Assessment of simulation-based data augmentation technique by Uncertainty Quantification for spacecraft propulsion system PHM.

Shotaro Hamato[1], Himeko Yamamoto[2], Noriyasu Omata[3], Yu Daimon[4] and Seiji Tsutsumi[5]

[1,2,3,4,5] *Japan Aerospace Exploration Agency Research and development directorate, Sagamihara, Kanagawa, 252-5210, Japan*

hamato.shotaro@jaxa.jp

ABSTRACT

One of the challenges of applying Prognostics and Health Management (PHM) in industrial systems is the lack of labelled training data including anomalies and faults. This study proposes training data generation by a physics-based numerical model and uncertainty quantification (UQ) considering input uncertainty and model form uncertainty, and demonstrates the proposed methodology in a spacecraft propulsion system. A one-dimensional numerical model of the spacecraft propulsion system has been developed in which ignition delay and trapped bubble dynamics are modeled. Sources of uncertainty originating in input variables of the numerical model are identified by domain experts. The probability distributions of them are modeled as uniform distributions, and training data are generated through the propagation of these probability distributions using a Monte Carlo approach. The generated training data were compared with available experimental data and showed good agreement in time-series and frequency-domain response. The 95% confidence interval (C.I.) of total uncertainty, integrating input uncertainty and model form uncertainty, was evaluated through UQ. The generated data enables the use of unsupervised methods for anomaly detection. The C.I. can be used as the normal space for anomaly detection.

1. INTRODUCTION

Data-driven approaches are widely used for PHM (Prognostics and Health Management) (Tahan, M. et al., 2017; Wang, D., Tsui, K.-L., & Miao, Q., 2018; Lee, J. et al., 2014) and anomaly detection. However, in real-world industrial systems, anomaly data available as training data for machine learning is quite limited. In more critical systems such as spacecraft and launch vehicles, even normal data available is limited. When conducting PHM in such systems, where the normal and anomaly data are not sufficient, it is important to leverage domain knowledge of the target system.

In some cases, numerical models of the target system incorporating domain knowledge are available. By utilizing the numerical model to generate training data for data-driven approaches, it becomes possible to conduct health management even for systems where training data is insufficient(Omata, N. et al., 2022; Satoh, D. et al., 2020).

It is desirable that the data generated from the models reproduce the behavior of the target system. However, in practice, numerical models contain errors. Constructing the perfect numerical model replicating the behavior and the underlying physics is extremely challenging. In addition, since real-world systems are affected by the uncertainties that come from the surrounding environment and the system itself, observations from the system have stochastic perturbations. Consequently, it is essential to take such uncertainties into account for training data generation. The quality of the numerical model is also important. Uncertainty Quantification (UQ) can evaluate the numerical simulation results in terms of input uncertainty and model error. This assessment gives us insights that how we can improve the plausibility of the generated training data by adding another physics to the numerical model for more accurate prediction or conducting additional experiment for more plausible input uncertainty. That activity helps to guarantee the quality of the generated training data.

In this study, we focus on the spacecraft propulsion system as the target system and generate training data by using the numerical model. The numerical model of the spacecraft propulsion system is evaluated through the UQ technique.

2. SPACECRAFT PROPULSION SYSTEM

Spacecraft propulsion system plays an essential role in docking to the International Space Station (ISS). The system consists of propellant and oxidizer tanks, feed lines, combustion chambers connected to the end line of the feed lines, and nozzles. Multiple nozzles are installed on the different faces and generate thrust in different directions. The propellant/oxidizer supply to the combustion chamber is controlled by opening and closing valves. If a valve failure occurs, thrust in a particular direction may be lost. During critical phases such as landing or docking, failure may cause dangerous situations that endanger not only the system itself but also the nearby systems and human lives. Therefore, it is crucial to identify the failure valves as soon as possible. In real spacecraft, the dynamic pressure response on feed lines is available. Some previous research was conducted to identify fault locations based on the pressure response characteristics.

In this study, we focus on the spacecraft for the Martian Moons eXploration (MMX) project. The mission of this project is to make a sample return from the Martian moon Phobos. In this mission, the spacecraft plans to land on the Phobos ground, and the propulsion system plays an important role. The propulsion system dynamic characteristics were validated in a system-level Static Firing Test (SFT). In the flight, the system is equipped with two main thrusters and twenty small thrusters (RCS: Reaction Control System). In the SFT, the test scope was limited to the two main thrusters and ten RCSs. The schematic diagram of the propellant (hydrazine) feed lines in the SFT is illustrated in Fig.1.

The configuration of the MMX propulsion system in the SFT was modelled(Daimon, Y. et al., 2024). The modeling approach is based on fundamental tests using a single combustor and physical understanding using Computational Fluid Dynamics (CFD). The developed model was implemented as a Modelica-based one-dimensional CAE model, which incorporates the combustion pressure prediction model(Inoue, C. et al., 2021), ignition timing prediction model(Daimon, Y. et al., 2023), and trapped bubble effect model(Yamamoto, H. et al., 2023). This numerical model enables one to predict the dynamic pressure response during valve operation sequences, and the results have been validated against the experimental results.

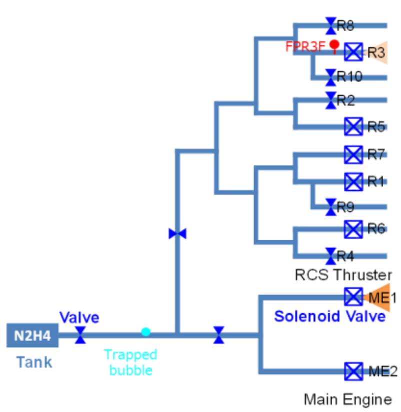

Fig.1 Schematic diagram of MMX propellant feed lines.

3. SIMULATION-BASED DATA AUGMENTATION

In the previous research(Tominaga, K. et al., 2023), since the valve failure affects the frequency response and modes of the pressure dynamics, our target data is the time series and frequency response data of the pressure dynamics in feed lines. In the field of image recognition, data augmentation is performed by simply rotating, scaling, and flipping images. However, in this study, the dataset we want to generate is numerical data obtained as the results of physical phenomena. Consequently, it is necessary to consider the underlying physics and uncertainties behind the data.

The measurement data obtained from the real-world system contains uncertainty. For example, the measurement results vary even in the same operational conditions due to the external factors such as the surrounding environment and the internal uncertainties the system itself has. In contrast, numerical models generate unique and deterministic outputs for the obtained input conditions.

Sources of uncertainty, which are expected to vary among multiple test cases, are recognized from the input variables of the MMX propulsion system model through a discussion with propulsion system experts, focusing on only the propellant feed lines. The recognized sources of uncertainty are shown in Table 1.

Table.1 Sources of uncertainty in the propellant feed lines of the MMX propulsion system model.

Variable	Distribution
Fuel speed of sound	Uniform
Valve opening/closing time delay	Uniform
Fuel inlet tank pressure	Uniform
Outlet tank pressure	Uniform
Trapped bubbles	Uniform

The speed of sound in the propellant varies depending on its temperature and the vapor deposition of dissolved helium, which could result in a two-phase flow. The valve opening and closing duration changes due to the pressure difference across the valve and manufacturing tolerances. Although the propellant tank pressures are regulated, they are not perfectly constant but vary stochastically. In the SFT, since some valve outlets were connected not to combustion chambers but to large tanks, the back pressure of the tanks are identified as the sources of uncertainty. Small gas bubbles may be trapped in the branching section or at the pressure sensor mounting. These trapped bubbles are difficult to measure their volume. The volumes of the bubbles are estimated based on data assimilation to the experimental results. The identified bubble locations are shown in Fig.1.

To model the identified sources of uncertainty as probability distributions, enough experimental data is required. However, the available experimental data on each operational condition were limited. Therefore, each probability distributions of the sources of uncertainty are defined as uniform distributions based on the limited experimental data and expert knowledge. From the parameter space defined as uniform distributions, 1024 points are sampled using Latin Hypercube Sampling (LHS). A Monte-Carlo simulation was conducted to generate synthetic training data. The target operational mode is that only the R3 thruster was operated for three pulses, where a single pulse is defined as the set of the opening and closing. The pressure time-series data was measured at the sensor FPR3F in the propellant feed lines. The location of the sensor and the operated thruster are illustrated in Fig.1. A Fast Fourier Transformation (FFT) was applied to the closing duration of the final pulse in the measured time-series data to examine not only the time-series data but frequency-domain response.

4. Uncertainty Quantification

In this study, we performed Uncertainty Quantification (UQ) with reference (Roy, C. J. & Oberkampf, W. L., 2011). In this method, uncertainty consists of three uncertainties.

1. Numerical approximation
2. Input uncertainty
3. Model form uncertainty

Numerical approximation is a numerical error related to iteration error for iteration computation and discretization error, for example. Input uncertainty evaluates the uncertainty propagated from the input variables to the output variables. Specifically, modeling the sources of uncertainty as probability distribution, sampling from these distributions and conducting Monte Carlo simulations enable to obtain the probability distribution of the output variables. Model form uncertainty is the evaluation of the error between the varying experimental results and the numerical solutions. Illustrating the input uncertainty and the experimental results in Cumulative Density Function (CDF) and calculating the area covered with these two CDFs obtain the evaluation of the error between the experiment and the simulation. This area is called Area Validation Metric and is equal to the expectation of the difference between these two probability distributions.

For anomaly detection of the valve failure, since frequency-domain response is important (Tominaga, K. et al., 2023), we chose the first modal frequency of the closed duration in the operation sequence as Quantity of Interest (QoI). The results of FFT can be affected by sidelobes and frequency resolution due to signal length and window function settings. Since the pressure time-series dynamics in closed duration closely resemble a dumped sinusoidal waveform, we evaluated the modal frequency by modal decomposition using a superposition of damped sinusoidal waveforms. By optimizing the parameters of the damped sinusoidal waveform to match the original time-series data, the frequency of the signal can be estimated more reliably. In this study, we evaluated input uncertainty, model form uncertainty, and total uncertainty.

5. Results and Discussion

5.1. Data augmentation

The synthetic training data are compared with the experimental data. Figure 2 illustrates the pressure time-series dynamics on the sensor FPR3F in the propellant feed lines. Randomly chosen a single data out of 1024 cases is shown.

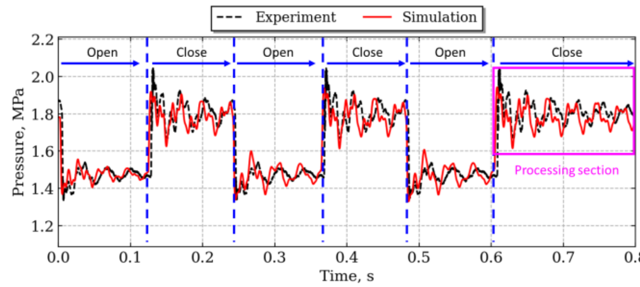

Fig.2 Pressure time-series dynamics on the sensor FPR3F in the propellant feed lines.

Since the tank pressure measured during the test campaign was used as the simulation input, the mean pressure in the opening and closing duration agrees with the experimental results. Not only the mean but the frequency and the amplitude agree as well.

The generated 1024 cases are compared with the experiment. In this figure, 128 samples out of 1024 samples are randomly chosen for ease of visibility. Figure 3 shows the pressure time-series dynamics of the 128 cases on the sensor FPR3F.

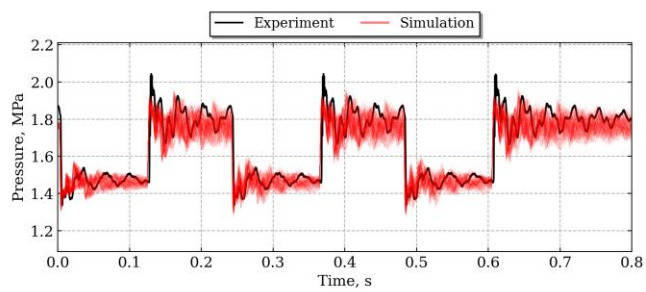

Fig.3 Pressure time-series dynamics on the sensor FPR3F. Comparison between the synthetic 128 cases and the experimental data.

The mean pressure in the opening duration agrees with the experiment. However, the mean pressure in the closing duration is lower than the experimental value. The tank pressure affects the mean pressure in the closing duration, which is regulated in the morning of each test day and gradually decreases as the experiments proceed. The range of tank pressure was determined from the minimum and maximum values observed during a test day. Since the experimental result shown in Fig.3 was obtained from the first test of the day, the mean pressure of the experiment is higher than the generated data. The surge pressure of the generated data is lower than the experiment, which is influenced by multiple factors such as valve closing speed and trapped bubbles. In this study, since the focus is on modal frequencies in the frequency domain, the errors on the mean pressure and surge pressure are not critical. Overall, the generated data generally encompasses experimental result, indicating that assumption of the sources of uncertainty was appropriate.

In addition to the time-series dynamics, the frequency-domain response was examined. The duration shown in Fig.2, ranging from 0.6 to 0.8s, was processed by FFT. Figures 4 illustrate the frequency-domain response from the sensor FPR3F. As with Fig.3, 128 samples were randomly chosen out of 1024 samples in Fig.4.

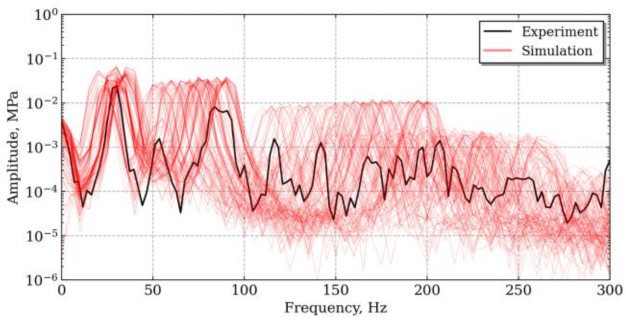

Fig.4 Frequency-domain response of the pressure dynamics on the sensor FPR3F. Comparison between the synthetic 128 samples and the experimental data.

The frequency response exhibits modes associated with the propellant feed line geometry and the trapped bubbles. Additional modes emerge by the effect of the bends and branch points in the feed lines. The biggest mode around 28Hz corresponds to a primary mode of the feed line, which is reproduced in the generated data. The modal frequency varies from 20 to 40Hz depending on the sources of uncertainty. The modal frequency around 90Hz also comes from the feed line geometry, which varies from 50 to 100Hz. The modal frequencies over 100Hz have minor discrepancies between the experimental and simulation results. Since these higher modes are influenced by the location and volume of the trapped gases, the assumption of the trapped gas location may be different from the experimental situation. Although minor discrepancies in the higher frequency area exist, the generated training data successfully reproduced both the time-series and frequency-domain characteristics. The generated training data enables one to conduct unsupervised learning-based anomaly detection.

5.2. Uncertainty Quantification

UQ was performed for the first modal frequency from FPR3F in Fig.4. Modal decomposition by a superposition of dumped-sinusoidal waveform was applied to the closing duration in Fig.2 to evaluate the first modal frequency. Figure 5 shows the CDF of the first modal frequency extracted from the 1024 samples and the model form uncertainty. Since only the single experimental data was available, it is represented as a single straight line in Fig.5.

From the CDF of the numerical simulation, the 95% Confidence Interval (C.I.) ranges from 29.09 to 31.15Hz. The error between the generated data and the experimental data was evaluated based on area validation metric and was found to be 1.29Hz. The 95% C.I. of the total uncertainty based on the input uncertainty and model form uncertainty ranges from 27.79 to 32.45Hz. This C.I. can be utilized as a normal space for anomaly detection. Increasing the amount of experimental data and modeling the sources of uncertainty with more plausible probability distributions will improve the fidelity of the synthetic dataset and the C.I., and enhance the reliability of PHM in data-limited systems.

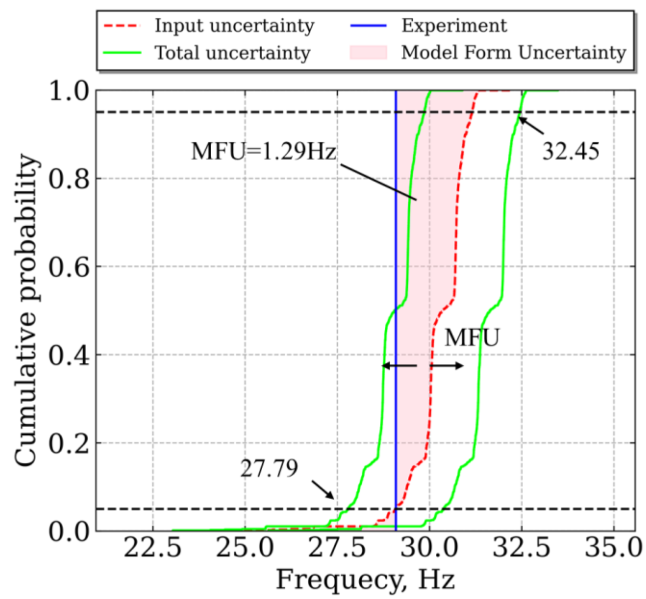

Fig.5 Assessment of the generated training data in the form of input uncertainty and model form uncertainty.

6. Conclusion

In this study, we focused on the spacecraft propulsion system and investigated synthetic data generation for PHM in systems where available normal and anomaly data for machine learning are quite limited. Synthetic training data was generated by the 1D-CAE numerical model incorporating domain knowledge. We applied Uncertainty Quantification (UQ) to assess the generated training data in comparison to the experimental data. Sources of uncertainty from the input variables in the numerical model of the system-level static firing test were identified through discussions with propulsion system experts and modelled as probability distributions. From these sources of uncertainty, Latin Hypercube sampling and Monte Carlo simulations on the numerical model enable one to generate synthetic training data. Comparison with the experimental data in the same operational conditions showed that the generated data successfully reproduced the time-series dynamics and frequency-domain response characteristics. These generated training data can be utilized for unsupervised learning-based anomaly detection. In addition, UQ was conducted for the first modal frequency, which is important for valve failure detection, to evaluate input uncertainty, model form uncertainty, and total uncertainty. Taking into account both the effect of the sources of uncertainty and the inherent error of the 1D-CAE model, the 95% confidence interval of the predictions was evaluated. This confidence interval can be used as a normal space for anomaly detection. Modeling the probability distribution of uncertainty sources more plausibly by increasing the amount of available experimental data will improve the reliability of unsupervised learning-based and confidence interval-based anomaly detection. The proposed approach provides a means of health monitoring for data-limited systems such as spacecraft. In future application, it is expected to contribute to autonomous health management and state-dependent operation in Orbital Transfer Vehicles (OTVs) and planetary exploration spacecraft.

ACKNOWLEDGEMENT

The authors sincerely thank Mr. Takamine Imada, Prof. Yusuke Maru, and Mr. Junichi Nakatsuka of the MMX project team for their generous provision of valuable experimental data obtained from system-level firing test.

REFERENCES

Daimon, Y. et al. (2023). One-dimensional modeling of ignition timing for hypergolic bipropellant thrusters. *Proceedings of the Aerospace Europe Conference - EUCASS - CEAS - 2023*. https://doi.org/10.13009/EUCASS2023-820.

Daimon, Y. et al. (2024). LIQUID propulsion system simulation validated by the mmx system firing tests. *Space Propulsion Conference 2024*, GLASGOW.

Inoue, C., Oishi, Y., Daimon, Y., Fujii, G., & Kawatsu, K. (2021). Direct formulation of bipropellant thruster performance for quantitative cold-flow diagnostic. *Journal of Propulsion and Power*, 37 (6), 842–849. https://doi.org/10.2514/1.B38310.

Lee, J. et al. (2014). Prognostics and health management design for rotary machinery systems—reviews, methodology and applications. *Mechanical Systems and Signal Processing*, 42 (1–2), 314–334. https://doi.org/10.1016/j.ymssp.2013.06.004.

Omata, N. et al. (2022). Model-based fault detection with uncertainties in a reusable rocket engine. *2022 IEEE Aerospace Conference (AERO)*. https://doi.org/10.1109/AERO53065.2022.9843212.

Roy, C. J. & Oberkampf, W. L. (2011). A comprehensive framework for verification, validation, and uncertainty quantification in scientific computing. *Computer Methods in Applied Mechanics and Engineering*, 200 (25–28), 2131–2144. https://doi.org/10.1016/j.cma.2011.03.016.

Satoh, D., Tsutsumi, S., Hirabayashi, M., Kawatsu, K., & Kimura, T. (2020). Estimating model parameters of liquid rocket engine simulator using data assimilation. *Acta Astronautica*, 177, 373–385. https://doi.org/10.1016/j.actaastro.2020.07.037.

Tahan, M., Tsoutsanis, E., Muhammad, M., & Abdul Karim, Z. A. (2017). Performance-based health monitoring, diagnostics and prognostics for condition-based maintenance of gas turbines: a review. *Applied Energy*, 198, 122–144. https://doi.org/10.1016/j.apenergy.2017.04.048.

Tominaga, K. et al. (2023). Anomaly detection method for spacecraft propulsion system using frequency response functions of multiplexed fbg data. *Acta Astronautica*, 212, 235–245. https://doi.org/10.1016/j.actaastro.2023.07.022.

Wang, D., Tsui, K.-L., & Miao, Q. (2018). Prognostics and health management: a review of vibration based bearing and gear health indicators. *IEEE Access*, 6, 665–676. https://doi.org/10.1109/ACCESS.2017.2774261.

Yamamoto, H., Kawatsu, K., Daimon, Y., & Fujii, G. (2023). Development of a dynamic response model considering trapped gas effect for spacecraft liquid propulsion systems. *Journal of Evolving Space Activities*, 1. https://doi.org/10.57350/jesa.34.