

SACLA におけるグリッド電圧波形データを用いた サイラトロン の異常検知ソフトウェアの開発

DEVELOPMENT OF SOFTWARE TO DETECT ANOMALIES OF THYRATRONS USING GRID VOLTAGE WAVEFORM DATA IN SACLA

佐藤伸行^{#, A)}, 岩井瑛人^{B), C)}, 前坂比呂和^{C), B)}, 安留健嗣^{C)}, 近藤力^{B), C)}, 稲垣隆宏^{C), B)}

Nobuyuki Sato^{#, A)}, Eito Iwai^{B), C)}, Hirokazu Maesaka^{C), B)}, Kenji Yasutome^{C)}, Chikara Kondo^{B), C)}, Takahiro Inagaki^{C), B)}

A) CHUO ELECTRONICS CO., LTD.

B) Japan Synchrotron Radiation Research Institute (JASRI)

C) RIKEN SPring-8 Center

Abstract

We have developed a failure prognosis algorithm of thyratrons, utilized as high-power switches for pulsed high-power RF sources in the XFEL facility, SACLA, by using a machine-learning-based waveform analysis. Since the lifespan of a thyatron is relatively short and individually different, the thyatron replacement often arises when the thyatron deteriorates or fails due to the limited product lifespan. For the stable operation of SACLA, therefore, it is important to detect signs of thyatron failure and to prevent a sudden failure of the thyatron. We investigated the possibility of estimating the degree of thyatron degradation by applying a machine-learning method to voltage waveforms of the control grid of a thyatron. Since some features of the waveform were found to be sensitive to the age of a thyatron, we developed software to estimate the degree of deterioration of thyratrons. The developed software can distinguish the new thyatron and the failed one almost perfectly and estimate the age of a thyatron with sufficient accuracy.

1. はじめに

X線自由電子レーザー施設 SACLA では、パルス大電力高周波源の大電力スイッチとしてサイラトロンが多数使用されている。サイラトロンは寿命が数万時間と比較的短く個体差が大きい。また高価であり、原価高騰と為替変動によりここ数年で購入費用が倍になっている。寿命によりサイラトロンが故障すると、電子ビームの品質や加速器の稼働率に影響する[1]。トラブルを未然に防ぐため、計画的に交換をしているが、一方で費用低減のため可能な限り長く使用することが望ましい。

SACLA 線型加速器の上流部分では強いバンチ圧縮を行うため、最も品質の良いサイラトロンが要求される。そこで、新品のサイラトロンはまず上流部分で使用され、ユーザ利用の合間の調整時間に早ければ 2,000 時間程度でこれを交換し、下流に再利用する。交換を行う上では、稼働時間、自爆回数、出力電圧(クライストロン電圧)のタイミングジッタなどのデータを参考にすほか、制御グリッドの電圧波形に異常があるかをひとつずつ確認して、交換する個体を決めている。

より正確にサイラトロン の余命を予測できれば、定常的にモニターし続けることによって、故障の直前まで使用した上で計画的に交換することができ、施設の稼働率向上や費用低減に繋がる。自爆回数については機械学習により故障予兆を推定する取り組み[2, 3]が行われているが、自爆回数の増加以外の故障モードが存在する、偽陽性が多いなどの課題がある。我々はより精度の高い予測を行うため、グリッド電圧波形に着目した。

SACLA で使用しているサイラトロンには 3 段の制御グ

リッド(G1, G2, G3)があり、G1, G2 の順にトリガパルスを加することで導通する。メンテナンス時には主に G1 および G2 の電圧波形をオシロスコープで測定し観察する。サイラトロンが劣化すると、グリッド電圧波形に、1) 残留電圧変動、2) 立ち上がり最大値の増加と立ち下りの発生(G1)、3) サージ電圧の増加、4) 着火ディレイおよびジッタの増加、といった変化が発生することが分かっている[4]。我々はこの変化からサイラトロン の劣化度を推定できないかと考えた。

グリッド電圧波形のデータは故障や定期交換のたびに測定し蓄積されている。このデータと機械学習の手法を組み合わせることでサイラトロン の劣化度を推定することができるか検討を行った。機械学習の出力からは、サージ電圧の増加と劣化度に関連性が高いことが示唆された。これはサイラトロン の動作メカニズムからも妥当なものである。この結果を元に、サイラトロン の劣化度を推定するとともに、推定稼働時間を併せて出力するソフトウェアを開発した。

本報告では、これらの取り組みに関して、機械学習による分析、開発したソフトウェア、および、評価結果について述べる。

2. 使用するデータ

2.1 対象のデータ

本取り組みでは、サイラトロン の入力部である G1 と G2 の電圧波形(以降、それぞれ G1 波形、G2 波形と呼ぶ)を使用した。また、出力であるクライストロン電圧の波形(以降、Vk 波形と呼ぶ)も併せて使用し、同じ枠組みで検討することにした。これは、グリッド電圧波形にディレイ・ジッタ等が発生する時、同様に Vk 波形にもディレイ・ジッタ等が現れるためである。

[#] nobusato@ccc.co.jp

G1、G2、Vk 波形の例を Fig. 1 に示す。これは新品のサイラトロンにトリガパルスを印加した時の波形で、G1、G2、Vk の順で特徴的な波形が出ていることが分かる。

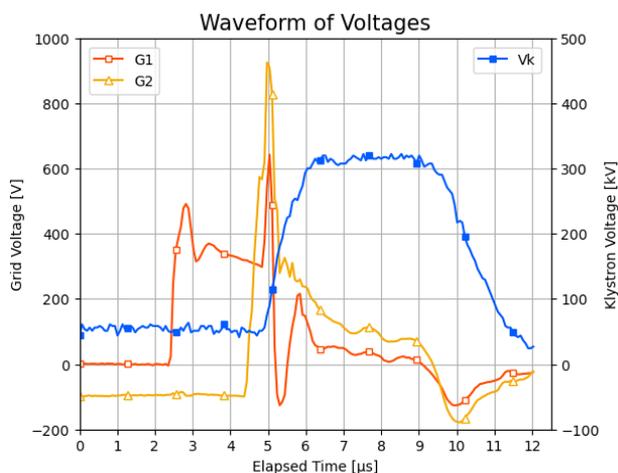


Figure 1: Example of waveforms of thyatron. The left axis corresponds to G1 and G2, and the right axis to Vk.

2.2 データの測定

データの測定はテストスタンドという試験設備で行う。G1、G2 の順にトリガパルスを 1 度ずつ印加することを 1 ショットと呼ぶ。テストスタンドでは、連続で 100 ショットの入力を行い、その時の種々の波形を A/D コンバータを介して測定し、記録する。

データの測定はサイラトロンの故障や定期交換時に行われる。これらのデータは個体のシリアル番号、設置場所、稼働時間、取り外し理由、故障状態などの情報とともに保管されるため、データとその属性が紐づけられており機械学習をする上で便利なデータである。

2.3 データの形式

G1、G2、Vk 波形の各ショットのデータの形式は下記となっている。

- サンプリング周波数 14.875 MHz (周期約 67.2 ns)。
- 2,048 サンプル (約 137.7 μ s) の電圧値。
- G1、G2 は実値の 1 / 1,500 (単位: V)。
- Vk は実値の 1 / 5,000 (単位: kV)。
- 小数点以下 6 桁。
- テキストファイルに 1 ショットあたり 3 ファイル (波形種類ごとに 1 ファイル)、1 行当たり 1 データで記録。
- 電圧値の正負が反転している。

2.4 データの選別

サイラトロンの個体ごとに 100 ショットの波形を取得しているが、そのうち正常に測定できなかったデータ (ジッタ、テストスタンド由来のノイズ等) は除いた (平均 10 % 程度)。なお、除去はルールベースの処理により行った。

2.5 データの加工

各波形は、サイラトンへのパルス入力開始時から一定時間に渡りサンプリングされており、振幅の大きい特徴的な波形は先頭に近い部分に現れる。よって、機械学習で用いる前にデータを切り出す。

検討の結果、G1 波形については 21~180 サンプル

(160 個の値、 $\sim 10.8 \mu$ s)、G2、Vk 波形については 41~180 サンプル (140 個の値、 $\sim 9.4 \mu$ s) を用いることとした。先頭の 1~20 サンプルないし 1~40 サンプルを用いていないのは、パルス入力前の時間帯であるためと、測定環境由来のノイズを取り除くためである。

また、前述のように測定環境の都合で電圧値の正負が反転しているため、データを再度反転させて使用した。

2.6 サイラトン個体のラベル付け

機械学習による分類を行う上で、学習に使用するデータにラベル付けを行う必要がある。本取り組みでは、“正常”と“故障”の 2 値分類を行うこととし、ラベル付けを行った。ラベル付けの方針としては、目視によりサージが顕著な個体を“故障”、それ以外を“正常”とした。“故障”に分類した個体は現行の運用でも故障間近あるいは故障した個体である。

2.7 データの追加

本取り組み開始時に測定済みのデータのみでは、稼働時間の偏りが大きかったため、故障などにより保管してあるサイラトンについて新たに波形の測定を行った。これは、より多様なデータがあるほど機械学習の精度向上が見込まれるためである。また、“正常”と“故障”がほぼ半数となるように、新品のサイラトンについても追加で波形の測定を行った。

2.8 自爆回数が多い個体の除外

本取り組みの当初は、データを取得済みのすべての個体の波形を使用して検討していたが、途中から自爆回数が多い個体の一部を除外することとした。これは、自爆回数が多い個体とその他の個体のデータを見比べた結果、自爆回数が多いけれども波形が新品と同様の個体があり、今回の方法では効果的に異常検知できないと判断したためである。自爆回数については既存の取り組み [2, 3] でケアすることとした。

2.9 データの個数

以上の条件で選別したデータを Table 1 に集計した。

個体数は計 36 個となり、うち、“正常”が 19 個、“故障”が 17 個となった。稼働時間を 5,000 時間ごとに区切ると

Table 1: Number of Thyratrons per Classification and Operating Hours

Operating hours	Classifications		
	Normal	Faulty	Total
0~5,000	13	0	13
5,000~10,000	4	0	4
10,000~15,000	2	1	3
15,000~20,000	0	5	5
20,000~25,000	0	7	7
25,000~30,000	0	1	1
30,000~35,000	0	3	3
Total	19	17	36

それぞれの区分けに 1 個以上の個体が含まれている。

2.10 データの特徴

典型的な G1 波形データを Fig. 2(a), (b) に示す。(a) は新品 (0 時間使用) の波形データ、(b) は劣化品 (13,700 時間使用) の波形データである。いずれも、約 100 ショット分のデータをショットごとに色分けし、重畳した図となっている。G2、Vk 波形は紙面の都合で省略する。

新品の G1 波形はほぼぴったり重なっており、ショットごとに揺らぎが発生していない。G2 波形も同様で、Vk 波形については多少の揺らぎが発生している。

劣化品の G1 波形には主に 6.5~9.5 μs 付近にショットごとの揺らぎが発生しており、これが残留電圧変動と呼ばれるものである。G2 波形も同様である。また立ち上がり電圧の増加や立ち下りの発生、大きなサージ電圧が発生しているのが読み取れる。さらに新品と劣化品でピークのタイミングを比較すると、特に G2、Vk 波形で後ろにずれており、ディレイ・ジッタが発生している。

劣化度が進むとそれぞれの変動がさらに大きくなる。

3. 機械学習の手法を用いたデータ分析

機械学習の手法を用いてデータの次元削減と分類、推定稼働時間の計算を行った。処理系として Python 言語のライブラリの一つである scikit-learn を用いた。

まず、主成分分析による次元削減を行い、その結果をロジスティック回帰やサポートベクターマシンによる分類、および、線形回帰による推定稼働時間の計算に用いた。

3.1 主成分分析

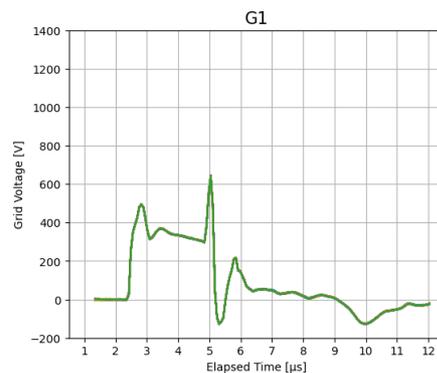
主成分分析[5]とは、相関のある多数の値から相関のない(少数の)値に変換する操作である。主成分分析により変換された後の値を主成分と呼ぶ。主成分を求めるための係数は、元になるデータの主な特徴を表している。なお、主成分分析では一般に前処理としてデータの標準化が行われることが多いが、本取り組みでは波形データの情報が失われるため行わない。

主成分分析は 2 段階で行われる。第 1 に、学習データから係数を求める。第 2 に、求めた係数を使用して学習データを含むデータから主成分を求める。全波形データのうち、“正常”と“故障”が均等になるように分けた半分を学習データとした。

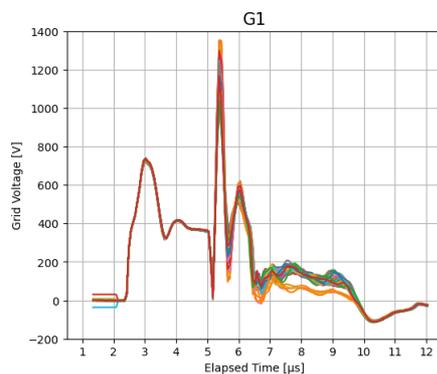
学習データに対する主成分分析として、G1、G2、Vk 波形それぞれについて、全個体・全ショットのデータを一度に入力して係数を求めた。結果の一部を Fig. 2(c) に示す。これは主成分分析の結果 (G1 波形) のうち第 1 主成分の係数ベクトルであり、元の波形のどこに着目したかを表す。5.3 μs 付近にピークがあり、ジッタの影響を捉えている。また、3.1 μs 付近のピークは、立ち上がりの最大値や着火ディレイを捉える。同様に G2 波形ではジッタ、Vk 波形ではディレイの影響を捉えている。

以上の分析により、主成分分析によって、特徴的な波形を主成分として取り出せる可能性が見いだされた。使用する主成分の数は、後段の評価を行った結果、6 で充分と判断した。このときの累積寄与率は 90 % 超となっていた。

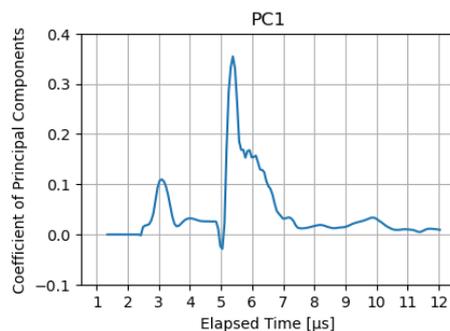
以降、ここで求めた係数を使用して、学習データから求めた主成分を主成分 A、同じ係数を使用してそれ以



(a) Waveforms of new thyatron (0 hours of use)



(b) Waveforms of deteriorated thyatron (13,700 hours of use)



(c) Coefficients of the first principal component for thyatron data

Figure 2: Typical waveform data of G1 grid voltage for thyatron and part of the results of principal component analysis. In (a) and (b), each waveform is superimposed for almost 100 shots.

外のデータから求めた主成分を主成分 B と呼ぶ。

3.2 ロジスティック回帰による分類

ロジスティック回帰[6]は事象の確率を元に 2 値分類を行う手法で、ある個体が“故障”である確率を求め、50 % 以上なら“故障”、それ以外なら“正常”と判定する。主成分 A を学習データ、主成分 B を検証データとした。

ロジスティック回帰のパラメータはグリッドサーチにより G1、G2、Vk 波形それぞれで選定した。検証データでは

ほぼ 100 %の正解率となったことから、ロジスティック回帰による分類を行えると判断した。

なお、入力 1 ショットごとに分類結果を出力する。

3.3 サポートベクターマシンによる分類

サポートベクターマシン[7]は 2 値分類の手法で、ある N 次元パラメータ群を分割する超曲面を計算し、分割された一方のエリア内を“故障”、他方を“正常”と判定する。主成分の数が 6 のため、N = 6 とした。ロジスティック回帰と同様に、主成分 A を学習データ、主成分 B を検証データとした。

サポートベクターマシンのパラメータはグリッドサーチにより G1、G2、Vk 波形それぞれで選定した。検証データでほぼ 100 %の正解率となったことから、サポートベクターマシンによる分類を行えると判断した。

なお、入力 1 ショットごとに分類結果を出力する。

3.4 線形回帰による稼働時間推定

前述の分類では、1 ショットごとの判定としたが、ショット間の波形の揺らぎ(残留電圧変動)も重要な要素である。

揺らぎを評価するため、主成分の共分散行列を計算し、その L2 行列ノルムおよび行列式について自然対数を取ったものと、サイラトン個体の稼働時間を比較したところ、いずれも正の相関があることが分かった。

そこで、線形回帰[8]を用いて、主成分の共分散行列から稼働時間を推定することにした。これにより、実際の稼働時間よりも推定稼働時間が大きい場合、劣化の度合いが高い、という知見が得られる。

学習データとしては、2.9 の個体のうち、稼働時間が 5,000 時間未満および 20,000 時間超の個体のデータを用いて主成分分析を行ったものを用い、検証には 2.9 の個体に加え、2.8 で除外した自爆回数が多い個体(11 個)の主成分も用いた。学習データの相関係数の評価結果を Table 2 に示す。Vk 波形に比べて G1、G2 波形のほうが相関係数の値が大きいという結果となった。

Table 2: Correlation Coefficient between Operating Hours and Natural Logarithms

	Correlation coefficient		
	G1	G2	Vk
Natural logarithm of the L2 norm	0.83	0.85	0.66
Natural logarithm of the determinant	0.95	0.94	0.59

このうち、相関係数が最も大きい G1 波形の行列式のデータを Fig. 3 に図示する。学習データに対応する点および回帰曲線を描画しておりはっきりとした相関が見られる。また検証データが回帰直線の周りに配置されており、未知のデータでもよく傾向を表していると言える。

4. 異常検知ソフトウェアの開発

以上の結果を元に、サイラトンの劣化度を推定するとともに、推定稼働時間を併せて出力するソフトウェアを開発した。以降、解析対象のデータを対象データと呼ぶ。

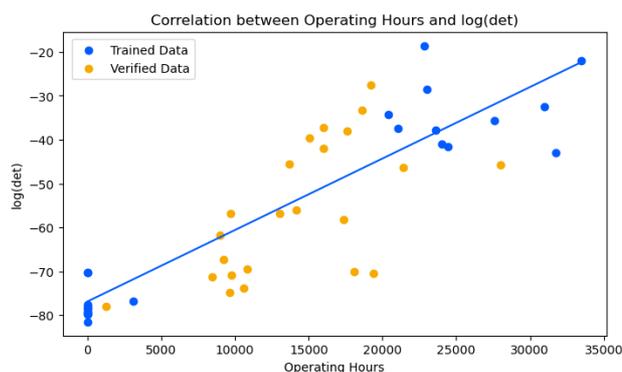


Figure 3: Correlation between operating hours and the natural logarithms of the determinants.

4.1 異常検知ソフトウェアの概要

異常検知ソフトウェアは Linux 上で動作するコンソールアプリケーションとし、Python 言語にて実装した。各機能は scikit-learn を用いて実装した。また、結果の表示は、ウェブブラウザで表示可能なように HTML 形式で出力することとした。複数個体のデータを一度に処理可能とし、結果全体のサマリー画面を用意した。

4.2 主成分分析の実装

2.9 の全個体のデータにより計算した主成分の係数を用いて、対象データの主成分を計算する。

4.3 ロジスティック回帰の実装

決定済みのパラメータで、2.9 の全個体のデータを用いて学習したロジスティック回帰モデルを用いて、対象データの主成分に対し、故障確率を計算する。計算は 1 ショットごとに行い、個体ごとに平均値および中央値を取り、集計結果とした。

4.4 サポートベクターマシンの実装

決定済みのパラメータで、2.9 の全個体のデータを用いて学習したサポートベクターマシンモデルを用いて、対象データの主成分に対し、“正常”と“故障”の分類を行う。分類は 1 ショットごとに行い、個体ごとに Softmax 関数を用いて“故障”の確率を計算し、集計結果とした。

4.5 線形回帰の実装

対象データの主成分の共分散行列に対し、L2 行列ノルムおよび行列式を計算し、それぞれの自然対数を取る。3.9 で学習した線形回帰モデルを用いて、計算した L2 行列ノルムの自然対数から推定稼働時間を計算し、それを集計結果とした。行列式の自然対数については、その値のみ集計結果に加えた。

4.6 サマリー画面の実装

サマリー画面の例を Fig. 4 に示す。サイラトン個体ごとにロジスティック回帰から算定した故障率の平均値、推定稼働時間(単位:時)、サポートベクターマシンから算定した故障確率を表示する。また、サポートベクターマシンから算定した故障確率のうち G1、G2、Vk 波形いずれかによるものが 0.5 を超えている場合、サイラトン個体の状態を“故障”、それ以外を“正常”と表示するようにした。なお、推定稼働時間の計算上、負の値になることが

あるが、補正は行わない。

このサマリー画面を見ることで、どの個体の劣化度が高いかが一目瞭然となる。また、複数の指標で劣化度を確認できるため、交換の判断材料が多くなっている。

サイラトロン個体の名称をクリックすると、個別詳細画面が表示されるようにした。

サイラトロングリッド波形解析実行サマリー

サイラトロン個体	判定結果									判定
	故障率回帰			推定稼働時間			SVM故障確率			
	G1	G2	Vk	G1	G2	Vk	G1	G2	Vk	
sn2010_33504h	1.00	1.00	0.99	33623	33847	51897	1.00	1.00	1.00	故障
sn2219_17592h	0.44	0.58	0.64	28083	25104	16994	1.00	1.00	1.00	故障
sn2424_28000h	0.95	0.90	0.95	21628	23507	-2070	1.00	1.00	1.00	故障
sn2899_31000h	0.99	0.99	1.00	25236	20769	15286	1.00	1.00	1.00	故障
sn3939_10600h	0.04	0.03	0.03	-2089	-3313	-20	0.00	0.00	0.00	正常
sn4086_13000h	0.08	0.05	0.04	15773	18472	12674	0.00	0.00	0.00	正常
sn4469_1536h	0.00	0.00	0.01	-5893	-7020	-936	0.00	0.00	0.00	正常
sn4557_0h	0.00	0.00	0.00	-10027	-8620	-10649	0.00	0.00	0.00	正常

Figure 4: Example of a summary window.

4.7 個別詳細画面の実装

個別詳細画面では、下記項目が 1 画面に表示される。画面については一部のみ紹介する。

- サイラトロン個体の判定結果。
- 入力波形。
- 主成分の散布図行列。
- 稼働時間に対する故障確率。
- 分類結果図。
- 主成分の共分散行列。
- 共分散行列の行列ノルムおよび行列式。
- 個別の集計結果。

冒頭では個体の判定結果をまとめて表示する。出力日、タイトル、個体名、入力元ディレクトリ名、しきい値設定ファイル名、各判定結果が文字列で表示される。

稼働時間に対する故障確率としては、学習データおよび対象データについて、ロジスティック回帰による故障確率を、横軸を稼働時間、縦軸を故障確率としてプロットする。稼働時間により故障確率が上がっていくこととの関連性を把握しやすくなる。

分類結果図の例を Fig. 5 に示す (G1 波形をロジスティック回帰で分類した結果)。学習データおよび対象データについて、第 1、第 2 主成分をプロットするとともに、“正常”と“故障”のエリアを区別した図となっている。緑色のエリアが“正常”、灰色のエリアが“故障”と分類されることを表しており、青色の点が灰色のエリアにあれば“故障”と判定されることを意味する。ただし、本取り組みでは主成分の数が 6 であり、本図では第 3~6 主成分が省略されている。エリア分けは、第 3~6 主成分について対象データの中央値を仮に当てはめた結果であり、模擬的な図となっているが、主成分分析で捉えた特徴が分布のどの位置にあるのか把握しやすい図となっている。

稼働時間に対する共分散行列の行列ノルム、行列式としては、横軸を稼働時間、縦軸を共分散行列の行列ノルム (上段) および行列式 (下段) とし、縦軸はそれぞれ自然対数を取ったグラフが表示される。推定稼働時間を求める目安となる。

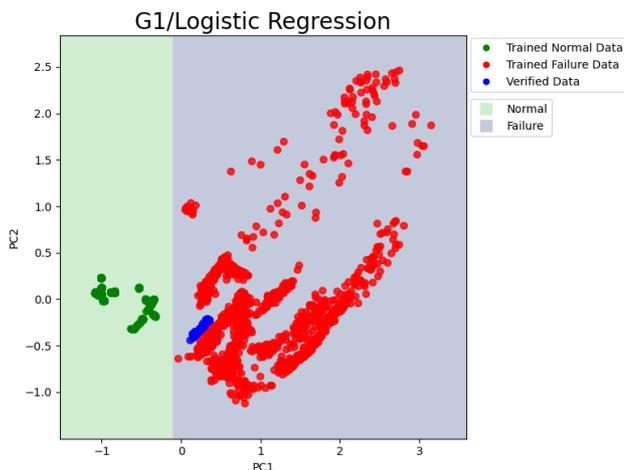


Figure 5: Example of a classification results chart.

このように図示することで解析結果を視覚的に理解しやすくしている。

4.8 実環境での評価

SACLA に設置された計算サーバーでの動作確認では、個体あたり 20 秒程度の時間で計算と結果の出力が完了した。これにより、故障確率などの指標と分析データの図を比較的短時間で閲覧できるようになった。また、新規に取得した未知のグリッド電圧波形データによりソフトウェアを評価したところ、サイラトロンの状態と目視による波形の様態、ソフトウェアの出力結果が一致し、新品と故障品をほぼ 100% の精度で分類できること、十分な精度で稼働時間推定ができること、システムが正常に動作していることが確かめられた。

5. 結論と今後の展望

SACLA で蓄積されているサイラトロンのグリッド電圧波形データを用いて、サイラトロンの劣化度を自動で評価するソフトウェアを開発した。機械学習の手法を用いることで、サージ電圧の増加などから劣化度を推定できることを示した。開発したソフトウェアにより、サイラトロンの交換作業にかかる保守業務を補助し、加速器施設の安定稼働に資することが期待される。

今後、このソフトウェアを SACLA の加速器全体に展開していく予定である。また、改善の余地としては、故障の前兆を実際に故障する 1 週間程度前に検知できれば、計画的な交換を行えるため、実環境で“故障”と判定された個体の追跡調査を行うことが考えられる。また、G1、G2、Vk 波形から最小限のデータへの絞り込みや、グリッド電圧波形の測定と推定の最適化など、効率化と性能向上についても引き続き取り組む予定である。

謝辞

本件では、SACLA 関係者の皆様より大変多くのご助力をいただきました。皆様に心より感謝申し上げます。

参考文献

- [1] T. Inagaki *et al.*, “SACLA 主加速器の運転と保守の状況”, Proc. PASJ2015, Tsuruga, Japan, Aug. 2015, pp. 925-929.

- [2] Y. Sato *et al.*, “SACLA における機械学習による故障予知・異常検知アプリケーションの試作”, Proc. PASJ2022, Kitakyusyu (Online meeting), Japan, Oct. 2022, pp. 14-18.
- [3] E. Iwai *et al.*, “SACLA での機械学習による加速器の運転合理化に向けた取り組み”, 「加速器」, Vol. 20, No. 2, 2023, pp. 90-99.
- [4] S. Nakazawa *et al.*, “SACLA でのサイレントトラブルと対策”, Proc. PASJ2017, Sapporo, Japan, Aug. 2017, pp. 1000-1004.
- [5] https://en.wikipedia.org/wiki/Principal_component_analysis
- [6] https://en.wikipedia.org/wiki/Logistic_regression
- [7] https://en.wikipedia.org/wiki/Support_vector_machine
- [8] https://en.wikipedia.org/wiki/Linear_regression