

KEK 電子陽電子入射器のデータアーカイビングシステムとビッグデータ解析 DATA ARCHIVING SYSTEM AND BIG DATA ANALYSIS FOR THE KEK e+/e- LINAC

宮原房史^{#,A)}, 早乙女秀樹^{B)}, 大房拓也^{B)}, 草野史郎^{A)}, 工藤拓弥^{B)},
佐武いつか^{A)}, 佐藤政則^{B)}, 王迪^{A)}

Fusashi Miyahara^{#,A)}, Hideki Saotome^{B)}, Takuya Ofusa^{B)}, Shiro Kusano^{C)}, Takuya Kudou^{C)},
Itsuka Satake^{A)}, Masanori Satoh^{A)}, Di Wang^{A)}

^{A)} High energy Accelerator Research Organization (KEK)

^{B)} Kanto Information Service (KIS)

^{C)} Mitsubishi Electric System & Service Co., Ltd

Abstract

The KEK e+/e- LINAC supplies beams to the SuperKEKB HER, LER, PF, and PF-AR, operating at a maximum beam repetition rate of 50 Hz, which is allocated to each circular accelerator according to its requirements. Since the injection efficiency and stability of each circular accelerator are influenced by the injected beam orbit and emittance, data on the injected beam and device parameters are essential for experimental understanding. In LINAC, synchronized data from beam position monitors, RF monitors, and pulsed magnets are archived for every beam. Since the volume of data is enormous, they are converted into binary files (HDF5), which allow data compression and fast readout. These data can be synchronously analyzed together with data stored in the EPICS Archiver Appliance, such as LINAC device data, environmental temperatures, and injection efficiencies into the circular accelerators. In addition, a web-based simple analysis tool has been developed to make the data accessible and analyzable by anyone.

1. はじめに

KEK 電子陽電子入射器(入射器)は SuperKEKB HER、LER および PF、PF-AR へビームを供給しており、最大 50 Hz のビーム繰り返しを各加速器の要求に応じてそれぞれに割り当てる運転を行っている。各加速器へ入射効率やビームの安定性に寄与するパラメータや稀に発生する異常な軌道の特定などにはビームと同期したデータが必要になる。入射器では全てのビームに関して、ビーム位置モニターと RF モニター、パルス電磁石の同期データをアーカイブしている。データ量が膨大となるためデータ圧縮と高速読み出しが可能なシステムが必要となる。またほぼ全ての機器は EPICS を用いて制御されており、機器のパラメータや測定値などのデータは Archiver Appliance を用いて保存されており、パラメータの時間的な変動などの確認や同期データと合わせた解析も可能となっており、統計的な解析だけでなく、機械学習を用いたデータ解析も進められている。誰もが容易に解析を可能とするために WEB ベースの簡易解析ツールの開発を行っている

2. データアーカイビングシステム

入射器は複数の円形加速器にビーム入射を行っており、イベントタイミングシステムを用いて入射する加速器に合わせてパルスごとの加速器制御を行っている[1]。イベントシステムは入射先の加速器に合わせたビームモード(実際はビームモードと目的別に分けられたタイミングで定義されたイベントコード)とショットごとに割り振られたショット ID を配信しており、それぞれ機器の制御・データ

へのビームモードタグ付けと同一ショットの判別に利用している。

入射器では各機器から直接得た 1 次データとデータ解析用にまとめられた 2 次データを保存している。さらに 1 次データは EPICS を通じて取得した EPICS PV 単位(ほぼ全ての機器の設定値やモニター値、温度などの環境の測定値)とビームショットごとのイベントベースの同期データ[2]の 2 種類がある。機器単位のデータの保存と読み出しは Archiver Appliance[3]を利用しており、各データは Angular を用いたビューワーで高速な可視化が可能である[4]。ビーム位置モニター(BPM)、RF モニター[5]、パルス電磁石[6]に関してはそれぞれ、全モニターに関して同一ショットでまとめた生データを保存している。各生データは 10000 ショットごとにテキスト形式でまとめられ、gzip で圧縮して保存している。

Archiver Appliance は EPICS の PV ごとにデータを自動収集し、効率的に長期アーカイブが出来る。Web UI や REST API を通じて容易にデータの読み出し、可視化が可能であり、複数クライアントからの並列アクセスに強いなどのメリットもある。設定値やリアルタイムのデータの確認、複数のパラメータの時間変化をみることで相関を確認することが出来るため非常に役立っている。また長期間のパラメータのドリフトの確認にも欠かせない。一方、PV 単位のデータで、ビームと非同期のサンプリングデータであるため詳細な解析には不向きである。各パラメータが非同期であるため、ビームと相関があるパラメータがあってもショットごとの位置(軌道や分散部におけるエネルギー)の揺らぎと同程度以下の影響しかない場合は、ビームの揺らぎに隠れてしまい、相関を見つけることが困難になる。また、時間的なパラメータのドリフトとの区別も難しい。

[#] fusashi.miyahara@kek.jp

同期データは同一ショットを比べるため、ビームに影響するパラメータや軌道と相関するパラメータをより正確に解析できる。また、全てのショットが保存されているため、前のショットの影響(例えば熱電子銃などは引き出し電流量でカソードの熱負荷が異なるため、PF 入射用の 0.3 nC のビームの前が陽電子生成用の 12 nC ビームか否かで電子銃直後のビームが異なる)なども調べることが可能である。同期データは任意のパラメータを指定した時間範囲のデータを生データから抽出可能であるが、BPM、RF モニター、パルス電磁石のデータは独立に保存されているため、データの読み出しと同期イベント抽出に時間を要する。この方法は数分程度の短期間のデータ抽出では問題ないが、数時間以上のデータを出力するには適さない。BPM、RF モニター、パルス電磁石の全データをまとめるとパラメータの数は約 2050 個となり、10000 ショットのテキストデータで 160MB 程度となる。そこで高速な同期データ読み出しを可能にするために全パラメータを記録したバイナリファイルを用意している。ファイルと読み出しに使用するデータのフォーマットごとに全パラメータ 100000 ショット(テキストデータで 1.6 GB)のデータ読み出しに要する時間を Table 1 にまとめた。入射器では同期データを 1 時間単位で圧縮した HDF5 形式のファイルにまとめている。Python のデータ解析用のライブラリ Pandas [7]は Dataframe を HDF5 [8]に変換

Table 1: Comparison of File Formats

File Format	Comp. Lib.	Data format	Comp. Ratio	Read out time [s]
TEXT (Original)	None	Pandas, Dataframe	100	12.2 ± 0.1
TEXT (compressed)	zlib, LV=1	Pandas, Dataframe	42	17.2 ± 0.1
HDF5	zlib, LV=1	Pandas, Dataframe	32	3.64 ± 0.02
HDF5	blosc, LV=1	Pandas, Dataframe	49	1.76 ± 0.02
ROOT	zlib, LV=1	ROOT TTree	31	0.61 ± 0.01
ROOT	zlib, LV1	ROOT RDataframe	31	1.43 ± 0.03

して保存することが可能で、ファイルを読み出し、Dataframe に代入することが出来る。HDF5 形式は部分読み出しが可能で、Dataframe に指定したパラメータのみを代入出来る(format='table'で変換した場合)ためメモリの節約が可能である。HDF5 の他に ROOT フォーマットの検討も行った。ROOT [9]はデータフォーマットに TTree と RDataframe が選択可能であり、どちらも HDF5 に比べて高速処理可能であるが、Pandas 等 Python の他のライブラリから直接読み出すことが出来ず、機械学習等で使用する場合は一度フォーマットを変換する必要があるため同期データ用バイナリファイルには HDF5 形式を採用した。

3. ビーム同期データ解析

アーカイブされた同期データと Archiver Appliance のデータは異常な軌道が発生した際の原因特定やビーム品質の向上に寄与するパラメータの発見や確認などに役立っている。これらのデータを用いた解析事例を3つ紹介する。

3.1 異常な事象の原因特定

同期データを用いた SuperKEKB 入射に伴うビームアボートの例を示す。入射器ではアボートが発生したとき、その数分前からアボート発生までの同期データを即座に解析し、異常な軌道の発生の有無を確認している。Figure 1 に LER アボートが起きた際のビーム軌道の変化とその原因となったクライストロン RF 出力の変動の様子を示す。赤色の点がアボート発生時のデータに対応しており、軌道、RF 振幅、位相がそれまでとは離れた値となっている事が分かる。この BPM は分散関数の大きな位置($\eta=1.67$ m)に設置されており、アボート時のビームエネルギーは通常より低くなっている。異常があったクライストロンはダンピングリングから出射されたビームをゼロクロス位相に乗せて、パンチ圧縮に用いるが、このときは減速位相に乗ってしまったことがわかる。この例の様に、入射器ではアボート発生時に同期データを解析し、RF、パルス電磁石の異常を確認するシステムを用意している。

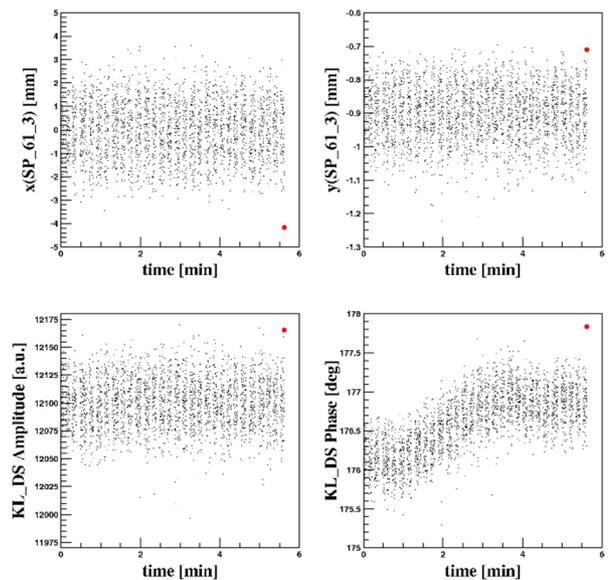


Figure 1: Beam orbit deviations observed just before a beam abort during LER injection (upper), and the RF amplitude and phase of the klystron where the problem occurred (lower). The red points correspond to the beams that triggered the abort.

3.2 説明可能 AI を用いた需要パラメータの抽出

入射器では説明可能 AI SHAP (Explainable AI: XAI) を用いて、ビームに重要なパラメータの抽出を試みている[10]。加速器は多数のパラメータがあり、それらは環境(温度等)により変化するため統計学的手法でビームに大きく寄与するパラメータを特定するのは容易ではな

い。XAI はある出力(モデルの予測値)に対して入力(モデルに与える特徴量)の重要度を評価する事が可能である。SHAPを用いて重要なパラメータを特定した事例を紹介する。入射器の第3スイッチヤードのビームダンプリンでは Axion-Like Particle (ALP)探索の予備実験が進められているが、この実験ではダンプリンのビームロスを経験可能な限り小さくすることが重要となる。SHAP を用いた解析から J-ARC 上流部の RF 位相がビームロス低減に大きく寄与する予測が得られ、実験的にも確認できた [11]。今後は解析の対象を増やし、安定運転に寄与するパラメータの抽出なども行っていく。

3.3 ビッグデータを用いた相関解析

加速器では物理的に相関が予想できるパラメータが多く、実際にその様な相関を実験的に確認することは重要である。また、XAI を用いた解析ではビーム軌道の様に入力が互いに相関を持つ場合、重要度の評価は難しくなる。機械学習を用いた解析と統計的な解析は相補的な関係となるため、両方を組み合わせることで理解が深まると期待できる。統計的な手法で HER の入射効率に影響するパラメータを特定した例を紹介する。

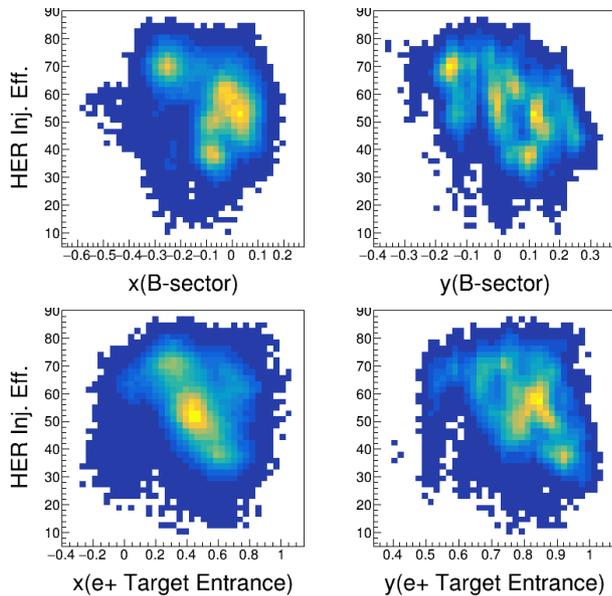


Figure 2: Correlation between beam position and injection efficiency into the HER. The upper figures show the correlation with the beam position measured in the upstream B-sector of the injector, while the lower figures show the correlation with the position measured by the BPM located just upstream of the positron target.

HER 入射ではエミッタンス増大を抑えながらビームを輸送することが重要である。入射器ではビームが加速管を通過する際のウェイク場でバンチ後方がキックされることがエミッタンス増大の理由の一つとなる。したがって、エミッタンス増大の場所の特定は重要な課題である。HER の入射効率とビーム位置の相関を Fig. 2 に示す。入射器上流部 (B-sector) ではいくつかの島構造が現れているが、陽電子標的直前ではきれいな相関が確認できる。図には示していないが、ビームの角度もきれいな相関が確認できる。電子ビームは陽電子標的に開けられた穴を通過

するため、ビームロス無く穴を通過するために軌道が制限される。陽電子標的以降のオプティクスは陽電子ビームの透過率が最大となるように調整されているため、標的からダンピングリング入射分岐路までの光学系は β 関数が大きくなっている。そのため軌道の変化も大きくなり加速管でのウェイク場の影響も他の区間に比べて大きくなっていると考えられる。HER 入射のビーム調整では、ターゲット以降の軌道調整を調整することでエミッタンスが改善することが実験的にも確認されており、この結果と矛盾が無い。

4. 簡易データ WEB 解析ツール

入射器では Archiver Appliance で収集された任意のデータは、WEB 上ですぐに確認できる。しかし、同期データに関してはその様なツールは用意されておらず、出力したデータを Python スクリプトや ROOT 等のツールで解析する必要があり、解析にはある程度の知識と経験が要求される。そのため、誰もが解析できる状況にはない。そこ誰もが容易に同期データを解析できる様に WEB ベースの解析ツールの開発を進めている。

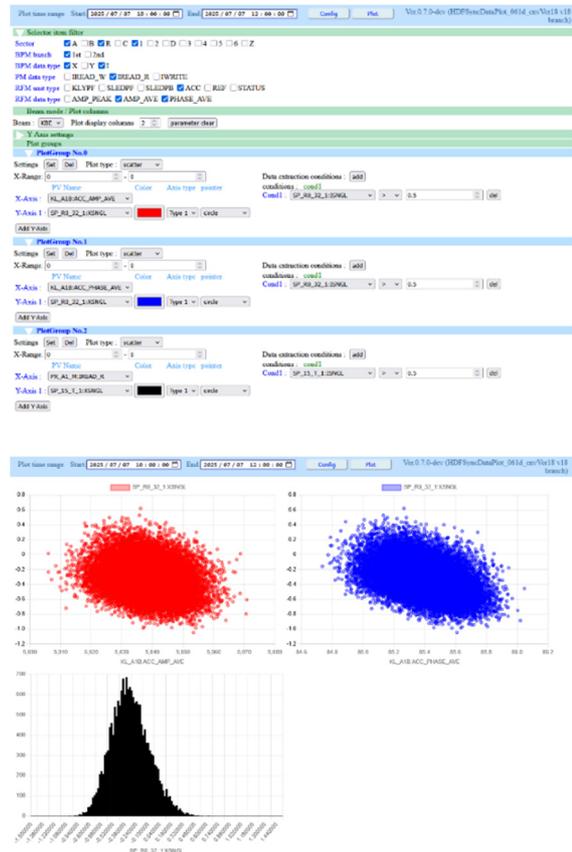


Figure 3: Input and output screens of the web-based simple analysis tool.

Figure 4 に解析ツールの入力と出力画面の例を示す。入力には任意のパラメータ (BPM で測定した位置、電荷量、RF モニターで測定した RF 振幅、位相、パルス電磁石電流値) が選択可能で、それぞれの相関や分布が確認できる。図では入射器上流の RF 振幅・位相と J-ARC 中央の分散関数の大きな位置でのビーム位置の相関を

プロットしている。解析には任意の条件を課すことが可能で、この例では電荷量の要求を課している。HDF5 フォーマットで保存された同期済みのデータを解析するため、結果を高速で出力可能である。今後は Archiver Appliance のデータとの相関解析も可能な様に機能を拡張する予定である。

5. まとめと今後の展望

KEK 電子陽電子入射器ではビームと同期した機器 (BPM、RF モニター、パルス電磁石) の生データおよびこれらが全てをまとめた HDF5 フォーマットで保存している。他のほぼすべての機器のデータは EPICS Archiver Appliance でアーカイブしている。これらのビッグデータは高速に解析が可能であり、機械学習を用いた重要パラメータの推定や相関解析などの統計的な解析で利用され、入射器の性能向上に役立っている。説明可能 AI SHAP を用いた解析ではビームロス低減の調整に重要なパラメータの抽出に成功しており、今後、他の調整や加速器の不安定要因の特定などにも適用していく予定である。また、誰もがビームデータを容易に解析できるツールを開発しており、今後も操作性の改善や機能の拡充を進めていく。

参考文献

- [1] H. Kaji *et al.*, “Construction and Commissioning of Event Timing System at SuperKEK”, Proc. IPAC’14, Dresden, Germany, June 2014, pp. 1829-1832.
- [2] F. Miyahara *et al.*, “KEK 電子陽電子入射器ビーム同期データ収集、解析システム開発”, Proceedings of the 19th Linear Accelerator Meeting in Japan, Online, Oct. 2022, pp. 927-930.
- [3] EPICS Archiver Appliance, <https://epicsarchiver.readthedocs.io/en/latest/>
- [4] I. Satake *et al.*, “KEK 電子・陽電子入射器における Archiver Appliance の運用状況”, Proceedings of the 17th Annual Meeting of Particle Accelerator Society of Japan, Online, Sep. 2020, pp. 735-738.
- [5] H. Katagiri *et al.*, “RF Monitor System for SuperKEKB Injector LINAC”, Proc. IPAC’18, Vancouver, Canada, April 2018, pp. 2128-2130.
- [6] Y. Enomoto *et al.*, “A New Pulsed Magnet Control System in the KEK Electron Positron LINAC”, Proc. IPAC’18, Vancouver, Canada, April 2018, pp. 2121-2123.
- [7] <https://pandas.pydata.org/>
- [8] <https://www.hdfgroup.org/solutions/hdf5/>
- [9] <https://root.cern/>
- [10] K. Uemura *et al.*, “KEK 電子陽電子入射器調整性能向上に寄与する重要パラメータの推定”, Proc. 21st Annual Meeting of Particle Accelerator Society of Japan (PASJ2024), Yamagata, Japan, Aug. 2024, pp. 637-642.
- [11] K. Uemura *et al.*, “EBES 実験のための説明可能 AI を用いた KEK LINAC 調整性能向上に寄与する重要パラメータの推定”, The Physical Society of JAPAN 2025 Spring Meeting, Online, March 2025.