

# 機械学習を用いたXRISM衛星搭載極低温検出器の異常検知アルゴリズムの開発

柏崎 未有<sup>\*1</sup>, 辻本 匡弘<sup>\*2</sup>

## Development of the machine learning-based anomaly detection algorithms for the low-temperature detector onboard the XRISM satellite

KASHIWAZAKI Miu<sup>\*1</sup>, TSUJIMOTO Masahiro<sup>\*2</sup>

### ABSTRACT

Early detection of anomalies in spacecraft telemetry data is important for ensuring the safe operation of the spacecraft. The XRISM satellite, launched on September 6, 2023 (UTC), utilizes the Automatic Telemetry Monitor Software (ATMOS) for this purpose. However, ATMOS serves as a general system primarily designed for analyzing time-series telemetry data. Consequently, each mission necessitates dedicated complementary systems. In this article, we present the results of an anomaly detection algorithm applied to the Resolve instrument onboard XRISM. Our methodology entails a machine-learning approach utilizing actual data collected during ground testing. We assess the efficacy of this approach in detecting anomalies in the detector noise spectra and the cold-stage temperature data.

**Keywords:** XRISM, *Resolve*, anomaly detection, machine learning.

### 概要

安全な衛星運用を実現するため、衛星テレメトリデータの異常を早期に検知することは極めて重要である。2023年9月7日（日本時間）に打ち上げられたXRISM衛星では、異常検知システムATMOS (Automatic Telemetry Monitor Software) が使用されている。しかしATMOSはテレメトリ時系列データの閾値判定を主とした汎用システムであり、データの特性に即した異常検知にはミッションごとの異常検知システムが相補的に必要である。本研究では、XRISM衛星に搭載されたミッション機器*Resolve*のデータに対する異常検知アルゴリズムの開発を行った。地上試験データを用いて機械学習的な手法を用い、*Resolve*装置のテレメトリデータに現れる二種の異常-検出器ノイズスペクトルに含まれる異常と、検出器の温度データに含まれる異常-の検知アルゴリズムの開発を通して、同手法の有用性を検証した。

### 1. はじめに

2016年に打ち上げられたX線天文衛星ASTRO-H<sup>[14]</sup>は、衛星姿勢制御系に起因する事故により、打ち上げから約1ヶ月で運用を停止した<sup>[8]</sup>。そのため、その後継機として開発されたXRISM衛星<sup>[15]</sup>では、異常検知システムの整備が特に重要と認識されている。

現在、プロジェクトで採用されている汎用システム

ATMOS (Automatic Telemetry Monitor Software) は、衛星テレメトリデータ（もしくはその演算）の時系列データを閾値判定するものである。即時性が高い反面、複雑な判定には向いていない。相補的なシステム、すなわち即時性が低いものの複雑な判定が可能なシステムが必要である。特に運用上重要なものは、閾値判定では分からない(i)データのパターンに基づいた異常、(ii)データを一次処理しなければ発見できない異常の検知

\* 2023年11月30日受付 (Received November 30, 2023)

<sup>\*1</sup> 東京大学大学院理学系研究科 (Graduate School of Science, The University of Tokyo)

<sup>\*2</sup> 宇宙航空研究開発機構 宇宙科学研究所 (Japan Aerospace Exploration Agency, Institute of Space and Astronautical Science)

である。これらの異常を検知するアルゴリズムを開発するにあたり、今回は機械学習を利用することにした。

機械学習的手法が衛星テレメトリデータの異常検知に有効であることは、いくつかの先行研究により示されている。まず Inui et al. (2009)<sup>[2]</sup>では、人工衛星の電力サブシステムで発生した異常事例について、いくつかの次元削減アルゴリズムおよび k-means クラスタリングアルゴリズムを用いて異常検知を行い、それぞれの手法の有効性を示した。また Kawahara et al. (2008)<sup>[9]</sup>では、宇宙機の姿勢変更マヌーバを模擬したデータにおける異常検知問題に対し、時系列データにおける異常発生時刻を、機械学習を利用した変化時刻検知の汎用的なアルゴリズムを部分空間法を用い検証し、手法の有効性を示した。しかし、機械学習においては汎用的な手法は存在しないので、それぞれの装置のドメイン知識を活用しながらそれに特化した異常検知アルゴリズムを開発する必要がある。

XRISM 衛星に搭載される *Resolve* 装置<sup>[16]</sup> は、X 線マイクロカロリメータ検出器を擁した X 線分光器である。6×6 ビ° クセルからなる X 線検出器の熱浴を極低温(約 50mK)に保持し、1つずつの X 線光子入射に伴う温度上昇の時系列データを軌道上で処理してエネルギーを決定する。そのデータの特質として、(1) 時間スケールや温度データの幅広いダイナミックレンジ、(2) 周波数やエネルギー空間で表現される検出器ノイズデータ、などがある。これらの特質を踏まえた、同装置に特化した異常検知が必要である<sup>[18]</sup>。

そのため、最初のステップとして、今回、2019 年から 2022 年にかけて行われた *Resolve* 装置の地上試験のデータを活用し、異常検知アルゴリズムを開発と検証を行った。この結果を参考にして、従来のしきい値判定に対する優劣の比較や、実際の運用に組み込む異常検知システムの要求分析と設計などを行うが、それは本稿の範囲外である。

極低温検出器を要する *Resolve* 装置に特に有用な、二

つの異常を検知する異常検知アルゴリズムの開発を行った。一つ目は、検出器のノイズスペクトルに含まれる異常である。周波数空間で表現された検出器ノイズスペクトルデータにおいて、特定の周波数帯域で現れ分光性能を劣化させる既知のノイズ源が存在する<sup>[4,5,17]</sup>。そのようなノイズが出現する時間帯を検知し、その時間帯を観測時間から除去して、観測性能の品質を保証する必要がある。二つ目は、検出器の温度データに含まれる異常である。50 mK の温度ステージは、断熱消磁冷凍機 (ADR; adiabatic demagnetization refrigerator<sup>[7]</sup>) により温度制御されている。温度安定性が検出器のゲイン安定性、すなわち分光性能に直結するので、安定度を損なう事象を検知する必要がある。ここで、ADR は約 2 日に一度 1 時間ほどの再磁化をしながら使用する。リサイクルの際は、50 mK の制御が特定のパターンをたどって外れる。これは正常な運用なので、異常とみなさない工夫が必要である。

本稿の構成は以下の通りである。まず、§ 2 で、本研究で用いた異常検知アルゴリズムを概観する。次に、§ 3 で、実データに適用させた結果を述べ、§ 4 で結論を述べる。

## 2. 異常検知アルゴリズム

異常検知アルゴリズムは、外れ値検知と変化点検知に大別される<sup>[10]</sup>。外れ値検知とは、正常なデータが従う確率分布から大きく外れたデータ点を検知することである。変化点検知とは、連続する時系列データが異常な変化をする時刻を検知することである。今回は、検出器ノイズスペクトルに含まれる異常には外れ値検知を、温度データに含まれる異常には変化点検知を用いた。以下、本稿で用いた異常検知アルゴリズムを概説する。詳細は参考文献<sup>[12,3]</sup>を参照されたい。

## 2.1. 外れ値検知

外れ値検出のアルゴリズムには、教師ありや半教師あり学習が利用される。教師あり学習とは、正常・異常のラベルがついた教師データからラベル付けのルールを学習し、未知のデータが異常であるかを推測する機械学習の手法の一種である<sup>[1]</sup>。一方、半教師あり学習では、正常データセットのみを機械学習の手法で学習し、学習した正常データのルールから外れたデータを異常と推測する。異常検知では、学習データとして正常データが圧倒的多数になったり、異常パターンが多様になったりするため、半教師あり学習が望ましいが、異常を明示的に学習させないので、判定精度は落ちる。以下、教師データを  $\mathbf{Y}_{\text{train}} = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_{N_{\text{train}}}\}$ 、テストデータを  $\mathbf{Z}_{\text{test}} = \{\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_{N_{\text{test}}}\}$  とする。

### 2.1.1. Stochastic Gradient Descent Classifier

Stochastic Gradient Descent (SGD) classifier は、確率的勾配降下法をもちいて、教師データの二つのカテゴリを類別する実数関数を求める教師あり学習の一種である<sup>[12]</sup>。ここで類別する実数関数とは、入力データ  $\mathbf{x} = \{x_1, x_2, \dots, x_n\}$  に対し

$$f(\mathbf{x}) = \omega_1 x_1 + \omega_2 x_2 + \dots + \omega_n x_n + \mathbf{b} \quad (1)$$

と表され、 $f(\mathbf{x})$  の符号によって正常と異常が分類される。

この関数のパラメータを定めるため、教師データ  $\mathbf{Y}_{\text{train}}$  に対し、以下の式が最小となるパラメータを探索する。

$$E(\boldsymbol{\omega}, \mathbf{b}) = \frac{1}{N_{\text{train}}} \sum_{i=1}^{N_{\text{train}}} L(\mathbf{y}_i, f(\mathbf{y}_i)) + \alpha R(\boldsymbol{\omega}) \quad (2)$$

ここで、 $\boldsymbol{\omega} = \{\omega_1, \omega_2, \dots, \omega_{N_{\text{train}}}\}$  である。L は教師データに対する損失項であり、教師データに対し分類を間違えるとスコアが増えていく。R( $\boldsymbol{\omega}$ ) は正則項であり、式が複雑になるほど値が増えることで、教師データに対する過学習を抑制する。

SGD Classifier は、式 (2) を最小とするパラメータ探索のため、確率的勾配降下法を用いる。確率的勾配降下法では、まず初期パラメータ  $\theta$  においての実数関数の勾配を求める。そして勾配の符号方向に、以下の式のように次の点を取りパラメータを更新する。

$$\theta := \theta - \zeta \frac{df(\theta)}{d\theta} \quad (3)$$

$\zeta$  は一定値である。これをランダムに取り出したサンプルに対して繰り返し、勾配が 0 になるパラメータを、最適なパラメータとする。

SGD Classifier を用いる際、scikit-learn ライブラリ<sup>[21,22]</sup> を用いた。scikit-learn における、SGD Classifier の主要なパラメータ<sup>1</sup> は、表 1 のとおりである。

実際のモデルの検討では、初めに線形モデルを試し、分けるデータが線形モデルで分類できるかを検証した。そして、うまくいかなかった場合は非線形モデルを試す方針を立てた。そこで一般的に使われる線形モデルの一つであり、多データでも対応できる SGD Classifier を検討した。

Table 1: SGD Classifier の主要なパラメータ

パラメータ	内容
penalty	正規化の種類
loss	損失関数の種類
alpha	正規化の強さ

### 2.1.2. クラスタリングを用いた再構成誤差法

クラスタリングを用いた再構成誤差法は、半教師学習を用いた異常検知アルゴリズムの一種である。このアルゴリズムでは、正常なデータには一つ以上の特定の形やパターンがあると仮定する。そして、正常データはパターンごとに、データ空間でいくつかのクラスター(群)を形成すると期待される。どの正常なクラスターに

<sup>1</sup>[https://scikit-learn.org/stable/modules/generated/sklearn.linear\\_model.SGDClassifier](https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.SGDClassifier).

も属さないデータは、異常なデータとみなす。

再構成誤差法では、以下のような手順により異常検知を行う<sup>[2]</sup>。まず正常な教師データ  $Y_{\text{train}}$  を  $k$  個のクラスタにクラスタリングする。このときそれぞれのクラスタ中心を  $\mu_1, \mu_2, \dots, \mu_k$  とする。任意のテストデータ  $z_i \in Z_{\text{test}}$  について、一番近いクラスタ中心が  $\mu_j$  であるとき、 $z_i$  に対する再構成ベクトル  $\hat{z}_i$  を

$$\hat{z}_i = \mu_j \quad (4)$$

とする。このとき元のベクトルと再構成ベクトルとの差として定義する再構成誤差  $E_{\text{clustering}}$  は

$$E_{\text{clustering}} = \|\hat{z}_i - z_i\|^2 \quad (5)$$

と表される。異常なデータは再構成誤差が大きくなり、これが閾値  $s$  よりも大きいとき異常とみなす。すなわち、

$$E_{\text{clustering}} > s \quad (6)$$

となるとき、テストデータ  $z_i$  を異常と判定する。

## 2.2. 変化点検知

### 2.2.1. ChangeFinder

変化点検知アルゴリズムの一種として ChangeFinder が知られている。時刻  $t$  以前の時系列データを SDAR (sequentially discounting auto regressive) モデルにあてはめ、時刻  $t$  の観測値が実際に観測される確率を算出し、スコアリングすることで変化点検知を行う<sup>[3]</sup>。

ここで SDAR モデルとは、自己回帰モデル (AR; auto regressive モデル) という時系列モデルのパラメータ推定の計算量を軽減した時系列モデルである。AR モデルでパラメータを推定するには、変化点検知アルゴリズムでは AR モデルの係数を時間ごとに変化するモデルを仮定するので、推定の計算量は多くなる。そこで SDAR アルゴリズムでは、現在の値と過去に求めたパラメータのみを用いて推定し、計算量を軽減している。

AR モデルは定常な時系列データのみに対応するのに対し、SDAR モデルは非定常なモデルにも対応できるよう推定値に対して忘却パラメータ  $r$  を設定できる。これを用いると、共分散行列  $\Sigma$  の最尤推定量を更新する際、

$$\hat{\Sigma} := (1 - r)\hat{\Sigma} + r(\mathbf{y}_t - \hat{\mathbf{y}}_t)(\mathbf{y}_t - \hat{\mathbf{y}}_t)^T \quad (7)$$

として更新し、 $r$  が 0 に近づくほど、定常な時系列モデルとなる。そして更新した共分散行列をもとに、他のパラメータも定める。

次にアルゴリズムを説明する。まず時刻  $t$  までの時系列データ  $Y_{1:t-1} = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_{t-1}\}$  に対し、SDAR モデルを構築する。この時構築された SDAR モデルから得られる確率密度関数を  $p(\mathbf{y}_t | Y_{1:t-1})$  とする。この確率密度関数に対し、時刻  $t$  のデータ  $\mathbf{y}_t$  における変化点スコア  $\text{score}_y(\mathbf{y}_t)$  を

$$\text{score}_y(\mathbf{y}_t) = -\log p(\mathbf{y}_t | Y_{1:t-1}) \quad (8)$$

とする。次にウィンドウ幅  $l$  を定め、 $z_t$  を

$$z_t = \frac{1}{l} \sum_{i=t-l+1}^t \text{score}_y(\mathbf{y}_i) \quad (9)$$

のように求める。 $\text{score}_y(\mathbf{y}_i)$  で変化点を検知すると、その変化点はその後の検知に与える影響が大きくなってしまふ。そこで  $\text{score}_y(\mathbf{y}_i)$  の移動平均系列  $Z_{1:t} = \{z_1, \dots, z_t\}$  を計算し、再度その SDAR モデルを考え変化点検知を行うことにする。よって、この操作で得られる確率密度関数  $q(z_t | Z_{1:t-1})$  を用いて、時刻  $t$  における変化点スコア  $\text{score}_z(t)$  を

$$\text{score}_z(t) = -\log q(z_t | Z_{1:t-1}) \quad (10)$$

と計算する。最後に時刻  $t$  における変化点スコアに対し、閾値  $s$  を設定し、

$$\text{score}_z(t) > s \quad (11)$$

となる時刻  $t$  を、変化点として検出する。

微分値に閾値を設定することで変化時刻を検出する



といった方法は、温度変化量の大きさに score が依存してしまうという問題がある。一方、ChangeFinder の score は温度の変化量の大きさの影響を強く受けないため、温度変化点検知に有力である。

### 3. 異常検知アルゴリズムの開発

#### 3.1. 検出器ノイズスペクトルの異常検知

##### 3.1.1. 現象の背景

Resolve 装置で X 線を検出する際、約 500 Hz 以下のノイズが X 線イベントのエネルギー分解能に影響する。その影響は、周波数が低いほど大きい。この周波数帯域にパワーを持つノイズ源を監視し、その異常を検知する必要がある。地上試験を通していくつかのノイズ源が特定されたが、ここでは検出器外部からの電磁・微小擾乱干渉に起因する以下の 2 つのノイズ（それぞれ MTQ ノイズ、beat ノイズと呼ぶ）を対象とする。

(1) 磁気トルカ (MTQ; magnetic torquer) による電磁干渉ノイズ。MTQ は XRISM 衛星に搭載される姿勢制御系のサブシステムの一つである。MTQ はコイルに電流をかけて地球磁場と作用させることで、XRISM 衛星に蓄積された角運動量を散逸する。この時、127 Hz のパルス幅変調で磁場強度を調節するので、127 Hz とその高調波でのピークがノイズスペクトルに生じる (Fig. 1)。また、検出器アレイを構成する 36 個のピクセルのうち、0, 9, 18, 27 番ピクセルにより強い線ノイズが立つことが知られている。これらのピクセルからのアナログ信号線がバンドルの端にあることから、外的影響を受けやすいためと考えられている [4]。

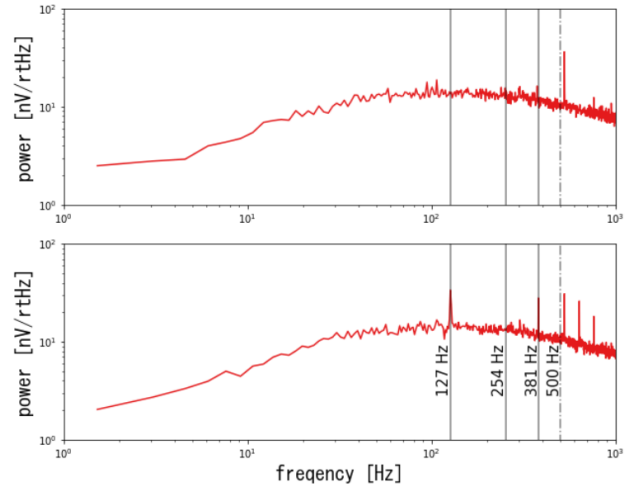


Fig. 1: 2022/06/09 に取得した MTQ ノイズなし (上)、2022/06/09 に取得した MTQ ノイズあり (下) の 8k noise spec の一例

(2) 機械式冷凍機による微小擾乱干渉ノイズ。Resolve 装置には検出器を冷却するため、複数の冷凍機が搭載されている。そのうちの 2 種類が、Stirling 冷凍機 (4 台) と Joule-Thomson 冷凍機 (1 台) という機械式冷凍機である [19,20]。Stirling 冷凍機の駆動周波数は  $f_1 = 14.77 - 15.16$  Hz, Joule-Thomson 冷凍機は  $f_2 = 51.27 - 53.35$  Hz の間で調整できる設計になっている。ある  $f_1, f_2$  を選択した時、二つの波の重ね合わせにより、

$$\begin{aligned} & \cos(2\pi n f_1 t) + \cos(2\pi m f_2 t) \\ &= 2 \cos\left(2\pi \frac{n f_1 - m f_2}{2} t\right) \cos\left(2\pi \frac{n f_1 + m f_2}{2} t\right) \end{aligned} \quad (12)$$

という波が生じる。すなわち、時空間において、振幅が周波数  $\left|\frac{n f_1 - m f_2}{2}\right|$  で変動する波 (うなり) が生じる。通常、フーリエ空間ではこのうなり周波数は現れない (単に  $n f_1$  と  $m f_2$  に分解されるだけである) が、振幅に対して非線形応答があると、うなり周波数  $\left|\frac{n f_1 - m f_2}{2}\right|$  にもパワーが分配される。実際、Resolve 装置では、 $(n, m) = (24, 7), (15, 4)$  のときの beat ノイズが実際に観測されている (Fig. 2) [5]。この時、うなりの周波数が  $f_1$  や  $f_2$  よりも低周波数帯域にくるので、特に重要である。

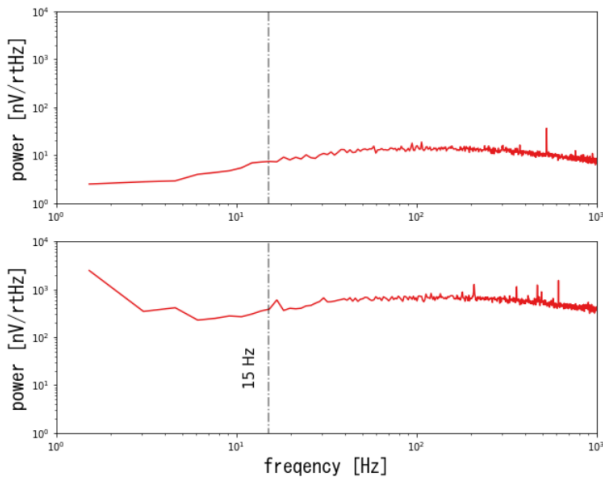


Fig. 2: 2022/06/09 に取得した Beat ノイズが発生していない時刻の 8k noise spec の一例 (上)、2019/12/11 に取得した Beat ノイズが発生している時刻の 8k noise spec の一例 (下)

### 3.1.2. 開発項目

Resolve 装置の運用では、1日に1回程度で、ノイズスペクトル (8k noise spec) が全ピクセルに対して取得される。これらから、上述の2つのノイズが発生しているかを検知するアルゴリズムを開発する。

異常検知に際し重要となる指標として、実際に異常であったもののうちどれだけ異常と予測できたかの割合を示す「真陽性率」、実際は正常であったのに異常と予測された「偽陽性率」がある<sup>[13]</sup>。今回は、検討したモデルで真陽性率の目標値を0.9と設定し、これを超える検知を行うことができるのか、そして真陽性率0.9を超える検知を行う場合、偽陽性率はどの程度の値になるかを検討した。

学習データとして、Resolve 装置の地上実験で、2021/05/06 から 2022/06/17 までの期間 (衛星初期電気試験) で得られたデータを用いた。8k noise spec 一つのサンプルは、36ピクセル×4096周波数ビン (1.5 Hz から 6.25 kHz まで 1.5 Hz の周波数分解能) のデータである。

### 3.1.3. 異常検知

**MTQ ノイズ** このノイズが強く現れる周波数を含む 10–400 Hz の帯域を、線ノイズが強くあらわれる4つのピクセルのみに対して、学習及びテストをする。このノイズは、すべてのピクセルに同時に出現するので、ピクセルを絞ってもノイズ発生時間帯を特定するには問題ない。

8k noise spec のデータとして、MTQ が確実に駆動している時刻の 82 サンプル、それ以外の時刻の 11597 サンプルを用意した。ここからランダム抽出によりテストデータを構成した。テストデータは、MTQ ノイズが立っていないデータ 115 サンプル、MTQ を駆動時に 127 Hz とその倍波でピークが立っている 25 サンプルからなる。それ以外のサンプルは、教師データとパラメータ推定用データとした。これらのデータでは、MTQ を駆動している時間帯は既知なので、その時に取得された 8k noise spec を異常、それ以外で取得されたものを正常とラベル付けをした。ただし、試験中に MTQ を駆動した時間が少なく、異常データのサンプル数が少なかったため、モデルの最適パラメータ探索の精度や学習精度が低下した。ゆえに今回は、SMOTE (Synthetic Minority Oversampling Technique)<sup>[6]</sup> をもちいて異常サンプルを増やし、正常サンプル数と均一な数にした。これにより、MTQ ノイズサンプルは正常サンプルの同数のサンプル数に増やした。ただし、SMOTE 等を用いたオーバーサンプリングを行うことで、元データの相関を損なう可能性や、過学習を引き起こす可能性が考えられることには、注意が必要である。

過学習の影響を減らすため、グリッドサーチを用いて、パラメータチューニングを行った。それぞれのデータを標準化させたのち、SGD Classifier のパラメータを決めるため、パラメータ検証用データに対し、グリッドサーチを行った。その結果、SGD Classifier のパラメータのうち、penalty は L1、alpha を  $1 \times 10^{-7}$ 、loss を squared\_hinge とした。

以上のように構築したアルゴリズムに、検証用データを適用した。異常データ 25 サンプルと正常データ 115 サンプルを用いてテストした結果を混合行列として表 2 に示す。真陽性率は 1.0、偽陽性率は 0.04 となり、0.9 以上の真陽性率での検知が可能であることが示された。加えて、偽陽性率も十分に小さく、高い確度で異常を検知できることが示された。

Table 2: SGD Classifier での MTQ ノイズ検知の混合行列

		判断された結果	
		異常	正常
実際の結果	異常	25	0
	正常	5	110

**Beat ノイズ** 今回は、対象ノイズが強く現れる 15 Hz 以下を学習・テストする周波数帯とした。Beat ノイズは多種のパターンを持つ。そこで再構成誤差法<sup>[2]</sup>を利用することにした。この手法は、半教師あり学習を用いた異常検知アルゴリズムの一つであり、beat ノイズのように表現の種類が多様な異常サンプルも判別しやすいと期待される。再構成誤差法には、次元削除もしくはクラスタリングを用いたアルゴリズムがある。複雑な教師データの場合、後者のほうが精度が良い結果が得られることが先行研究<sup>[2]</sup>より明らかになっているので、今回はクラスタリングを用いた再構成誤差法を用いた。クラスタリングのアルゴリズムとしては k-means を用いる。

今回用いる 8k noise spec の全サンプルは、beat ノイズが発生している時刻の 1356 サンプル、それ以外の時刻の 11593 サンプルである。ここからランダム抽出でテストデータを構成した。テストデータは、beat ノイズが発生している 116 サンプル、発生していないサンプル 678 サンプルである。残りを、教師データとパラメータ（クラスタ数）推定用データとした。これらのデータでは、beat ノイズが発生している時間帯は既知なので、その時に取得された 8k noise spec を異常、それ以外で取得されたものを正常とラベル付けした。

まず教師用データに対して標準化をした。そして、そ

の標準化をクラスタ数探索用、テスト用に対して用い、標準化を行った。その後、クラスタ数推定用データに対しクラスタリングを行い、エルボー法によって最適なクラスタ数を探索した。その結果、クラスタ数を 2 に設定して 2 値分類を行った。

結果、AUC (area under curve) 値は 0.92 となった。ここで AUC 値は ROC (receiver operating characteristic) 曲線の下部の面積であり、1 に近いほど誤判定が少なく検出率が高い良い検出器であることを示す指標である。再構成誤差に対する閾値を真陽性率が 0.9 となる点として定めるとき、偽陽性率 0.35 となった。これは改善すべき課題である。

## 3.2. 温度データの異常検知

### 3.2.1. 現象の説明

Resolve 装置の検出器が搭載された極低温ステージの温度制御をする ADR<sup>[7]</sup> が正常な作動をすることは、Resolve 装置の高エネルギー分解能保持に必要な不可欠である。ADR はエントロピー一定で消磁をすることで温度が下がる断熱消磁を利用しており、ある程度消磁が続くと、冷却を継続するため再磁化をする必要がある。この再磁化を ADR recycle という。ADR recycle が行われるときは、検出器ステージの温度が Fig. 3 のようなパターンで変化することが観測されている。

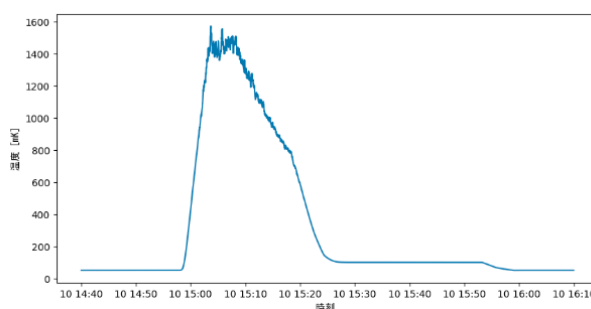


Fig. 3: ADR recycle 中の検出器ステージの温度変化の一例

Resolve 装置の検出器ステージの温度は、いくつかの時刻で温度変化していることが観測され、中には ADR

recycle での温度変化パターン以外の異常パターンでの温度変化が見られた<sup>[1]</sup> また ADR recycle 自体も、磁化電流を流す超伝導線がクエンチするなどして異常が発生する可能性がある。

### 3.2.2. 開発項目

Resolve 装置の検出器ステージの温度が正常な ADR recycle 以外のパターンでの温度変化をする時刻もしくは、ADR recycle が失敗した時刻を検知する。このとき必要となるのが、以下のようなアルゴリズムである。

- (1) 検出器ステージの温度が時間変化をした時刻を検知する検出器
- (2) (1) で検知した時刻以降の温度変化が異常であるかを検知する分類器

今回は、そのうち (1) の検出器ステージが温度変化をした時刻を検知する検出器の作成のみの開発を行った。

### 3.2.3. 開発と性能

**用いたデータセット** 今回用いたデータセットは、Resolve 装置の地上実験（2022/1/24 から 2022/06/17 までの期間）で得られた HK データのうち、検出器ステージの 50 mK モニター用温度データを用いた。

与えられた時系列データは、1 s ごとにサンプルされた連続データであるが、欠損値も散見された。欠損区間は、null 扱いとした。全サンプル数は 12446251 s 分である。変化点検知の際に用いるパラメータの調整は難しいため、すべての区間を変化点検知の対象とすると、検知精度が低下する。そこで変化点検知にかける前に、検知したい区間をある程度特定し、変化点の検出の精度を向上させる必要があった。そこで、以下の手順を踏み、変化点検知を行う区間を絞った。

- (1) 12446251 秒間の時系列温度データを  $1800 \text{ s} \times 6914$  に区切り、これを一区間とした。ADR recycle にかかる時間が約 60–90 分であることを鑑みた。
- (2) 一区間の温度データのうち、温度が 52 mK 以上となる時刻が一つでもある区間を抽出した。温度計グリッチなどによる影響を除くためである。
- (3) (2) で抽出した区間のうち、ADR recycle 中であることを示す 500 mK 以上となる区間を除いた。
- (4) (3) で抽出した区間の前後の区間をつなげた 7200 s を新たな区間とした。この区間に対して、k-means を用いてクラスタリングを行い、2 クラスに分類した。そのうちの欠損値が含まれるデータが分類されるクラスタ内のデータを除いた。

データを排除するために実行した。Step 2 と Step 3 では想定される対象外のデータを取り除いたが、Step 4 では二値分類を行うことで、50 mK に保たれているデータ内での異常な温度変化や、ADR Recycle による変化以外のものを検出することを目的とした。Step 4 の結果、欠損値が含まれるデータが含まれるクラスタと、50 mK に保たれているデータに分類された。したがって、欠損値を含むクラスタのデータは対象外とし、除外した。

**ChangeFinder を用いた検出器ステージの温度変化点検知** § 3.2.3 で分割した区間の温度データに対し、ローパスフィルタをかけてから ChangeFinder アルゴリズムを適用した。

変化点検知の際、微分値に対して閾値を設定し変化点を検知する方法が主流だが、この方法では閾値以下の変化点を検知できないことに加え、変化点以降の変化の傾きが大きい時刻に高く微分値が出力される。それに対し ChangeFinder では、変化点のみを高く出力されるので、今回のような変化開始時刻の検知に適している。



次にローパスフィルタを用いた理由として、一部 ADR recycle が生じる前後で温度の微小な変化が多く、ADR recycle 開始時の変化が的確に検知されなかった。ゆえに変化点検知を正確に行うため、カットオフ周波数を 0.05 Hz としてローパスフィルタをかけた。加えて欠損値の温度データの前後 10 s の変化点スコアに 0 をかけ、補間したデータに対し変化点検知を行わないようにした。カットオフ周波数は、いくつかのデータを用いて最適値を決めた。ADR リサイクル以外の変化点、特に変化温度が 10 mK 程度の異常な温度変化を検出できることを重視した。

ChangeFinder のパラメータ (§ 2.2.1) は、 $r$  を 0.05、SDAR モデルのオーダーを 1、 $L$  を 1.5 とし、変化点検知を行った。Fig. 4 が ADR リサイクル時の算出された変化点スコア、Fig. 5 が温度揺らぎ時の算出された変化点スコアである。算出された変化点スコアに対し、スコアが 20 以上になる時刻を検出した結果、143 区間で変化点が 660 個検出された。しかしこの時、Fig. 5 のような温度揺らぎのところで、変化点検知が過剰に反応してしまった。この条件では、単なる温度の揺らぎまで変化点として検出されている。そのため一段階目で検知した変化点が、ADR による温度変化や異常な温度変化であるか、単なる温度揺らぎであるかを区別するため、二段階の変化点検知を実行した。

まず ADR や異常な温度変化は、一定の状態から急激に温度が変わるものである。対照的に、温度揺らぎは連続的に発生する。このため、一段階で検出された変化のタイムスタンプが本当に ADR や異常な温度変化であるのかを確かめるため、二段階目では、検知した時刻に対し前 1000s、後ろ 100s を取り、合計 1100s の区間で再度変化点検知を行った。この時は ChangeFinder のパラメータを、 $r$  を 0.01、 $L$  を 7.5 とした。これにより温度揺らぎでの部分は定常性を持つので、検知点スコアが Fig. 4 のように相対的に小さくなる傾向が見れた。つまり、最初の段階で取り出した区間が適切でなく、変化点と判断する範囲がごくわずかな時間内に制限されてしまい、定常状態を正しく把握できず、温度揺らぎに誤検出が発生していたものが、二段階目の処理によっ

てしまい、定常状態を正しく把握できず、温度揺らぎに誤検出が発生していたものが、二段階目の処理によって正確に判定できるようになった。これにより閾値を 50 に設定することで、温度揺らぎによる変化点検知の数が削減でき、見た目で判断したものを抽出できる結果になった。

この時得られた変化点時刻は 289 点となった。検知した点を確認すると、全体の約 1/4 は ADR recycle による温度変化時刻や異常な温度変化時刻であり、本来目的とした時刻を検知していた。しかし一方で、残りは温度揺らぎやグリッチによる変化点等本来検知する必要のない温度変化時刻を検知していた。これらを取り除くためには、さらにヒューリスティックな手法を重ねる必要があり、ChangeFinder のみによる温度変化検知では、今回目的とする変化時刻検知は難しいと判断された。

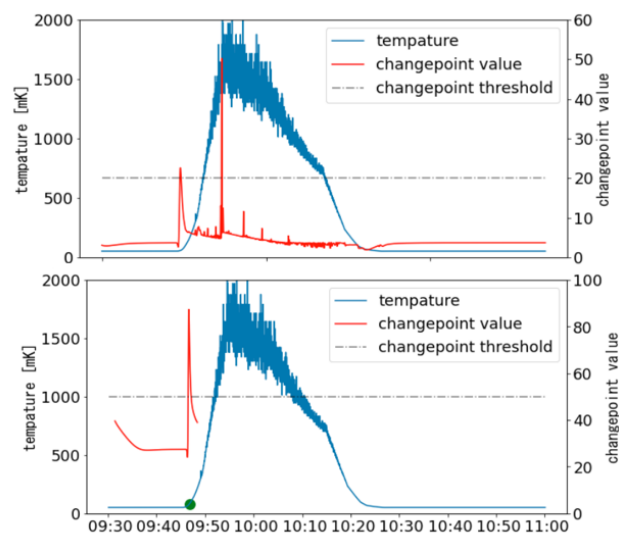


Fig. 4: 2022/1/29 での ADR Recycle 時の温度変化における、一段階目の変化点スコア (上) と二段階目の変化点スコア (下)。下図の緑の点が、二段階目で変化点検知をしている時刻である。

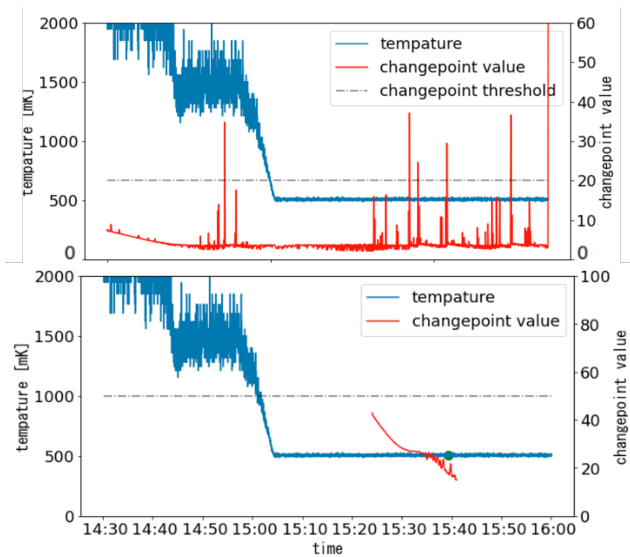


Fig. 5: 2022/2/5 での温度揺らぎ時の温度変化における、一段階目の変化点スコア（上）と二段階目の変化点スコア（下）。下図の緑の点が、二段階目に変化点検知をしている時刻である。

#### 4. 結果・考察

本稿では、XRISM 衛星に搭載されたミッション機器 *Resolve* のデータに対する異常検知アルゴリズムの開発について紹介した。地上試験データを用いて機械学習的な手法を用い、*Resolve* 装置のテレメトリデータに現れる二種の異常—検出器ノイズスペクトルに含まれる異常と、検出器の温度データに含まれる異常—を対象とした。

まず異常ノイズスペクトル検知について、今回は MTQ ノイズと beat ノイズという既知の異常を検出対象とした。MTQ ノイズに関しては、教師あり学習にもとづき SGD Classifier を用いて検出を行った結果、真陽性率 1.0、偽陽性率 0.04 での検出ができた。Beat ノイズに対しては、半教師あり学習にもとづき、クラスタリングを用いた再構成誤差法を用いて検出を行った結果、真陽性率 0.9、偽陽性率 0.37 での検出ができることが示された。MTQ ノイズに関しては、機械学習で一般的に用いられるモデルを使用した場合の結果であるが、実用化を十分目指せるレベルの結果であることが示された。Beat ノイズに関しては、適切な閾値を設定す

ることで、十分な真陽性率での検知ができることが示されたが、同時に偽陽性率が高くなり、誤検知が多くなることが分かった。ゆえに、先行研究<sup>[2]</sup>で検討された VQPCA のような、通常用いられるモデルよりも複雑なモデルを用いた再構成誤差法を検討し、偽陽性率を低下させることが必要と考えられる。

次に異常温度変化時刻検知について。検出器面の温度データを約 5ヶ月分処理し、289 点の温度変化を検知した。このときの検知した変化点には、正常な ADR recycle による温度変化とそれ以外の温度変化が混在していた。全体のうち 1/4 は ADR recycle による温度変化や異常な変化時刻（検出対象とするもの）であった。しかし残りは、温度揺らぎやグリッチによる温度変化であった。更に、アルゴリズムとして、ChangeFinder のパラメータを手動で調節する、適切なローパスフィルタを施すパラメータの異なる処理を 2 段階でかけるなど、ヒューリスティックに最適値を見つけていく過程が必要であった。異常ノイズスペクトルの検出で用いた古典的な機械学習手法とは異なり、異常温度変化時刻検知では時系列解析手法を用いたが、これを実用化するためには、ChangeFinder 以外の、別の変化点検知アルゴリズムも検討する必要があると考えられる。

#### 謝 辞

立教大学 山田真也氏、宇宙科学研究所 金丸善朗氏、匿名査読者に原稿へのコメントを頂いた。本研究を進めるうえでお世話になった、宇宙科学研究所 海老沢研教授、横浜国立大学 中村正吾准教授、及び XRISM *Resolve* チームメンバーの皆様に謝意を表す。

#### References

- [1] 塚本 邦尊, 山田 典一, 大澤 文孝, 中山 浩太郎, 松尾 豊. 東京大学のデータサイエンティスト育成講座: Python で手を動かして学ぶデータ分析. マイナビ出版. 2019. p198
- [2] 乾 稔, 矢入 健久, 河原吉伸, 町田和雄. 次元削減の再構成誤差を用いた異常検知手法の比較, JSAI, Vol.JSAI2009, 2009. pp. 1B12 - p. 1B12

- [3] 島田 直希. 時系列解析: 自己回帰型モデル・状態空間モデル・異常検知, Timeseriesanalysis. 共立出版. 2019. p. 155- p. 157
- [4] KURIHARA Miki, TSUJIMOTO Masahiro, ECKART Megan, et.al. Ground test results of the electromagnetic interference in the x-ray microcalorimeter onboard XRISM. In Space Telescopes and Instrumentation 2022: Ultraviolet to Gamma Ray, SPIE, Vol. 12181, 2022, pp. 1445–1458.
- [5] IMAMURA Ryuta, TSUJIMOTO Masahiro, AWAKI Hisamitsu, et. al., In X-Ray, Optical, and Infrared Detectors for Astronomy X, SPIE, Vol. 12191, 2022, pp. 763–770.
- [6] CHAWLA V.Nitesh, BOWYER W. Kevin, HALL O. Lawrence, PHILIP W. Kegelmeyer SMOTE: Synthetic minority oversampling technique. n Journal of Artificial Intelligence Research. Vol. 16, 2002, pp. 321–357.
- [7] SHIRRON J. Peter, KIMBALL O. Mark, JAMES L. Bryan, et.al., Space Telescopes and Instrumentation 2016: Ultraviolet to Gamma Ray, SPIE, Vol. 9905, 2016, p. 99053O.
- [8] 宇宙開発利用部会 X 線天文衛星「ひとみ」の異常事象に関する小委員会, X 線天文衛星 ASTRO-H 「ひとみ」異常事象調査報告書, 国立研究開発法人 宇宙航空開発研究機構, 2016, p. 22 - p. 49
- [9] KAWAHARA Yoshinobu et al. Change-Point Detection Algorithms based on Subspace Methods. Transactions of the Japanese Society for Artificial Intelligence. Vol.23 . p. 76 - p. 85
- [10] 井手 剛, 杉山 将, 異常検知と変化検知 = Anomaly detection and change detection, 講談社, 2015, p. 3
- [11] HASEBE Takashi, IMAMURA Ryuta, TSUJIMOTO Masahiro , et.al., Ground test results of the micro-vibration interference for the x-ray microcalorimeter onboard XRISM. spie. Vol. 1218. 2022. p. 1527 - p. 1539
- [12] BOISBERRANGER du Jérémie, BOSSCHE Van den Joris, et.al., scikit-learn. 2007. <https://scikit-learn.org/stable/index.html> 2023.4.17 参照
- [13] 日本統計学会. 統計学実践ワークブック. 学術図書出版社. 2020. p. 210
- [14] TAKAHASHI Tadayuki, KOKUBUN Motohide , MITSUDA Kazuhisa, et. al. The ASTRO-H (hitomi) x-ray astronomy satellite. SPIE. Vol. 9905. 2016. p. 99050U
- [15] TASHIRO Makoto, MAEJIMA Hironori, TODA Kenichi, et al. Status of x-ray imaging and spectroscopy mission XRISM. SPIE. Vol 11444. 2020. p. 176
- [16] ISHISAKI Yoshitaka, KELLEY Richard L., AWAKI Hisamitsu et.al., Status of resolve instrument onboard x-ray imaging and spectroscopy mission XRISM. SPIE. Vol 12181. 2022-9. p. 12181
- [17] HASEBE Takashi ,IMAMURA Ryuta , TSUJIMOTO Masahiro et.al, Ground test results of the micro vibration interference for the x-ray microcalorimeter onboard XRISM. SPIE. Vol.12181. 2022-8 p. 1527 - p. 1539
- [18] TSUJIMOTO Masahiro. Development of the data monitor system of the Resolve instrument onboard the XRISM satellite. JSIJ. Vol.12. 2023-2. p.1 - p. 6
- [19] ISHISAKI Yoshitaka ,KELLEY Richard L.,AWAKI Hisamitsu et.al., Status of resolve instrument onboard x-ray imaging and spectroscopy mission XRISM. SPIE, Vol 12181. 2022-9. p. 409 - p. 430
- [20] KANAO Ken'ichi , YOSHIDA Seiji, MIYAOKA Mikio et.al, Cryogen free cooling of ASTRO-H SXS helium dewar from 300 K to 4 K. Cryogenics. Vol. 88. 2017-12. p. 143 - p. 146
- [21] FABIAN Pedregosa. GAEL Varoquaux et.al. Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research. Vol.12 2011. p.2825 - p. 2830
- [22] LARS Buitinck, GILLES Louppe, MATHIEU Blondel et.al., API design for machine learning software: experiences from the scikit-learn project. arXiv preprint arXiv:1309.0238 2013 p. 108 - p. 122