Cover & Hart, 1967) as one of the methods to detect these anomalies. In this data challenge, we tried to classify anomalies by k-NN and to identify the location of the anomalies. For those classified as solenoid valve faults, we estimated the opening ratio of the solenoid valve from the similarity of the time series waveforms. As a result, the proposed method achieved a score of 99.05% based on the scoring rules given by the PHMAP 2023 Secretariat and our team won third place.

1. INTRODUCTION

The Data Challenge was held as part of PHMAP 2023 (PHM Society, 2023), an anomaly detection competition for spacecraft propulsion systems, and the authors participated as challengers. According to the view of the PHMAP 2023 Secretariat, PHM technology is expected to improve the anomaly detection technology for spacecraft propulsion systems, and it is said that the Data Challenge is said to have the following background. Telemetry data that can be obtained in orbit is limited due to sensor installation and downlink capacity constraints. A simulator for a simplified propulsion system developed with the cooperation of JAXA can acquire data covering various failure scenarios in real equipment. The data challenge is to detect anomalies from data in various scenarios generated by this simulator. The anomalies to be detected are bubble anomalies and solenoid valve faults, which are typical anomalies of spacecraft propulsion systems. Bubble anomalies are the generation of air bubbles in the spacecraft's pipes, and the existence of the air bubbles changes the speed of sound and causes a slight change in the pressure of the fluid. Also, when solenoid valve faults occur, the opening ratio of the solenoid valve decreases and the amount of fluid passing through the solenoid valve decreases. It is also necessary to detect unknown anomalies without confusion. In addition to detecting these anomalies, the tasks include identifying each abnormal location and predicting the opening ratio of the solenoid valve fault. This is the background to the data challenge. The authors are currently working on developing technology for time series anomaly detection. Our team has proposed various methods (Nakamura, Imamura, Mercer, and Keogh, 2020), (Nakamura, Mercer, Imamura, and Keogh, 2023). We participated in this data challenge with the aim of deepening our knowledge of anomaly detection technology through the competition. This paper proposes algorithms and approaches used in the Data Challenge. The paper is structured as follows. First, we describe the details of the data challenge in Chapter 2. Next, we define the formulas used in the proposed algorithm in Chapter 3, explain the results of the preliminary analysis in Chapter 4, and describe the anomaly detection method and the method for estimating the opening ratio of solenoid valve faults in Chapter 5. In addition, we describe the results in Chapter 6 and discuss them in Chapter 7. Finally, the conclusions are presented in Chapter 8.

2. PROBLEM DESCRIPTION

We classify normal, bubble anomalies, solenoid valve faults and unknown anomalies using pressure sensor data generated by the simulator. And we estimate the opening ratio for solenoid valve faults. It should be noted that the opening of the solenoid valve is a continuous value, with individual differences between spacecraft.

2.1. Experiment Scenarios

A simulator for a simplified propulsion system generates data covering various failure scenarios in real equipment. Figure 1 is a schematic of the experimental propulsion system (Tominaga, Daimon, Toyama, Adachi, Tsutsumi, Omata and Nagata, 2023). There are 4 solenoid valves (SV1 to SV4), and by opening and closing the solenoid valves, the working fluid water pressurized to 2MPa is repeatedly discharged. Pressure sensors (P1 to P7) are installed at 7 locations in the propulsion system. Pressure data is acquired at a sampling rate of 1 kHz from 0 to 1200 ms. The solenoid valves repeat the operation of opening for 100 ms and then closing for 300ms. Opening and closing are performed 3 times and measured for a total of 1200 ms (Figure 2, Tominaga at al, 2023). There are individual differences in the solenoid valves. 8 locations of BV1 and BP1 to BP7 can be considered as bubble generation locations. For simplicity, the amount of bubbles is constant in all cases. The opening of the solenoid valve is normally 0% or 100%, but when a fault occurs, the opening is between 0% and 100%. The test data also contains unknown anomalies that need to be detected. In this competition, there are spacecrafts from No.1 to No.4. The training data contains data from spacecraft No.1 to No.3, and the test data contains data from spacecraft No.1 and No.4.

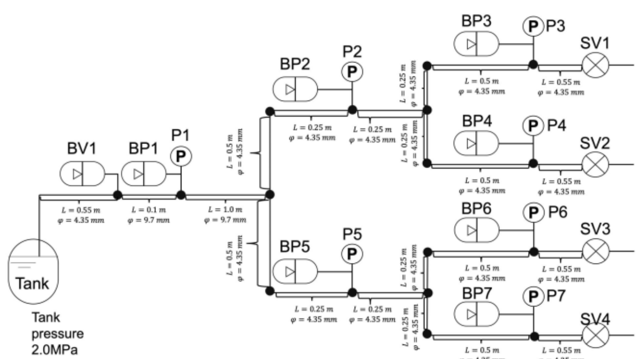


Figure 1. Schematic of experimental propulsion system. (Tominaga et al., 2023)

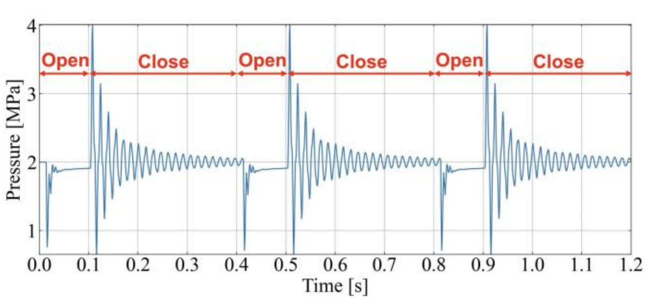


Figure 2. Typical pressure profile. (Tominaga et al., 2023)

2.2. Prediction Goals

There are five questions in this Data Challenge. Points are awarded for each correct answer. In addition, the spacecraft4 gets twice as many points for each correct answer.

1. Determine whether all test data are normal or abnormal. Each is worth 10 points.
2. For the data detected as abnormal, determine whether it is a bubble anomaly, a solenoid valve fault, or an unknown anomaly. Each is worth 10 points.
3. For the data identified as bubble anomaly, determine the location of the bubble from eight locations, BV1, and BP1 to BP7. Each is worth 10 points.
4. For the data identified as solenoid valve fault, determine which of the four solenoid valves (SV1 to SV4) failed. Each is worth 10 points.
5. For the solenoid valve identified as a fault, predict the opening ratio (0% <= opening ratio < 100%). A score is given by subtracting the difference between the correct answer and the estimated value from 20 points. The score can never be less than 0.

2.3. Dataset

This section describes the datasets given in this competition. As shown in Table 1, there are 35 normal training data, 16 solenoid valve abnormal data, and 8 bubble abnormal data for spacecraft 1, 2, and 3, respectively. For solenoid valve faults, there are 4 patterns of opening (0%, 25%, 50%, 75%) for each solenoid valve (SV1, SV2, SV3, SV4) and each has 1 data. Therefore, there are 16 data with a combination of 4 solenoid valves and 4 opening patterns. For bubble anomalies, there are 8 patterns, one for each location (BP1, BP2, BP3, BP4, BP5, BP6, BP7, BV1). Also, the anomalies do not occur at the same time. On the other hand, there are 23 test data each for spacecraft 1 and 4.

Table 1. Training Data (spacecraft:1,2,3)

Condition	Number	Label
Normal	35	-
Solenoid valve faults	16	Solenoid valve: SV1, SV2, SV3, SV4 Opening ratio: 0%, 25%, 50%, 75%
Bubble anomaly	8	location: BP1, BP2, BP3, BP4, BP5, BP6, BP7, BV1

3. DEFINITION

We define the similarity between data used in the proposed algorithm. First, the similarity D(i) between the i-th pressure sensors (P1 to P7) of the two data is defined as the following formula (1). where p(i,j) and q(i,j) are j[ms] values of the i-th pressure sensor in each data, and Start and *End* are the intervals to be analyzed.

$$\sum_{j=Start}^{End} |p(i,j) - q(i,j)| \tag{1}$$

Also, the similarity D using all seven pressure sensors is defined as the following formula (2).

$$\sum_{i=1}^{7} D(i) \tag{2}$$

In addition, we define the maximum standard deviation of each sensor value at a certain time. When the number of data is n, the maximum standard deviation SD is defined as the following formula (3) below. $P_{k,i}$ $(k = 1, \ldots, n, i = 1, \ldots, 7)$ is the sensor value at a certain time, and E_i is the expected value of $P_{1,i}, \ldots P_{n,i}$.

$$\max\{\left(\frac{1}{n-1}\sum_{k=1}^{n}(P_{k,i} - E_i)^2\right)^{\frac{1}{2}};\ i = 1, \ldots 7\} \tag{3}$$

4. PRELIMINARY ANALYSIS

This chapter presents the results of the preliminary analysis. As a result of comparison under the same conditions, it was found that the waveforms match from 0 to 103 ms. We also calculated the similarity D (formula (2)) between the normal and each anomaly to compare different conditions. Details are given in each section.

4.1. Comparison under same conditions

First, we searched for conditions that would reduce the variability between spacecraft in the same type of anomaly. We calculated the standard deviation SD (formula (3)) of each mode from 0 to 1200ms and confirmed the variation of each mode for each spacecraft.

- For the normal data, we calculated the standard deviation SD of 105 data including spacecrafts 1 to 3.
- For solenoid valve faults, we obtained the standard deviation SD of spacecraft 1 to 3 for each of 16 conditions based on abnormal location (SV1, SV2, SV3, SV4) and opening (0%, 25%, 50%, 75%).
- For the bubble anomaly, we obtained the standard deviation SD of spacecraft 1 to 3 for each of 8 conditions based on bubble location (BV1, and BP1 to BP7).

As a result, maximum standard deviation SD is 0 under the same conditions from 0 to 103ms. Figure 3 plots the maximum value of the maximum standard deviation SD under the same conditions (abnormal location and opening). Based on this fact, the waveforms match from 0 to 103 ms under the same conditions, and it is considered that individual differences between spacecraft can be ignored.

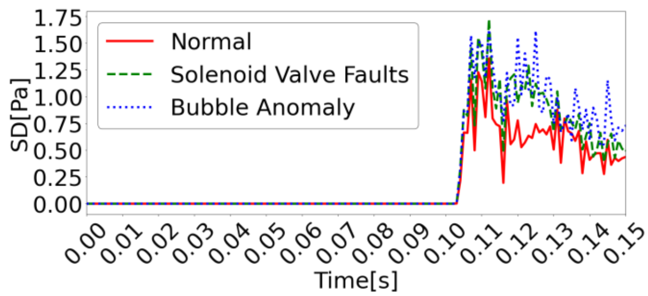


Figure 3. maximum value of the maximum standard deviation SD

4.2. Comparison of normal and each anomaly

In this section we compare different conditions. We calculated the similarity D (formula (2)) between normal and each abnormal mode in the interval from 0 to 99 ms (Figure 4). As a result, the similarity D was 8 or more and less than 73. Also, when the similarity D exceeds 60, it is the case where the opening of the solenoid valve is 0, and in all other cases the similarity D is 40 or less. It is considered that there is a sudden change near the opening ratio of 0. We also predict an upper limit on the possible values of the similarity D. We define $D_{SV1,0}$, $D_{SV2,0}$, $D_{SV3,0}$, and $D_{SV4,0}$ as the values of similarity D when the opening ratio of SV1, SV2, SV3, SV4 is 0, respectively. We set μ+3σ as the upper limit of the similarity D. μ is the average of $D_{SV1,0}$, $D_{SV2,0}$, $D_{SV3,0}$, $D_{SV4,0}$, and σ is the unbiased standard deviation of $D_{SV1,0}$, $D_{SV2,0}$, $D_{SV3,0}$, $D_{SV4,0}$. If the sample exceeds the upper limit, we treat it as an unknown anomaly.

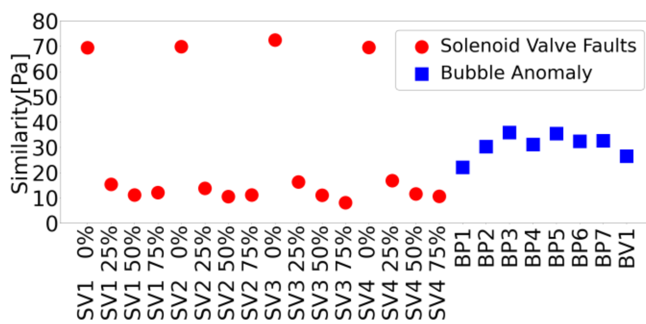
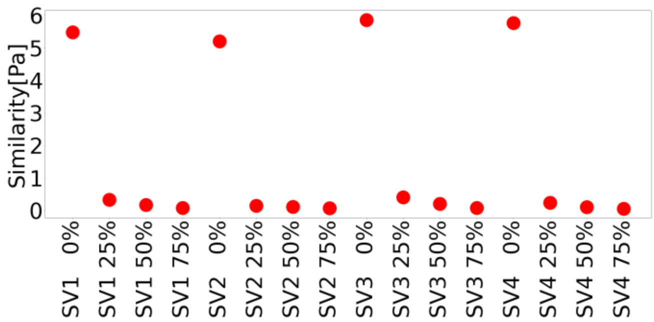
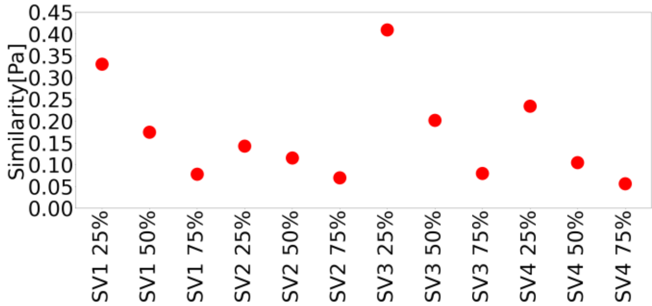


Figure 4. Similarity D between normal and each abnormal mode

We also calculated the similarity D(1) between each solenoid valve fault and normal from 0 to 99 ms for P1 (Figure 5). As a result, it was found that the similarity D(1) for any solenoid valve monotonically decreases as the opening ratio increases. We decided to use the similarity D(1) for opening estimation.



(a) Similarity D(1)



(b) Similarity D(1) (y-axis range: 0 to 0.45)

Figure 5. Similarity D(1) between each solenoid valve fault and normal

5. Anomaly Classification

In this chapter, we propose a method of classifying anomalies and a method of estimating the opening of the solenoid valve. As a premise, according to the results of the preliminary analysis in Chapter 4, it is thought that variations due to individual differences in each spacecraft can be ignored in the interval from 0 to 103ms, so we will focus on this interval (0 to 99ms). The procedure for classification and prediction is as shown in the flow chart in Figure 6. Details are explained in each section.

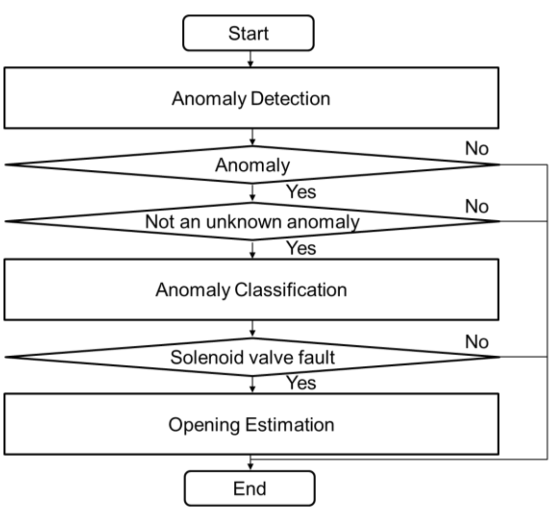


Figure 6. Overall flowchart

5.1. Anomaly Detection

First, we classify normal and abnormal. It is thought that the waveforms of normal data from 0 to 99 ms match each other. For each test data sample, the similarity D (formula (2)) to normal data is calculated in the interval from 0 to 99 ms, and those with D = 0 are classified as normal, and others as abnormal. Also, if the similarity D to normal exceeds the upper limit ($\mu+3\sigma$) obtained in Chapter 4, we classify the sample as an unknown anomaly. Figure 7 is this flow chart.

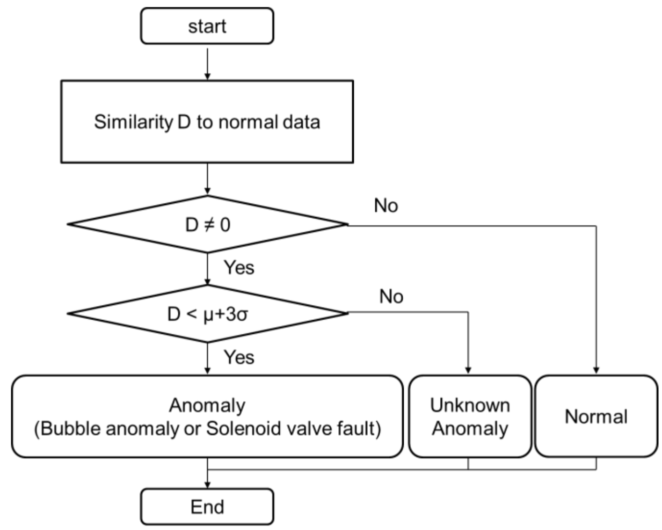


Figure 7. Flowchart (Anomaly Detection)

5.2. Anomaly Classification

We classify and locate anomalies using an interval from 0 to 99 ms. Figure 8 is this flow chart. We apply the k-NN algorithm (Cover & Hart, 1967) using the similarity D defined in Chapter 3, and each test data sample is classified into the mode with the lowest similarity D among the modes in the training data. The training data mode used for k-NN are 25 modes of normal, solenoid valve faults and bubble anomalies described below. We treat normal data as 100% open mode.

- Normal (100% open mode)
- 4 solenoid valves (SV1, SV2, SV3, SV4) and 4 patterns of opening (0%, 25%, 50%, 75%)
- Locations of bubble (BV1, and BP1 to BP7)

We interpret the samples classified as normal by the k-NN results as those with solenoid valve openings greater than 75% and less than 100%, and we name this a minor SV fault. For those classified as minor SV fault, we use the following procedure to locate the abnormal solenoid valve.

1. For each sensor (P3, P4, P6, P7) adjacent to the solenoid valve (SV1 SV2, SV3, SV4), we calculate the similarity D(i) ($i = 3{,}4{,}6{,}7$) defined in Chapter 3.
2. We determine the solenoid valve adjacent to the sensor with the maximum value of similarity D(i) among P3, P4, P6, and P7 as the abnormal location.

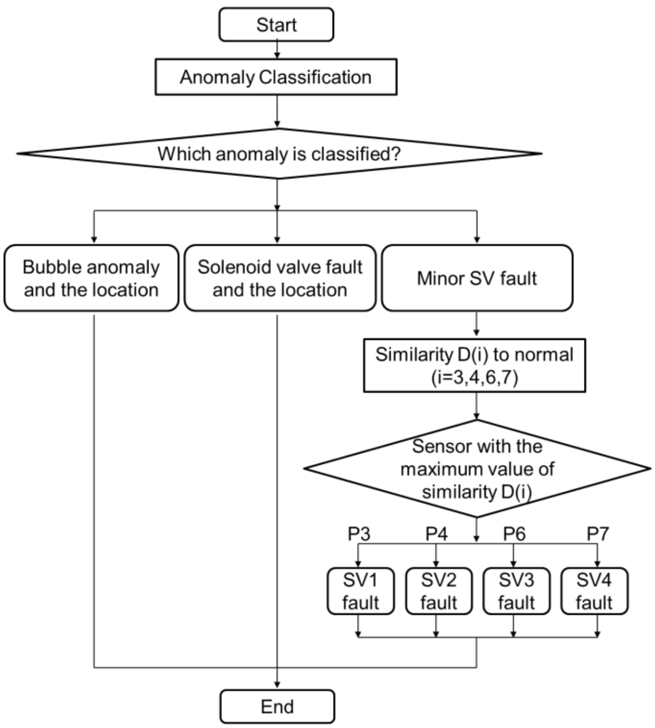


Figure 8. Flowchart (Anomaly Classification)

5.3. Opening Estimation

We explain the prediction method of the opening ratio for the solenoid valve diagnosed as abnormal. Figure 9 is this flow chart. The opening ratio (0, 25, 50, 75) is set for each abnormal solenoid valve in the training data. Also, since normal data can be interpreted as data with an opening ratio of 100%, we use these to estimate the opening ratio. Using the sensor P1 and the sensors (P3, P4, P6, P7) adjacent to the abnormal solenoid valve, we estimate the opening as follows.

1. We calculate the similarity D(1) for the sensor P1. From the analysis results in Chapter4, the similarity D(1) to the normal decreases monotonically with increasing opening. Using the value of D(1), we determine the range (O_1, O_2) in which the opening of the test data sample is included. (O_1, O_2) is (0,25), (25,50), (50,75), or (75,100).
2. We calculate the similarity D(i) ($i = 3{,}4{,}6, or\ 7$) between the test data sample and the training data sample with opening O_1 and O_2 as S_1 and S_2, respectively. The sensor used for calculation is the sensor adjacent to the abnormal solenoid valve.
3. We calculate the value that internally divides the openings O_1 and O_2 into $S_1 : S_2$. The opening is estimated by the following formula (4).

$$\frac{O_1 S_2 + O_2 S_1}{S_1 + S_2} \tag{4}$$

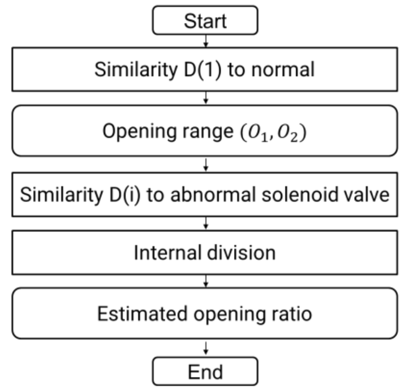


Figure 9. Flowchart (Opening Estimation)

6. CLASSIFICATION AND ESTIMATION RESULTS

First, we describe the classification results. As a result of implementing the algorithm proposed in Chapter 5 on the test data, 20 data out of 46 data were classified as normal, 10 data as bubble anomaly, 10 data as solenoid valve fault, and 6 data as unknown anomaly. Next, we describe the opening estimation results. We adopted the estimation results of the algorithm in Chapter 5 for 9 out of 10 data classified as solenoid valve faults. For the rest of the data, we regarded it as exceptional data from comparison of sensor waveforms adjacent to the solenoid valve. This data, classified as an SV1 anomaly, seems similar in waveform shape to the 25% open waveform for the adjacent P3 waveform (Figure 10, Figure 11). However, when comparing the waveforms of P4, P6, and P7, it seems that the change from normal data is larger than the case of 25% opening, so we expect this data to be around 0% opening, which is lower than 25%.

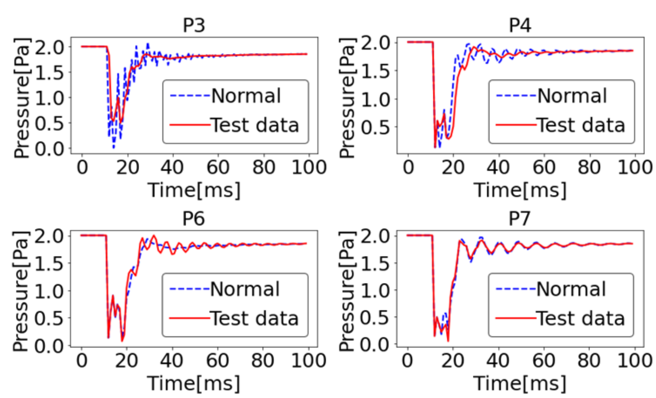


Figure 10. Sensor waveforms adjacent to the solenoid valve: exceptional test data

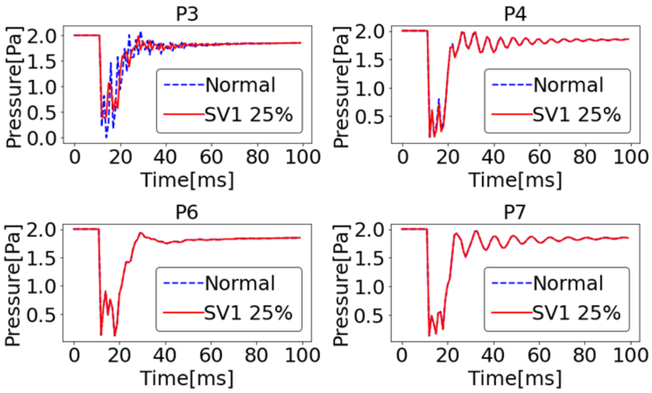


Figure 11. Sensor waveforms adjacent to the solenoid valve: SV1 fault data (opening ratio of 25%)

In addition, we adjusted the opening ratio by ±2% based on daily updated leaderboard scores. As a result, the proposed method achieved a score of 99.05% based on the scoring rules given by the PHMAP 2023 secretariat and our team won third place.

7. Discussion

We discuss the obtained score and prediction accuracy. We predicted the score obtained from the calculation method described in Chapter 2. If we correctly classify normal and abnormal and localize all anomalies, we expected to get a score of 82.14%. Also, if each solenoid valve opening differs from the correct value by 2%, we expected to obtain a score of 98.21%. And if each solenoid valve opening differs from the correct value by 1%, we expected to obtain a score of 99.11%. The final score was 99.05%, so it seems that we predicted each opening with a difference of about 1%. Next, we also consider unknown anomalies. When we visualized the waveforms of those classified as unknown anomalies, we found that all waveforms started at a pressure of 3 MPa (Figure 12). Since the propulsion system was pressurized at 2 MPa in the experimental scenario, we consider this fact to be some anomaly related to the pressure setting.

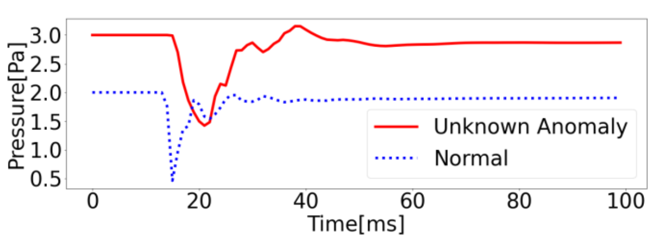


Figure 12. The waveform of sensor data (P1) classified as unknown anomaly

8. Conclusion

In this paper, we proposed an anomaly detection method for spacecraft propulsion systems. We tried to classify anomalies by k-NN and to identify the location of the anomalies. Also, for those classified as solenoid valve faults, we estimated the opening ratio of the solenoid valve from the similarity of the time series waveforms. In PHMAP2023 Data challenge, the proposed method achieved a score of 99.05%.

Acknowledgement

We would like to thank the organizers of the data challenge.

References

Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. IEEE Transactions on Information Theory, vol.13, pp. 21-27.

Nakamura, T., Imamura, M., Mercer, R., & Keogh, E. (2020). Merlin: Parameter-free discovery of arbitrary length anomalies in massive time series archives. 2020 IEEE international conference on data mining (ICDM). IEEE

Nakamura, T., Mercer, R., Imamura, M., & Keogh, E. (2023). MERLIN++: parameter-free discovery of time series anomalies. Data Mining and Knowledge Discovery, vol. 37, pp. 670-709.

PHM Society. (2023). PHM Conference Data Challenge.

Retrieved 18 July 2023 from https://data.phmsociety.org/phmap-2023-data-challenge

Tominaga, K., Daimon, Y., Toyama, M., Adachi, K., Tsutsumi, S., Omata, N., & Nagata, T. (2023). Dataset generation based on 1D-CAE modeling for fault diagnostics in a spacecraft propulsion system. Asia Pacific Conference of the Prognostics and Health Management Society 2023 (PHMAP 2023).

Yoshiki Kato received his M.S. degree from Chiba University in 2019. Since 2022, he has been with Mitsubishi Electric Corp., Japan. His research interests include data analysis and machine learning.

Taku Kato received his M.E. degree from The University of Tokyo in 2023. Since 2023, he has been with Mitsubishi Electric Corp., Japan. His research interests include data analysis and machine learning.

Tsubasa Tanaka received his B.E. degree from Hokkaido University in 2017. Since 2019, he has been with Mitsubishi Electric Corp., Japan. His research interests include data analysis and machine learning.