ISBN: 978-3-95450-261-5 ISSN: 2673-7000 doi: 10.18429/JACoW-NAPAC2025-MOP007

ANOMALY DETECTION OF SLOW-MOVING VARIABLES AT LANSCE FOR IMPROVED BEAM QUALITY*

E.-C. Huang†, N. Yampolsky
Los Alamos National Laboratory, Los Alamos, NM, USA

Abstract

Modern accelerator facilities operate with a large amount of variables, many of which can influence beam quality. While most of these variables are constrained within predefined boundary conditions, slow fluctuations over extended periods—from tens of minutes to a full day—can still significantly degrade beam performance. Due to their gradual nature and the difficulty in distinguishing meaningful trends from background noise, such variables often go unnoticed and remain unoptimized by operators for days.

This study investigates the use of machine learning algorithms to identify and analyze these slow-moving variables. By applying advanced time-series analysis and feature importance ranking, the proposed approach reveals hidden correlations between slow variable drifts and one key beam quality metric: the ring loss at the Los Alamos Neutron Science Center. The results demonstrate the potential of machine learning to detect subtle anomalies and offer actionable insights to mitigate persistent beam quality issues that can disrupt operations for weeks at a time.

INTRODUCTION

Modern accelerator facilities like the Los Alamos Neutron Science Center (LANSCE) rely on hundreds of thousands of interdependent variables to maintain beam stability and performance. While many of these variables are normally bounded by alarms based on standard tolerances, a significant class of slow-moving variables — those that drift subtly over timescales ranging from tens of minutes to one day — often escape the alarm detection. These slow drifts can introduce long-term degradation in beam quality, particularly impacting target metrics such as the ring loss, which is critical to operational efficiency. Unlike transient faults, which are typically caught by the alarm system, these trends are masked by noise, rarely cross thresholds, and may involve high-dimensional correlations.

In addition to degrading performance, slow drifts can also mislead operator judgment. Because these changes unfold gradually, their effects can overlap with operator interventions. For instance, a recent tuning action may seem to improve or worsen beam quality, when in fact a slow-moving variable was the true cause. This makes it harder for operators to evaluate the impact of their changes and can delay effective optimization.

The situation worsens for LANSCE as we are migrating from the obsolete DataWatcher [1] system developed in-house decades ago to the modern Control System Studio/Phoebus [2]. Furthermore, old equipment often lacks the extensive diagnostics ready for an alarm system.

To address this, we developed a lightweight analysis framework that targets two distinct types of slow anomalies. First, we use dynamic correlation analysis to detect undesired oscillations. Separately, we apply L1-based feature selection to rank variables based on their relevance to the drift of key performance indicators, such as the ring loss.

We integrate this analysis into a real-time GUI interface called *BlameIt*, enabling seamless queries of LANSCE archivers, interactive diagnostics, and targeted analysis. Our results show that this approach is well-suited to detecting operationally relevant anomalies that traditional alarms or residual-based models tend to miss.

DATA DESCRIPTION AND METHODS

LANSCE is currently at its 3rd generation of data archiver using the EPICS Archiver Appliance (EAA) [3] since 2024, while InfluxDB [4] was used briefly between 2019 to 2024, and DSRP, a home-grown system, was used since the 90s. A data warehouse was developed to unify the archived data access for all three periods. Asynchronous data acquisition and improved regular expression search has been added to the local fork of arvpyf [5]. Currently, LANSCE operates with more than 200k PVs while >130k channels are archived with various intervals, mostly between 1s to 1 minute. However, for analysis simplicity and intuitive understanding for operators, we limit our analysis to PVs with the LANSCE standard PV names, which contain strictly 10 capital alphanumeric characters marked by the PV's respective beamline, device type, device number, and the channel ID. PVs with non-standard names are generally derived or intermediate PVs that are often strongly collinear with at least one standard PV.

Furthermore, since we are focusing on slow-moving effects, the analysis by default excludes most binary PVs. Binary PVs generally causes sudden jumps that could easily trigger an alarm or fast protect. To reduce PVs with high collinearity with the target PV, say the ring loss, the analysis by default also excludes most beam diagnostics PVs, like beam currents and individual loss monitors. These variables co-moves with the target PV but are not the contributor. After these selection process, the analysis is left with around 5k variables.

Given the high dimensionality of the system and the presence of strong collinearity among many of them, our anomaly detection strategy focuses on two complementary goals: (1) the detection undesired oscillatory behavior and (2) the identification of gradual drifts. Furthermore, since the goal is to deploy it as a near-online tool for the

* Work supported by LANL's LDRD program under 20240474MFR
† en-chuan@lanl.gov

ISBN: 978-3-95450-261-5 ISSN: 2673-7000 doi: 10.18429/JACoW-NAPAC2025-MOP007

control room, our strategy is shaped by a few design priorities:

- Robustness over fine-tuning: We prefer to reliably detect a meaningful subset of impactful anomalies, rather than the comprehensive set of parameters.
- Interpretability over complexity: Techniques are chosen to provide actionable insights to human operators.
- Efficiency for real-time use: The entire pipeline must run fast enough for near-real-time interaction by the control room.

To address these goals, we employ two distinct methods: dynamic correlation and Lasso regression.

Dynamic Correlation Analysis

Using a sliding time window (typically 60–120 minutes depending on the targeted oscillatory period), the correlation of between each PV under analysis and the target PV is calculated. In an ideal running condition when the beam is optimized, the target PV should have minimum correlation with the fluctuation of control variables (co-moving diagnostics channels are excluded from the analysis by default). A general correlation method has also been tried but it showed less sensitivity, since some of the patterns might be intermittent.

Furthermore, PVs within a subsystem maintain high mutual correlation, and therefore, they will all appear as highly correlated with the target PV, providing an insight for operators regarding the source of the fluctuation. This method is unsupervised and well-suited to detecting oscillatory behavior without needing a training phase.

Lasso Regression

To identify which PVs are most influential to slow variations, we apply Lasso(L1) regression, implemented via `tensorflow`, using a history of PV time series and the target PV as the metric. Due to the inhomogeneous data types, a robust scaler is used via the interquartile range. Further data cleaning like removing data with sparse values are applied. Since accelerators contain many groups of highly colinear PVs, the L1 penalty encourages sparsity, highlighting one PV in each colinear group. In comparison, Ridge/L2 regression with the square of the coefficients as penalties normally keeps all the PVs in the colinear group in the model. While L2 provides a more complete set of information, the L1 method provides a much less clustered results for operators and the ranking of each colinear group is more intuitive. For an online analysis tool, we aim to provide a relatively small set of PVs that can guide the operators toward the correct direction instead of a complete list that is hard to determine the next action.

Trade-off Between Algorithms

These methods were selected after brief evaluation of several alternatives, including simple correlation studies, pairwise correlation matrices, Fourier transforms, encoder-decoder schemes, Elastic Net (a combination between L1 and L2), and point-biserial correlation for binary PVs. In practice, these approaches were less sensitive to oscillations, too computationally intensive, sensitive fine-tuned variables, less interpretable, or ineffective for slow drift detection in our operational setting. Our final approach reflects a tradeoff favoring clarity, efficiency, and operational usability.

Though these methods are used independently, they are presented together in *BlameIt*. Operators can select examine trends over arbitrary time windows with a range cut over the target PV. They can also select PV groups based on beamline, device types, and/or data types, and view correlation structure and variable importance. This design supports both near-real-time anomaly detection and retrospective analysis, enhancing situational awareness without requiring complex model training or high compute resources.

RESULTS AND CASE STUDY

The original concept of this project started in the 2021 run cycle when the beam losses of the LANSCE Proton Storage Ring (PSR) oscillated with a ~50-min interval intermittently. Operators, calling this PSR breathing, checked every suspicious or not suspicious PVs they could think of, and each system owner rechecked their system to no avail. The breathing persisted for more than 3 weeks while the production beam was delivered in this suboptimal status. A preliminary version of the dynamic correlation analysis was quickly developed and pointed to the voltages of the PSR benders. While the bender currents, closely watched in operation as they determine the bending strength, remained stable, the voltages were given much less attention. Furthermore, the oscillations on the voltages were well within the alarm range. After this discovery, the magnet team quickly identified that a 1.5°C fluctuation in the cooling water was the culprit. The acceptable water temperature range was 9.4°C in the specification, and therefore, it did not trigger an alarm.

Figure 1 shows *BlameIt* being used on one day of the data mentioned above. Users select the desired time and value ranges of the target PV. Since this effect was not observed in the linear accelerator, only PVs in the high energy beam transport and the PSR were selected. Furthermore, some beam diagnostics variables, like current monitors and radiation detectors, were excluded from the analysis. Though they are highly correlated with the target PV (the ring loss), they are not contributing to it. Furthermore, the users can choose to exclude binary PV, normally marked by a data type of “L” or “B”. After the initial setup, users can grab data and start the analysis. Depending on the amount of data requested, the data acquisition could take a few seconds to several minutes. However, the improved PV selection via regular expression and concurrent requests have significantly improved the speed. The results are show at the bottom right corner, sorted by the dynamic correlation for this analysis. One can choose which PVs, all normalized, to plot to compare with the target PV.

ISBN: 978-3-95450-261-5 ISSN: 2673-7000 doi: 10.18429/JACoW-NAPAC2025-MOP007

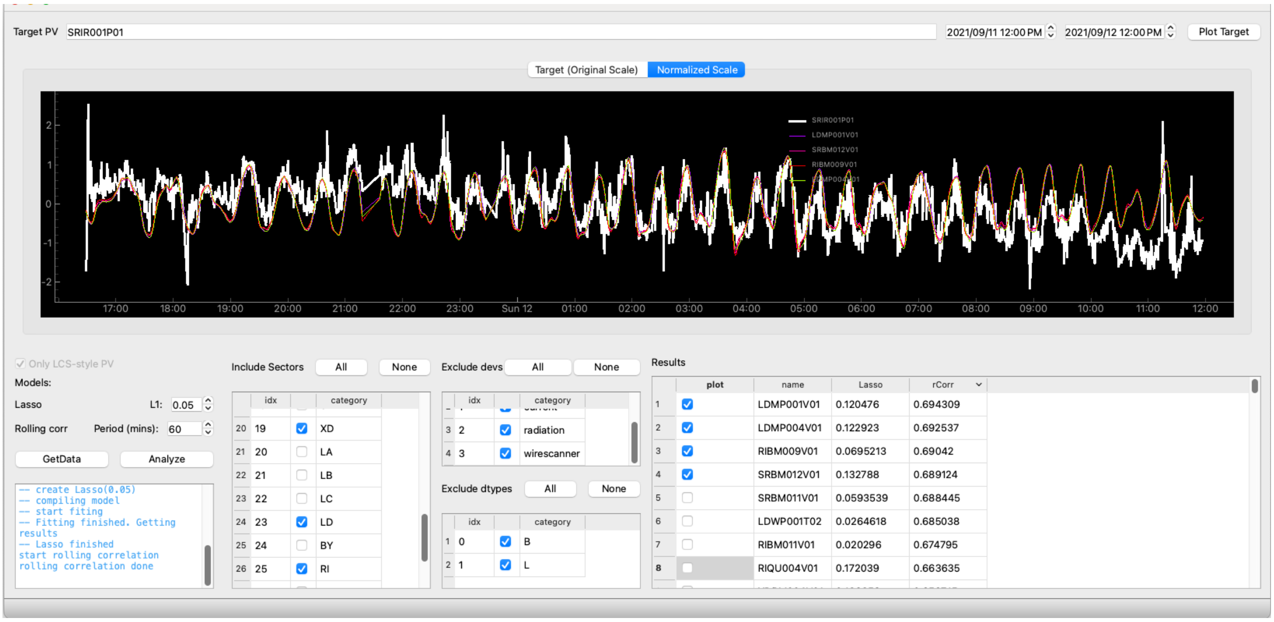

Figure 1: A screenshot of the *BlameIt* software with an analysis focusing on the PSR loss (SRIR001P01) during the period discussed in the text.

FUTURE WORK

Different algorithms can be quickly implemented via *BlameIt* at LANSCE to identify different types of anomalies. We would like to base new models on a selected good region and provide potential PVs contributing to the selected bad region. This could potentially be done via comparing the weighting factors for an L2 model. While we try to tackle the anomaly issues one type at a time, the goal is to implement different anomaly detection algorithms in real time and generate a meaningful and interpretable results for operators.

CONCLUSION

This study demonstrates the effectiveness of machine learning approaches in detecting slow-moving variables that gradually degrade beam performance at accelerator facilities. Our lightweight analysis framework, combining dynamic correlation analysis and Lasso regression, successfully identified the root cause of persistent beam quality issues that existing alarm system missed. The "PSR breathing" case study validates our approach—when operators faced weeks of intermittent beam loss oscillations, our analysis method could potentially identify the problem.

The *BlameIt* interface integration represents a significant operational advancement, allowing operators to interactively explore correlations between system variables and beam quality metrics in near-real-time. While developed for LANSCE, these methodologies are applicable to any complex system where subtle, slow-moving variables impact performance. As accelerator facilities grow in complexity, such tools will become essential for maintaining optimal performance and minimizing downtime.

REFERENCES

[1] M. Pieck, C. Hatch, E. Westbrook, and H. Watkins, "LANSCE's instrumentation and controls system modernization", in *Proc. IPAC'23*, Venice, Italy, May 2023, pp. 4783-4786. doi:10.18429/JACoW-IPAC2023-THPL133.

[2] *Control System Studio/Pheobus*. [Online]. Available: https://www.controlsystemstudio.org/

[3] M. Shankar, L. Li, M. Davidsaver, and M. Konrad, "The EPICS archiver appliance," in *Proc. of ICALEPCS*, Melbourne, Australia, Oct. 2015, pp. 761-764.

[4] InfluxData, *InfluxDB - Time Series Database*. https://www.influxdata.com/products/influxdb

[5] *Archiver Python Frontend*. [Online]. Available: https://github.com/NSLS-II/arvpyf