Estimating the health of turbine engine based on the relationship between torque margin and density altitude

Kosei Ozeki[1], Takahiko Masuzaki[2], Takeru Shiraga[3], Koji Wakimoto[4] and Nakamura Takaaki[5]

[1,2,3,4,5] *Mitsubishi Electric Corporation, Kamakura, Kanagawa, 2478501, Japan*

ozeki.kosei@aj.mitsubishielectric.co.jp
masuzaki.takahiko@dc.mitsubishielectric.co.jp
shiraga.takeru@ea.mitsubishielectric.co.jp
wakimoto.koji@df.mitsubishielectric.co.jp
nakamura.takaaki@dy.mitsubishielectric.co.jp

Abstract

We present an anomaly detection method developed for the PHM North America 2024 Conference Data Challenge. This competition is aimed at estimating the health of helicopter turbine engines (PHM Society, 2024). The task includes the estimation of the torque margin (regression) and the health state (binary classification) of turbine engines. We developed an estimation model using a hybrid algorithm that combines data-based machine learning and domain knowledge-based processing. Our method achieved scores over 0.99 for both the testing and validation datasets. based on the calculation rules provided by PHM Society. These results were ranked first among all the participating teams.

1. Introduction

The Data Challenge was held as part of PHM 2024 (PHM Society, 2024). This competition is an anomaly detection for the health of helicopter turbine engines, and the authors participated as challengers. Turbine engines deteriorate over time and require regular maintenance (BHT-407-FM-3, 2018). However, engine data is rarely collected automatically, making time-consuming trend analysis difficult. Therefore, early detection of deterioration may not be possible, and the timing of maintenance may be missed (Bechhoefer, and Kessler, 2022). To address this issue, machine learning-based methods such as SVM (Cao, Xu, Huang, and Yang, 2022), (Chakraborty, Sarkar, Ray, and Phoha, 2010) and Deep Learning (Huber, Palmé, and Chao, 2023), (Luo, and Zhong, 2017), as well as mathematical approaches (Bechhoefer, and Hajimohammadali, 2023), (Zhou, Zhou, Li, and Ca, 2023), (Tolani, Yasar, Chin, and Ray, 2005), are being tried to monitor the performance and health of the engine. We have been working on the development of anomaly detection technology and have proposed various methods. (Nakamura, Imamura, Mercer, and Keogh, 2020), (Nakamura, Mercer, Imamura, and Keogh, 2023). At PHMAP 2023 (PHM Society, 2023), The Data Challenge was held to detect anomalies in spacecraft propulsion systems. Our team proposed a time series classification method using the k-NN algorithm and achieved a score of 99.05%. This score was ranked third among all participating teams. This method primarily relies on machine learning algorithms, and we consider that using domain knowledge is effective for further improving accuracy (Kato, Kato, and Tanaka, 2023). In this Data Challenge, we proposed an anomaly detection approach that actively leverages domain knowledge.

2. Problem Description

In the Data Challenge, the task is to evaluate the health of helicopter turbine engines. This chapter presents the tasks and datasets of the Data Challenge.

2.1. Task

This task involves two problems: torque margin regression and the health state classification. The torque margin is an indicator of engine health. The health state is classified as healthy or faulty. Both regression and classification are evaluated separately, with the final score being the average of these two. Participants are required to estimate the engine health using the testing and validation datasets and submit their model's results. They are also required to provide a confidence level (class_conf) for their classification results. A severe penalty is imposed for high-confidence false negatives (instances where the engine is predicted to be healthy but is in fact faulty) because overlooking an engine failure could lead to expensive repair costs.

2.2. Dataset

The dataset for this competition consists of seven engines (assets) of the same model number. Each engine is instrumented to capture the outside air temperature (oat), mean gas temperature (mgt), pressure altitude (pa), indicated airspeed (ias), net power (np), and compressor speed (ng). For these operational conditions, there is a design (target) torque (trq_target). The real output torque (trq_measured) is also measured. The torque margin (trq_margin) calculated from equation (1) is assessed by comparing the output torque to the design (target) torque. The health state (faulty) is labeled into two values (0 = healthy, 1 = faulty).

$$trq\_margin = \frac{trq\_measured - trq\_target}{trq\_target} \times 100 \qquad (1)$$

Three types of datasets are provided: training, testing, and validation. The training dataset is comprised of data from four out of seven engines. The remaining three engines are used for the testing and validation datasets, with the torque margin and the health state hidden. All datasets are shuffled, and individual and temporal information is hidden.

3. PRELIMINARY ANALYSIS

This chapter presents the results of our preliminary analysis. We surveyed the meanings of the variables and domain knowledge from related materials. Subsequently, we also analyzed the dataset using this domain knowledge.

3.1. Density altitude (da)

We estimated the density altitude (da), which is not included in the disclosed data, using explanatory variables. Density altitude is defined as the pressure altitude corrected for ambient temperature. A higher temperature increases the density altitude, thereby affecting aircraft performance. The density altitude is approximated using the following equation, which incorporates the pressure altitude and outside temperature from the disclosed data.

$$da = 1.2376 * pa + 118.8 * oat - 1782 \qquad (2)$$

3.2. Relationship between variables

According to related materials (Airplane Academy, 2024), the explanatory variables for the disclosed data have a relationship shown in Figure 1. Pilots operate helicopters by adjusting the engine output according to external conditions, and by managing the output torque and propeller rotation speed. The target torque is determined by physical conditions such as outside air temperature, pressure altitude, airspeed, and altitude. Additionally, we formulated the following hypotheses about the relationships between these variables, based on domain knowledge.

Hypothesis 1. The outside air temperature (oat) and pressure altitude (pa) change gradually over time. In other words, they do not change suddenly or abruptly.

Hypothesis 2. Faulty engines tend to have a small torque margin. On the other hand, healthy engines also have a reduced torque margin under high density altitude conditions.

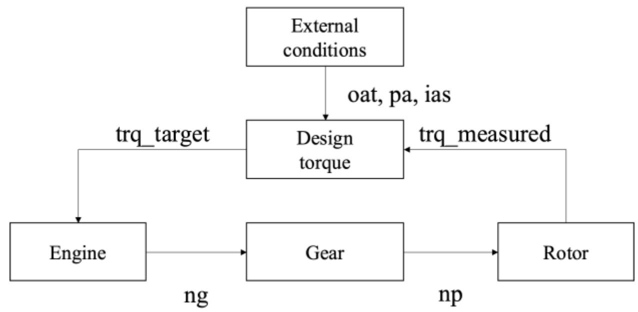

Figure 1. The relationship among explanatory variables

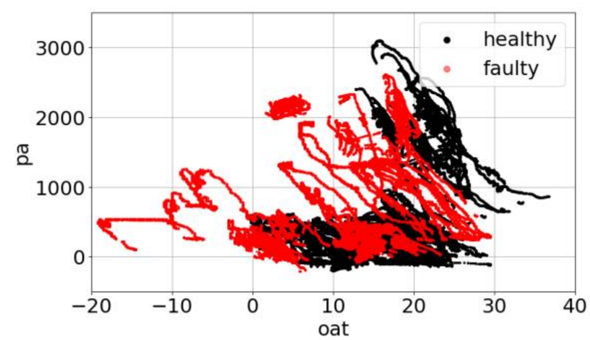

Figure 2. The Training Dataset (oat – pa)

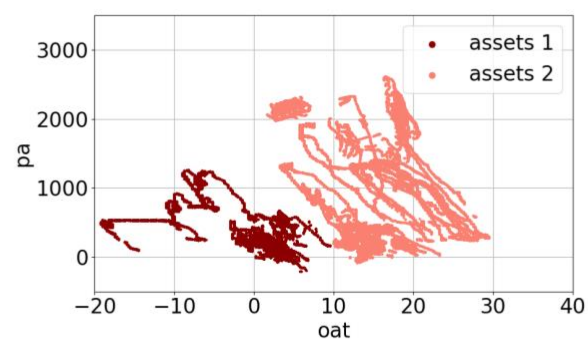

Figure 3. The sequence of faulty assets

3.3. Analysis of the dataset

Figure 2 shows the relationship between outside air temperature (oat) and pressure altitude (pa) in the training dataset. From Figure 2, the faulty data points continue linearly. Based on "Hypothesis 1", we inferred that these data were measured as the same individual. Also, as shown in Figure 3, we estimated that this sequence of fault assets is classified into two assets. Figure 4 shows the relationship between density altitude (da) and torque margin (trq_margin) in the training dataset. From Figure 4, we found two assets and estimated the following relationships.

- "Hypothesis 2" is valid for "Asset 1".
- "Hypothesis 2" is not valid for "Asset 2".

Similarly, Figure 5 shows the relationship between the density altitude (da) and the torque margin (trq_margin) in the testing and validation datasets. Based on Figure 5, we estimated that both the testing and validation datasets support "Hypothesis 2". Due to the differing characteristics of the assets in the training dataset and the testing/validation datasets, we anticipated a decline in the performance of data-based machine learning. As a countermeasure, we hypothesized that incorporating domain knowledge-based processing into machine learning would be effective.

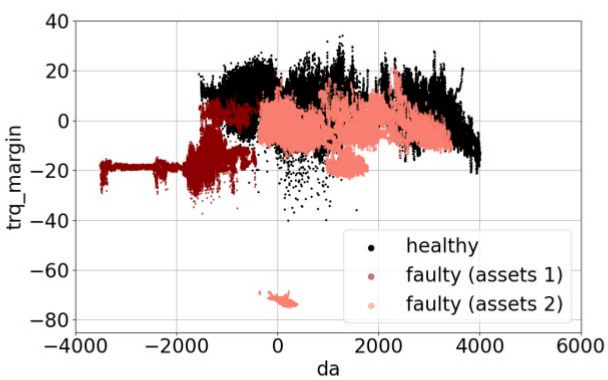

Figure 4. The Training Dataset
(da - trq_margin)

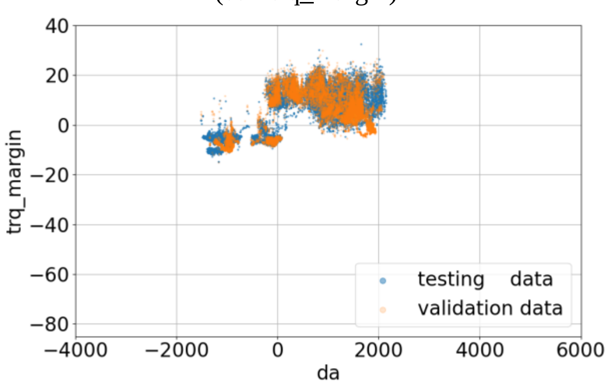

Figure 5. The Testing and Validation Datasets
(da - trq_margin)

4. PROPOSED ALGORITHM

This chapter presents the hybrid algorithm combining data-based machine learning and domain knowledge-based processing. The steps in the regression and classification process are shown in the flowchart in Figure 6. Details of each step are explained in the corresponding sections.

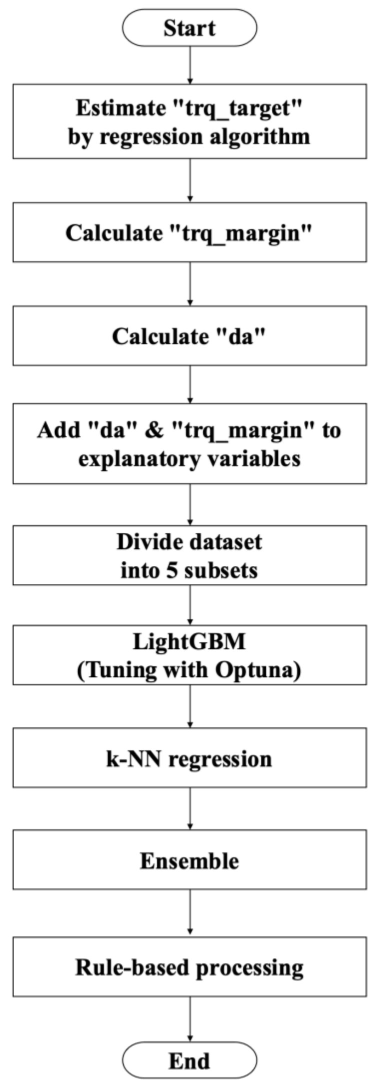

Figure 6. flowchart of Anomaly Detection

4.1. Regression algorithm

We hypothesized that estimating target torques determined by physical conditions would provide greater accuracy than estimating torque margins directly from domain knowledge. Therefore, the regression algorithm is structured to estimate the target torque (trq_target_pred) using a second-order polynomial regression based on the data. The torque margin, calculated from equation (3), is assessed by comparing the output torque (trq_measured) to the estimated target torque (trq_target_pred). In our regression algorithm, density altitude is not utilized as an explanatory variable. We hypothesized that the linear regression model could compensate for the density altitude linearly using other explanatory variables.

$$trq\_margin = \frac{trq\_measured - trq\_target\_pred}{trq\_target\_pred} \times 100 \qquad (3)$$

4.2. Classification algorithm

The classification algorithm is structured to estimate the health state (faulty or normal), using a combination of data-based algorithm and domain knowledge-based processing. Domain knowledge-based processing involves the k-NN algorithm and the rule-based processing. In the training of the classification algorithm, the explanatory variables include the density altitude and the torque margin, which is estimated by the regression algorithm. As data-based algorithm, we experimented with deep learning, k-NN, logistic regression, and LightGBM (Ke, Meng, Finley, Wang, Chen, Ma, Ye, and Liu, 2017). Among these, LightGBM exhibited the highest testing data score. Through visualizing this high scored estimation result, we were able to recognize that "Hypothesis 2" is valid for the testing data. Therefore, we adopted LightGBM as a data-based algorithm and were able to improve the accuracy through rule-based processing described later. In LightGBM, the parameters are optimized using Optuna (Akiba, Sano, Yanase, Ohta, and Koyama, 2019). The dataset is randomly divided into five subsets. Four of these subsets are used for training, and the remaining one is used for validation. This process is repeated five times, each time with a different validation set. Based on "Hypothesis 1" from the preliminary analysis, it is postulated that contiguous points are likely to share similar health states. Therefore, in the k-NN regression (k=5), each of the divided datasets is adjusted to account for temporal continuity. The results from these five sets are averaged (ensembled) to estimate the health states. Each dataset is not standardized. In the rule-based processing, the machine learning results are corrected as necessary based on domain knowledge of "Hypothesis 2". From "Hypothesis 2" it is postulated that faulty engines tend to have a small torque margin or healthy engines also have a reduced torque margin under high density altitude conditions. Therefore, we designed our rule-based processing with a threshold based on density altitude and torque margin.

Table 1. Rule-based processing

No	conditions	consequence
1	(da < 500) and (trq_margin < -0.01 da + 2.5)	faulty
2	(da ≧ 500) and (trq_margin ≧ -0.01 da + 2.5)	healthy

Table 2. Score for the testing dataset

No	Classification Algorithm	Addition of explanatory variables	Class 0 conf	Testing Data Score
1	LightGBM	No	0.7	0.7973
2	LightGBM	Yes	0.7	0.8636
3	LightGBM + k-NN	Yes	0.7	0.8641
4	LightGBM + k-NN + rule based	Yes	0.7	0.9016
5	LightGBM + k-NN + rule based	Yes	1.0	0.9990

5. ESTIMATION OF THE HEALTH STATE

Through machine learning based on the training dataset and processing based on domain knowledge, we estimate the health status of the testing dataset and the validation dataset. Table 1 shows the rule-based processing applied to the testing and validation datasets. Table 2 shows the testing data score obtained for the classification algorithm, explanatory variables, and confidence level. In Table 2, "Addition of Explanatory Variables" marks "Yes" for those including "da" and "trq_margin". Those not including them are marked as "No". "Class 0 conf" refers to the setting value for "class_conf" for data estimated to be healthy. Initially, the reliability was maintained at 0.7, considering the impact of false-negative penalties when the accuracy was low. However, it was adjusted to 1.0 when we determined that sufficient accuracy had been achieved. From Table 2, an improvement in the score has been achieved through processing based on domain knowledge. Figures 7 to 9 show the classification results of the LightGBM, k-NN, and rule-based algorithm, which were trained on explanatory variables including "da" and "trq_margin". For the validation dataset, we applied the same algorithm as the testing dataset to estimate the health status. The results are shown in Figure 10."

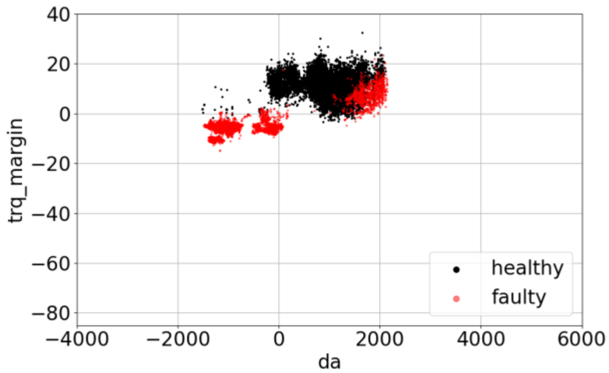

Figure 7. LightGBM result for the testing dataset

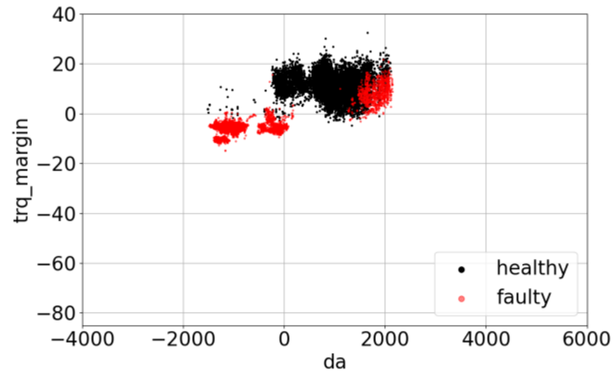

Figure 8. LightGBM + k-NN result for the testing dataset

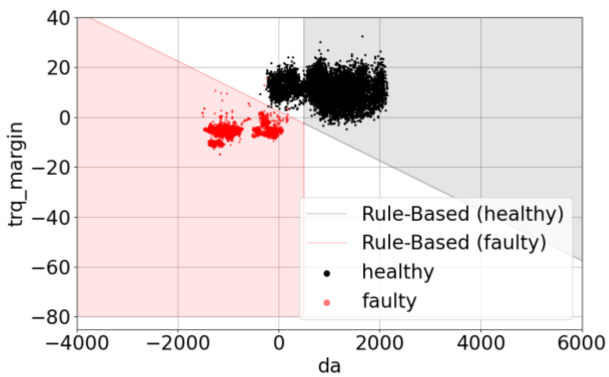

Figure 9. LightGBM + k-NN + rule-based result for the testing dataset

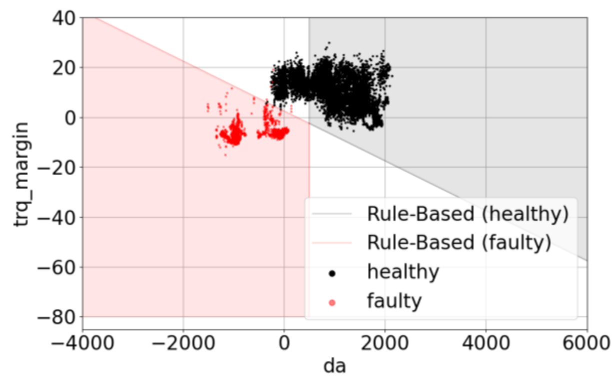

Figure 10. LightGBM + k-NN + rule-based result for the validation dataset

6. RESULTS AND DISCUSSION

Our method achieved scores over 0.99 for both the testing and validation datasets based on the calculation rules provided by the PHM 2024 Society. These were the highest scores among all teams. The preliminary analysis showed differences in asset characteristics between the training and testing/validation datasets. The health of the testing and validation datasets were estimated by combining the results of machine learning based on the training dataset and the rule-based processing derived from domain knowledge. We consider that even when estimating for new assets, it is possible to achieve a highly adaptable estimation of health by using data from past similar assets.

7. CONCLUSION

In the PHM North America 2024 Conference Data Challenge, the task was to estimate the torque margin (regression) and the health state (binary classification) of helicopter turbine engines. This paper proposes an estimation model using a hybrid algorithm that combines data-based machine learning and domain knowledge-based processing. Our method achieved scores over 0.99 for both the testing and validation datasets. These results were ranked first among all participating teams.

ACKNOWLEDGEMENT

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

REFERENCES

BHT-407-FM-3 (2018). Bell 407 Flight Manual, Section 4, Performance

Bechhoefer, E., & Kessler, J. (2022). Integrated Vehicle Health Monitoring on a Truck. *Proceedings of the Annual Conference of the PHM Society 2022*. vol. 14, no. 1. doi: 10.36001/phmconf.2022.v14i1.3162

Bechhoefer, E., & Hajimohammadali, M. (2023). Process for Turboshaft Engine Performance Trending. *Proceedings of the Annual Conference of the PHM Society 2023*. Vol. 15, no.1. doi: 10.36001/phmconf.2023.v15i1.3490

Zhou, J., Zhou, H., Li, Y., & Ca, F. (2023). Research on Anomaly Detection Method for the Executive Motor of Six-Rotor Unmanned Aerial Vehicle. *2023 IEEE 7th Information Technology and Mechatronics Engineering Conference (ITOEC)*. pp. 1168-1173. doi: 10.1109/ITOEC57671.2023.10291299

Huber, L. G., Palmé, T., & Chao, M. A. (2023). Physics-Informed Machine Learning for Predictive Maintenance: Applied Use-Cases. *In 2023 10th IEEE Swiss Conference on Data Science (SDS)*. pp. 66-72. doi: 10.1109/SDS57534.2023.00016

Cao, J., Xu, J., Huang J., & Yang, W. (2022). Towards Practical onboard PHM System for Aircraft Engines. *International Conference on Computational Science and Computational Intelligence (CSCI).* pp. 419-422. doi: 10.1109/CSCI58124.2022.00081.

Luo H., & Zhong, S. (2017). Gas turbine engine gas path anomaly detection using deep learning with Gaussian distribution. *2017 Prognostics and System Health Management Conference (PHM-Harbin)*, pp. 1-6, doi: 10.1109/PHM.2017.8079166.

Chakraborty, S., Sarkar, S., Ray A., & Phoha, S. (2010). Symbolic identification for anomaly detection in aircraft gas turbine engines, *Proceedings of the 2010 American Control Conference,* pp. 5954-5959, doi: 10.1109/ACC.2010.5531246.

Tolani, D., Yasar, M., Chin, S., & Ray, A. (2005). Anomaly detection for health management of aircraft gas turbine engines. *Proceedings of the 2005, American Control Conference,* pp. 459-464 vol. 1, doi: 10.1109/ACC.2005.1469978.

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)*. doi:10.1109/ICDM50108.2020.00147

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. doi:10.1007/s10618-022-00876-7

Kato, Y., Kato, T., & Tanaka, T. (2023). Anomaly Detection in Spacecraft Propulsion System using Time Series Classification based on K-NN. *Proceedings of the Asia Pacific Conference of the PHM Society 2023*. vol. 4, no. 1. doi: 10.36001/phmap.2023.v4i1.3675

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T. Y. (2017). LightGBM: A highly efficient gradient boosting decision tree. *Advances in Neural Information Processing Systems 2017*. vol. 30. doi: 10.5555/3294996.3295074

Akiba, T., Sano, S., Yanase, T., Ohta, T., & Koyama, M. (2019). Optuna: A Next-generation Hyperparameter Optimization Framework. Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '19). doi: 10. 48550/arXiv.1907.10902

PHM Society. (2024). PHM North America 2024 Conference Data Challenge. Retrieved 13 Sep 2024 from https://data.phmsociety.org/phm2024-conference-data-challenge

PHM Society. (2023). PHM Asia-Pacific 2023 Conference Data Challenge. Retrieved 13 Sep 2024 from https://data.phmsociety.org/phmap-2023-data-challenge

Airplane Academy. (2024). Torque, ITT, NP, and NG Explained on Turboprop Engines. Retrieved 13 Sep 2024 from https://airplaneacademy.com/torque-itt-np-and-ng-explained-on-turboprop-engines/

BIOGRAPHIES

Kosei Ozeki received his Ph.D. degree from Hokkaido University in 2024. Since 2024, he has been with Mitsubishi Electric Corp., Japan. His research interests include sound analysis and machine learning.

Takahiko Masuzaki received his Ph.D. degree from Kyoto University in 2009. Since 2006, he has been with Mitsubishi Electric Corp., Japan. His research interests include data analysis and machine learning. He is a member of the IEEE.

Takeru Shiraga received his M.S. degree from Waseda University in 2017. Since 2017, he has been with Mitsubishi Electric Corp., Japan. His research interests include data analysis and machine learning.

Koji Wakimoto received his B.S. and M.S. degrees from Kyoto University in 1983 and 1985 respectively and received his Dr. Eng. degree from the University of Tokyo in 1996. Since 1985 he has been with Mitsubishi Electric Corp., Japan. His research interests include data analysis, computer vision and geospatial information systems. He is a member of the IEEE.

Takaaki Nakamura received his M.S. degree from Kyushu University in 2002. Since 2002, he has been with Mitsubishi Electric Corp., Japan. His research interests include data analysis and machine learning.