# Systemic symptom detection in telemetry of ISS with explainability using FRAM and SpecTRM

Shota Iino[1], Hideki Nomoto[1], Takashi Fukui[2], Sayaka Ishizawa[2], Miki Sasaki[2], Yohei Yagisawa[2], Takayuki Hirose[1], Yasutaka Michiura[1] and Hiroharu Shibayama[3]

[1] *Japan Manned Space Systems Corporation, Tokyo, 100-0004, Japan*
*iino.shota@jamss.co.jp*
*nomoto.hideki@jamss.co.jp*
*hirose.takayuki@jamss.co.jp*
*michiura.yasutaka@jamss.co.jp*

[2] *JAPAN NUS Co., Ltd., Tokyo 160-0023, Japan*
*fukui-t@janus.co.jp*
*ishizawa-syk@janus.co.jp*
*yagisawa-yhi@janus.co.jp*

[3] *Space Engineering Development Co., Ltd., Tokyo 164-0001, Japan*
*shibayama.hiroharu@sed.co.jp*

## ABSTRACT

Explainability is important for machine learning-based anomaly detection of safety critical systems. In this respect, we propose a new systemic symptom detection method by combining two methodologies: the Functional Resonance Analysis Method (FRAM) and the Specification Tools and Requirement Methodology-Requirement Language (SpecTRM-RL) with machine learning-based normal behavior prediction model. The method was verified with data of thermal control system of Japanese Experimental Module of the International Space Station, and the result found that the proposed method enables flight controllers and specialists to obtain additional information for identifying causes of anomaly with the method.

## 1. INTRODUCTION

In International Space Station (ISS), several systems are operating to maintain the environment. Although those systems have high reliability, there are some anomalies for systems. Several flight controllers are monitoring the status of systems for 24 hours, 365 days a year. If an anomaly is detected, flight controllers assess the trends of telemetries and impacts for operations. Experienced flight controllers can detect symptoms of anomaly by unusual combinations of telemetries (funny data). However, it is difficult to define those unusual combinations because the numbers of combinations will be huge (at least 2^30 for 30 telemetries for just binary type parameters such as TRUE/FALSE). But machine-learning based model enables anomaly prediction by combinations without wasting huge state space.

## 2. AUTOMATIC ANOMALY DETECTION

Automatic anomaly detection methods were proposed by several authors. Especially machine-learning based anomaly detection methods are widely used. As an example of application in aerospace, Wang et al. (2019) proposed diagnostic health monitoring for in-orbit spacecrafts. Whereas such methods provide high accuracy for anomaly detection, explainability for operators is lacking. To apply automatic anomaly symptom detections methods to ISS operations, it is required to provide flight operators with the rationale for the prediction because they cannot take actions without justification. GalaxAI demonstrated an interpretable end-to-end analysis of spacecraft telemetry data providing mission specialists and operators with an interpretable view of the data analysis process (Kostovska et al., 2021). Zeng et al. (2022) proposed an anomaly detection framework to discover the complex relationships of telemetries of spacecraft based with known causal relationships. Those methods are effective for limited number of telemetries with known anomaly events. However, there are several kinds of anomalies for ISS systems including unknown events. Therefore, we propose a method to provide enough information for ISS flight controllers and specialists to assess the current situations of systems and possible causes. This paper explains the process of symptom detections for ISS operations and designs an automatic method to detect symptoms of anomaly with additional information for explaining reasons of detections.

## 3. Methodology

We present a systemic symptom detection method combining the Functional Resonance Analysis Method (FRAM) and the Specification Tools and Requirement Methodology-Requirement Language (SpecTRM-RL) with machine learning-based anomaly detection. Figure 1 shows flow of our proposed method.

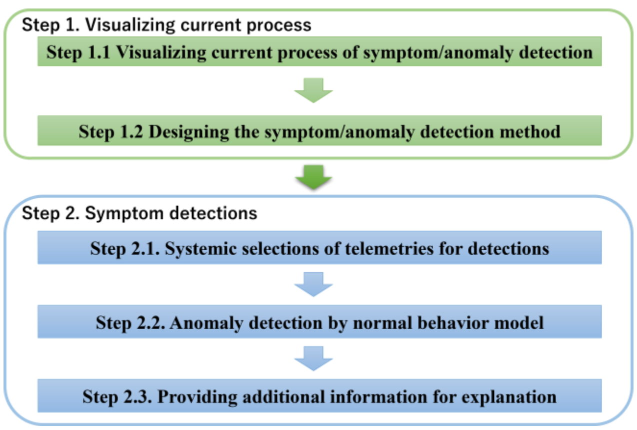


**Figure 1. Flow of the proposed method**

### 3.1. Visualizing current process of heuristic symptom detection

Firstly, it is important to understand current process of heuristic symptom detection for ISS operations to design efficient symptom detections for flight controllers. To understand the process, we visualized functions for current operations with Functional Resonance Analysis Method (FRAM). FRAM was proposed by Hollnagel (2012), which can be utilized to investigate complex interaction among identified functions of target systems; models are defined with six aspects listed in Table 1 and graphically represented. FRAM model provides insights about the dynamics of the target systems or how they actually work.

**Table 1. Six aspects of FRAM.**

<table>
<tr><th colspan="3">Class</th><th>6 aspects</th></tr>
<tr><td rowspan="5">Input</td><td colspan="2">Trigger</td><td>Input</td></tr>
<tr><td colspan="2">Prior condition</td><td>Precondition</td></tr>
<tr><td rowspan="3">Posterior condition</td><td>change output</td><td>Control</td></tr>
<tr><td rowspan="2">stop output</td><td>Resource</td></tr>
<tr><td>Time</td></tr>
<tr><td colspan="3">Output</td><td>Output</td></tr>
</table>

### 3.2. Designing symptom detection method

As a result of requirement analysis, we found the following two important factors of automatic symptom detection for ISS operations: One is importance of understanding complex relationship of functions to select proper telemetries from hundreds of telemetries. Systemic approach is required. The other is importance of additional information to assess telemetries related to the anomaly. We designed the detection method to provide systemic selections and additional information for assessing the trends of telemetries with machine learning-based detection methods. There are three steps: 1) systemic selections of telemetries for detections understanding interactions of functions, 2) machine learning-based anomaly detection, and 3) providing additional information for explanation with assessing the combinations of conditions in normal and abnormal period.

### 3.3. Systemic selections of telemetries for detections

There are more than one hundred telemetries related to the Thermal Control Assembly - Low (TCA-L) of the Japan Experimental Module (JEM). To make anomaly detection models for specific anomaly events, it is necessary to select objective and explanatory valuables. Iino et al. (2022) made functional model to understand relationships of telemetries using FRAM. This paper extends the FRAM modeling for larger scope with further interviews with specialists.

### 3.4. Anomaly detection by normal behavior model

Out-of-limits approach is widely applied for aerospace systems. An alarm is issued when the value of each telemetry is over predefined thresholds. There are two limitations. One is that it is difficult to detect symptoms of anomaly for combinations of multiple telemetries. The other one is that ranges of telemetry values vary depending on the operation mode and conditions. Anomaly detection using normal behavior model-based approach has been proposed as an alternative to out-of-limits approach. This study also adopts the latter approach. The specific procedure of normal behavior model based approach in this study consists of the following three steps: (1) A normal behavior model representing the relationship between the objective variable $y_t$ and the explanatory variables $x_t$ at a certain time step $t$ in the normal state, $y_t \approx f(x_t)$, is constructed by supervised regression machine learning. (2) The explanatory variable data at each time step of the test period are input to the constructed normal behavior model, and the value that the objective variable should take when the system is normal, $\hat{y}_t = f(x_t)$, is predicted. (3) The prediction error, i.e., difference between the predicted value and the actual value, $\epsilon_t = y_t - \hat{y}_t$, is compared to a certain threshold value $\delta$, and it is judged as a sign of anomaly when $\epsilon_t > \delta$.
In this study, Random Forest (RF), proposed by Breiman et al. (2001), was employed for the first step due to its ability

to learn the nonlinear relationship between the objective variable and many explanatory variables with low computational cost. Our RF model was implemented by using scikit-learn library in Python. For the next step, to avoid false alarms in normal conditions and missed alarms in abnormal conditions, appropriate threshold values were set as $\delta = n\sigma$ ($n = 1, 2, 3, 4$), where $\sigma$ is the standard deviation of the prediction error under normal conditions.

### 3.5. Model selection by comparing results of models

Several models for anomaly detection can be made for different purposes. For model selections, there are some criteria for model selections quantitatively and qualitatively. Different evaluation indicators have both merits and demerits. Therefore, we applied Pugh Concept Selection to select models from multidimensional views. Pugh Concept Selection, proposed by Pugh (1981), is a tool to control convergence to the best solution considering multiple criteria. To compare models from engineering view, FRAM model can be used for understanding which models are appropriate for purposes. To compare the performance of predictions of objective variable qualitatively, Root Mean Square Error (RMSE) was calculated for each model. Lastly the time for detecting symptoms of anomaly and the clarity of detection is also important from operational views. Considering those multidimensional views, we selected models for alert simulations.

### 3.6. Providing additional information for explanation

After receiving alerts of symptom detections, flight controllers and specialists need to understand the reasons of detections before performing any actions for trouble shootings. Traditional machine learning-based anomaly detection methods is black-box typed algorithms for explanations. Therefore, we performed additional assessment to provide additional information for narrowing down possible causes with SpecTRM-RL. SpecTRM-RL is one of the formal methods proposed by Leveson, N. Firstly, we make two-dimensional table consisting of each parameter to lateral direction and time-series to vertical direction. True (T) or False (F) are filled based on the condition of each parameter. The condition of each parameter was defined with the average and standard deviation as shown in equation (1).

$$m_i - 3\sigma_i \leq x_i \leq m_i + 3\sigma_i, \qquad (1)$$

where $x_i$, $m_i$, and $\sigma_i$ represent the value, average, and standard deviation of the $i$-th parameter, respectively. Then SpecTRM-RL algorithm identifies combinations of True, False or wildcard (*). Comparing the combinations of each parameters before and after symptoms of anomaly happens will help flight controllers or specialists to assess the trends of telemetries.

## 4. Results

### 4.1. Modeling current process of symptom detections

Figure 2 shows FRAM modelling of the process. Flight controllers monitor telemetries of assigned ISS operations. Then they find unusual trends for individual telemetry or anomaly by alerts of each telemetry if the observed values are over threshold. After symptom or anomaly detection, specialists assess the impact and perform trouble shootings for each anomaly. Our motivation is to enable them to assess symptoms with combinations of telemetries and provide additional information for further assessment.

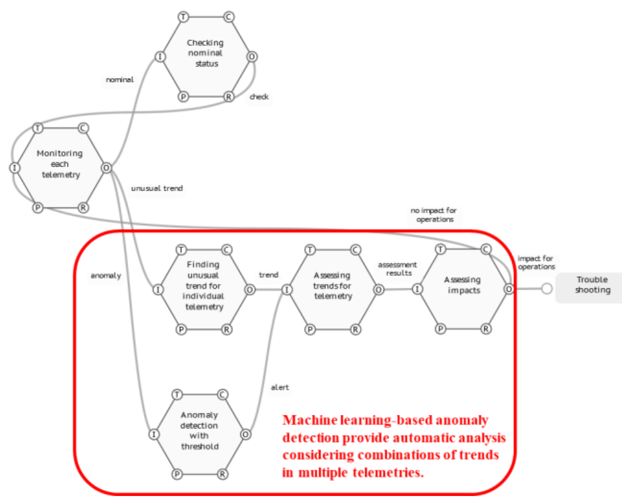


**Figure 2. FRAM modeling of symptom detection process**

### 4.2. Experimental setup

Data of ISS system in 2012 was utilized for verification. In 2012, a pump of TCA-L of the Japan Experimental Module (JEM) failed. We analyzed the downlinked data from ISS to the ground to verify our proposed method.

### 4.3. FRAM models for systemic selections of telemetries

Our FRAM model of systems related to TCA-L pump is shown in Figure 3. We made four patterns of FRAM models based on results of interviews with specialists. Selected telemetries and reasons of selections are listed in Table 2.

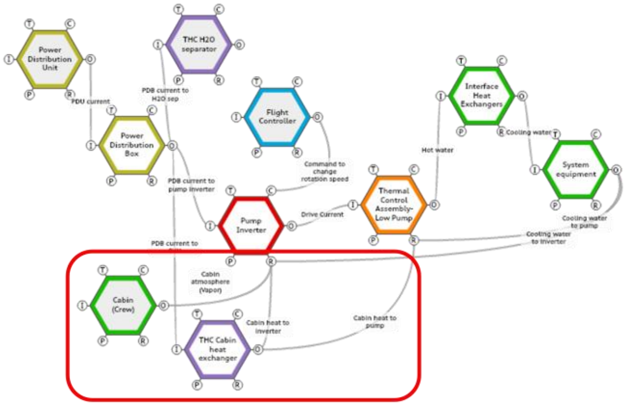


**Figure 3. FRAM model of related functions**

**Table 2. Selected telemetries for each model**

| Model No | Selected telemetries |
|---|---|
| Model 1 | 5 telemetries related to TCA-L pump. |
| Model 2 | 5 telemetries related to TCA-L pump.<br>2 telemetries related to dew point. |
| Model 3 | 2 telemetries related to dew point.<br>1 telemetry related to power. |
| Model 4 | 2 telemetries related to dew point.<br>3 telemetries related to THC. |

### 4.4. Comparison results

We calculated RMSE of each model for qualitative analysis as shown in Figure 4. Best RMSE is 0.42 for model 2 and second best is 0.958 for model 4.

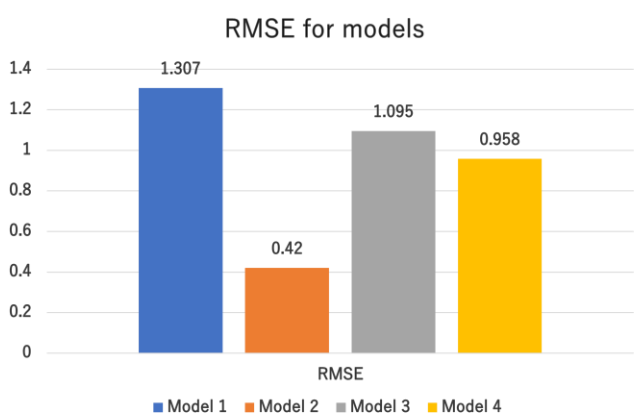


**Figure 4. RMSE of four models**

Observed (real) and predicted (pred) values of objective variable for each model were shown in

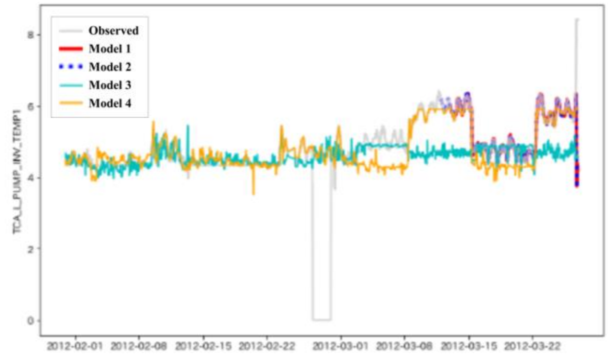


Figure 5. Blue, red, pink, yellow lines show the predicted values of model 1, 2, 3, and 4 respectively while grey line shows observed values.

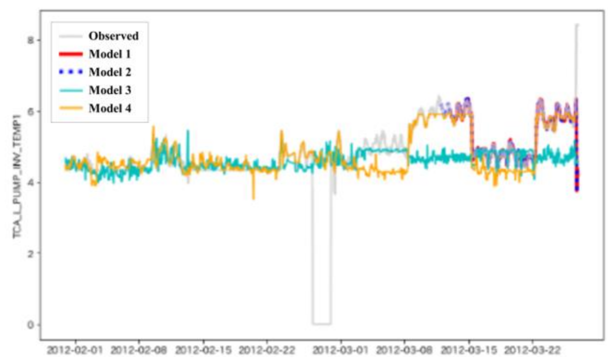


**Figure 5. Observed and predicted temperatures.**

Then, we analyzed the differences between observed and predicted values for symptom detections; we could observe symptoms of anomaly if the difference is bigger.

### 4.5. Alert simulation

We compared the results of models with Pugh Concept Selection as shown in **Error! Reference source not found.**. RMSE was lowest for model 2 while the performance of early symptom detection of anomaly was high in model 3 and 4. Discussing with specialists about the performances of models from several views**Error! Reference source not found.**, we selected model 4 for simulation as it is important to detect anomaly earlier with higher accuracy of predictions. Simulations with defined threshold were performed. We compared the simulation results with the threshold of two, three, or four-sigma. Consequently, four-sigma was chosen because the balance in the numbers of alerts was better than others. Results of model 4 is shown in Figure 6. Red points are the values over the threshold. Alerts can be released to flight controllers based on the simulations.

**Table 3. Results of Pugh Concept Selection**

| Approach | Methods | Model 1 | Model 2 | Model 3 | Model 4 |
|---|---|---|---|---|---|
| Engineering | FRAM | TCA_L | TCA_L +Dew point | Power +Dew point | CHX(cabin heat) +Dew point |
| Data | RMSE | 1.307 | 0.42 | 1.095 | 0.958 |
| Operation | Early symptom detection | 1 week~2 weeks before | 2 days~3 days before | 2 week~3 weeks before | 2 week~3 weeks before |
| | Levels of symptoms of anomaly | Low | Low | High | High |

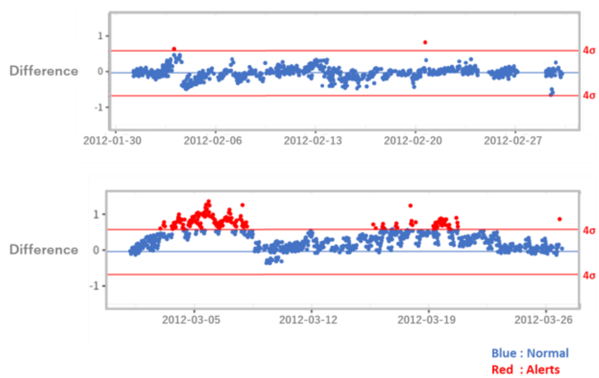


**Figure 6. Simulation results**

### 4.6. Providing additional information for explanation with SpecTRM

We conducted a statistical analysis on the data of each telemetry. We set the threshold of each condition to 4σ based on the analysis. Figure 7. shows the result of SpecTRM-RL analysis. Under normal condition, there are seven combinations whereas under abnormal condition, there are two combinations. Three parameters had all false values in abnormal conditions; cabin temperature, temperature of pump inverter and cabin heat exchanger coolant out temperature had all false.

| Normal condition | | | | | | | |
|---|---|---|---|---|---|---|---|
| Service module PPH2O < 11 | T | T | F | T | T | T | * |
| Cabin temperature< 22 | F | T | * | * | T | F | F |
| Condense out pressure of water separator < 2 | T | T | T | T | T | T | T |
| Temperature of pump inverter is between 3 & 4.5 | T | F | F | * | * | F | T |
| Cabin heat exchanger coolant out temperature < 12 | T | T | F | T | F | * | F |
| Cabin heat exchanger flow rate< 280 | T | T | T | T | T | T | T |
| Number of data | 21 | 23 | 25 | 109 | 248 | 332 | 389 |

| Abnormal conditions | | |
|---|---|---|
| Service module PPH2O < 11 | F | T |
| Cabin temperature< 22 | F | F |
| Condense out pressure of water separator < 2 | T | T |
| Temperature of pump inverter is between 3 & 4.5 | F | F |
| Cabin heat exchanger coolant out temperature < 12 | F | F |
| Cabin heat exchanger flow rate< 280 | T | T |
| Number of data | 10 | 118 |

**Figure 7. Observed and predicted temperatures.**

## 5. Discussion

### 5.1. Additional information for explanation

During normal period, there are several variations for defined conditions. However, during the abnormal period, we found unique characteristics for cabin temperature and cabin heat exchanger coolant out temperature. Those telemetries are related to dew point. Low temperature of those sensors will affect for condensations of pump inverter. Japanese astronaut in ISS conducted the trouble shooting tasks for this anomaly and found the short of pump inverter due to overcurrent of power (JAXA 2013). It validated our analysis results. Figure 8 shows the highlighted functions related to possible causes in FRAM model.

In the context of safety-critical systems, it is crucial for flight controllers and specialists to understand the rationale behind anomaly detections to make any actions. Therefore, levels of detailed explanation should be discussed with flight controllers and specialists carefully in the future.

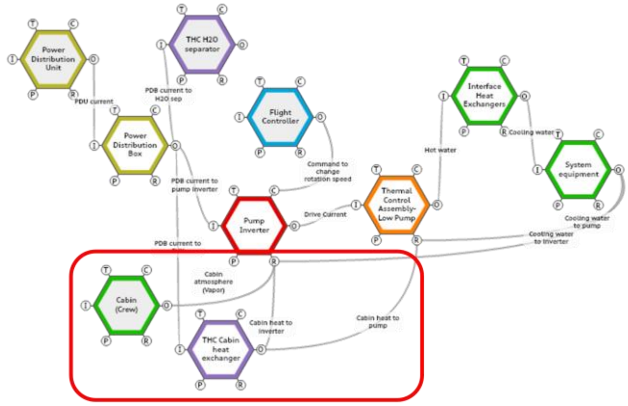


**Figure 8. Functions related to possible causes in FRAM.**

### 5.2. Defining threshold of each condition

In the experiment, we defined threshold with four sigma of each condition based on the statistical analysis. Another way of setting the threshold is using automatic model-agnostic interpretation methods such as SHAP (Lundberg 2017). For practical use, further discussions with flight controllers and specialists should be required.

### 5.3. Applying to other systems

The proposed methods can be applied to data of other systems of ISS or any systems, which have telemetry data. As other systems of ISS have different characteristics of data, further verification of the method will be required for practical use.

## 6. Conclusion

Currently, flight controllers find symptoms of ISS systems with watching trends of each telemetry manually. Automatic symptoms/anomaly detections with machine learning-based methods will help them to detect symptoms of anomaly early. As flight controllers and specialists need to know why the detections happened for performing any actions, providing additional information to assess trends of related telemetries for anomaly. We proposed a new method to provide additional information for explanations with FRAM and SpecTRM-RL. The proposed method was verified with an experiment of ISS systems. It enables us to carry out systemic analyses overcoming the limitations of previous studies which have difficulty in handling complex multiple factors. The experimental results implied the effectiveness of the method. Further experiments with other systems and discussion with flight controllers and specialists were required for practical use. The proposed method is expected to use for several safety-critical systems in aerospace and other fields.

## Acknowledgement

The authors thank some JAXA Flight controllers and specialists of Japanese Experiment Module (JEM) of International Space Station for the interviewing.

## References

Hollnagel, E. (2012). FRAM: *The Functional Resonance Analysis Method: Modelling Complex Socio-technical Systems*. Ashgate Publishing Ltd

Gao, Y., Yang, T., Xu, M., & Xing, N. (2012, January). *An unsupervised anomaly detection approach for spacecraft based on normal behavior clustering*. In 2012 Fifth International Conference on Intelligent Computation Technology and Automation (pp. 478-481). IEEE

Japan Aerospace Exploration Agency (JAXA). (2013). *Japanese Manned Space Technology brought by Development and Operation of Internatiomal Space Station (ISS), Japanese Experimet Module KIBO*. JAXA-SP-12-015, 1 - 227

Kostovska, A., Petković, M., Stepišnik, T., Lucas, L., Finn, T., Martínez-Heras, J., ... & Kocev, D. (2021, July). GalaxAI: *Machine learning toolbox for interpretable analysis of spacecraft telemetry data*. In 2021 IEEE 8th International Conference on Space Mission Challenges for Information Technology (SMC-IT) (pp. 44-52). IEEE

Leveson, N. G., Reese, J. D., & Heimdahl, M. P. (2003). SpecTRM: *A CAD system for digital automation. In 17th DASC*. AIAA/IEEE/SAE. Digital Avionics Systems Conference. Proceedings (Cat. No. 98CH36267) (Vol. 1, pp. B52-1). IEEE

Lundberg, S. M., & Lee, S. I. (2017). *A unified approach to interpreting model predictions*. Advances in neural information processing systems, 30

Pugh, S. (1981, March). *Concept selection: a method that works*. In Proceedings of the International conference on Engineering Design (pp. 497-506)

Wang, C., Lu, N., Cheng, Y., & Jiang, B. (2019). A telemetry data based diagnostic health monitoring strategy for in-orbit spacecrafts with component degradation. Advances in Mechanical Engineering, 11(4), 1687814019839599

Zeng, Z., Jin, G., Xu, C., Chen, S., & Zhang, L. (2022). Spacecraft Telemetry Anomaly Detection Based on Parametric Causality and Double-Criteria Drift Streaming Peaks over Threshold. Applied Sciences, 12(4), 1803