CONTINUOUS AND AUTOMATED MONITORING OF THE HEALTH AND PERFORMANCE OF THE LHC BEAM POSITION MONITOR SYSTEM

J. Martínez Samblas, A. Boccardi, M. Gonzalez-Berges, M. Krupa, CERN, Geneva, Switzerland

Abstract

We have developed a new automated system to continuously monitor the health and performance of the LHC Beam Position Monitors (BPMs) at CERN. We preprocess turn-by-turn and bunch-by-bunch data through denoising and normalisation before analysing it using a combination of rule-based quality checks and anomaly detection algorithms. All real-time acquisitions are processed on a dedicated Python node, with results stored in HDF5 format and aggregated via the Offline Analysis Framework, which also provides visualisation and reporting. Our system provides robust diagnostics, daily summaries, and dashboards to ensure BPM reliability and maintain high data quality throughout LHC operations.

INTRODUCTION

The Large Hadron Collider (LHC) Beam Position Monitor (BPM) system comprises more than one thousand BPMs organised in 67 Front-End Software Architecture (FESA) devices around the machine [1–3]. It provides turn-by-turn (TBT) and bunch-by-bunch (BBB) measurements that are essential for orbit control, beam diagnostics, machine studies, and routine LHC operation.

Because of its operational importance, we must continuously ensure the quality and availability of BPM data. In practice, acquisitions can be affected by corrupted or incomplete data, inconsistent metadata, missing bunch information, all-zero signals, and waveform distortions such as clipping or saturation. In a system of this scale, manual inspection alone is not sufficient to provide systematic and timely diagnostics.

To address this, we have developed an autonomous monitoring framework combining quality checks, preprocessing, and anomaly detection, with structured outputs for diagnostics, daily summaries, and dashboards.

DATA AND PREPROCESSING

The data used by the monitoring pipeline are obtained in parallel with the standard LHC BPM acquisition channels. They are independent copies of the operational data streams and therefore do not affect the data delivered to operators or orbit feedback.

Input Data

Our framework processes waveform acquisitions produced by the LHC BPM FESA devices. For each acquisition, we retrieve the horizontal and vertical beam position measurements together with the metadata required for their interpretation, such as BPM identifiers, bunch IDs, acquisition dimensions, and timing-related information.

Quality Gates

We apply rule-based quality gates immediately after acquisition to exclude corrupted data, inconsistent metadata, and acquisition failures from anomaly detection. These cases are tracked separately to avoid false positives, for example after missed injections.

The first checks verify the structural integrity of the data. The measured arrays must match the expected dimensions for the corresponding acquisition mode, and the horizontal and vertical planes must have compatible shapes. In addition, the bunch IDs must be consistent across both planes, since any mismatch usually indicates a problem in the acquisition chain or in the associated metadata.

Further checks target clearly non-physical signals. We reject acquisitions containing only zero values, as they correspond to fully invalid measurements. Another validation step enforces that the first sample is equal to zero, which is a safety condition inherited from the LHC BPM FESA system. These gates veto unusable captures before advanced analysis proceeds.

Signal Preparation

After validation, we convert the accepted data into a form suitable for automated analysis. For each transverse plane, we arrange the BPM signals into a matrix M in which each row corresponds to one BPM and each column to one turn, using the turn-by-turn averaged signal as the main input representation. This matrix-based view provides a compact and generic format that we can use with different anomaly detection algorithms without method-specific preprocessing.

An optional denoising step is then applied using singular value decomposition (SVD) [4, 5],

$$M = U\Sigma V^T, \tag{1}$$

where U contains spatial patterns across BPMs, V the corresponding temporal patterns, and Σ the singular values that rank their importance. For denoising, we reconstruct the matrix using only the first r most energetic modes, typically $r = 4$, which capture the dominant correlated content, while weaker modes are discarded as mostly incoherent noise. Figure 1 shows that these first four modes already explain more than 99% of the variance in a representative acquisition, with only marginal gains from additional modes.

ANOMALY DETECTION

We apply anomaly detection independently to the horizontal and vertical planes using zero-value checks, spike detection, SVD-based mode-dominance analysis, Isolation Forest, and tune-based frequency analysis.

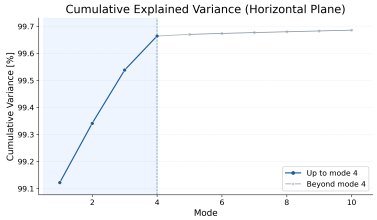

Figure 1: Cumulative explained variance of the leading SVD modes for a representative horizontal acquisition. The first four modes exceed 99% and are retained for denoising.

Rule-Based Methods

The first group of methods targets explicit non-physical patterns. Zero-value checks identify BPM signals that contain an abnormal number of exact zeros. We interpret such cases as signatures of invalid or partially corrupted data rather than genuine beam motion.

A second detector applies a Z-score-based criterion to identify spikes in each BPM turn series. For a BPM i, we evaluate the signal through a normalised deviation measure of the form

$$z_{i,j} = \frac{x_{i,j} - \mu_i}{\sigma_i}, \tag{2}$$

where $x_{i,j}$ denotes the position at turn j, and μ_i and σ_i are the mean and standard deviation of that BPM signal. We flag BPMs when they contain excursions beyond a configurable threshold. This method is intended to capture isolated abrupt transients that affect only a limited number of turns and may therefore escape more global analyses.

Advanced Detection Methods

For subtler deviations, we exploit the same SVD introduced in the preprocessing stage, but now also as a diagnostic tool. The leading modes usually describe coherent beam motion shared by many BPMs and carry almost all of the variance used in the denoised reconstruction. However, a sufficiently strong BPM-specific fault may also enter this high-variance subspace. We therefore verify not only how much variance is retained, but also whether the retained spatial modes remain distributed across the BPM system.

For each mode k, we first evaluate the spatial localisation metric

$$\max_i |u_{i,k}|, \tag{3}$$

that is, the largest absolute contribution of any BPM to the spatial pattern of that mode. Since each column of U has unit Euclidean norm, values close to one indicate that a single BPM contributes almost the entire norm of that mode. We also compute the normalised participation ratio [6], which estimates the effective fraction of BPMs contributing to a mode.

Figure 2 shows this check for the same representative horizontal acquisition. The retained denoising modes are not dominated by individual BPMs: the first mode, for example, has a participation ratio of 66%, indicating a spatial pattern spread over a large fraction of the BPMs. Highly localised modes are instead found outside the retained subspace, where they correspond to BPM-specific residual defects. Their rank and energy help distinguish strong BPM-specific effects, such as offsets, saturation, clipping, or scale errors, from weaker residual defects such as incomplete waveforms, intermittent zeros, and isolated spikes. We inspect all modes and flag the BPM giving the largest contribution when the localisation exceeds a configurable threshold; a localised retained mode is treated as a fault rather than accepted as beam motion. The detection can be applied iteratively to uncover additional localised BPMs.

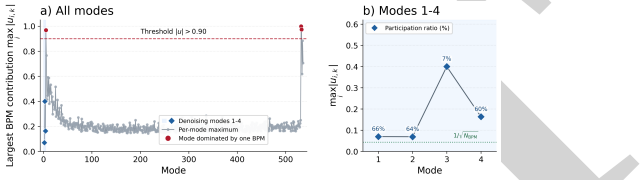

Figure 2: SVD mode localisation for the same acquisition. The retained modes are distributed across many BPMs, as quantified by the participation ratios in the zoom, whereas highly localised BPM-dominated modes appear outside the denoising subspace.

We also apply Isolation Forest [7], treating each BPM waveform as one sample in a high-dimensional space. Before fitting, we centre each waveform and scale it by its Median Absolute Deviation (MAD), which reduces sensitivity to offsets and amplitude differences. Rather than imposing a fixed contamination fraction, we flag BPMs when their anomaly score falls below a fixed global threshold, chosen empirically from representative data and then applied uniformly to all acquisitions.

Finally, frequency-based checks compare the dominant tune of each BPM with the main tune population of the acquisition. We obtain the tune from the Fast Fourier Transform (FFT) peak and refine it with Jacobsen interpolation [8] to achieve sub-bin accuracy. We flag BPMs when their tune deviates significantly from the main distribution.

To avoid unreliable decisions, we apply this tune-based test only when the dominant peak is sufficiently clear relative to the noise floor and the acquisition-wide tune spread remains small. Otherwise, we veto the tune-based test for that acquisition.

IMPLEMENTATION

The monitoring system is implemented as two complementary layers: an online processing stage hosted on the Unified Controls Acquisition and Processing platform (UCAP) [9], and an offline aggregation stage based on the Offline Analysis Framework (OAF) [10, 11]. This separation allows us to perform fast per-acquisition diagnostics during operation while also building higher-level summaries over longer periods. Figure 3 summarises the overall workflow and the interaction between these two layers.

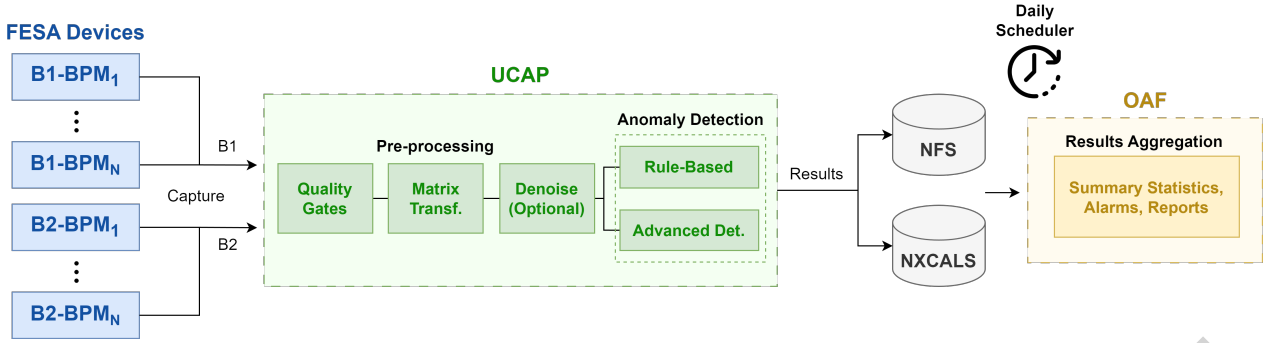

Figure 3: Overview of the automated monitoring pipeline for the LHC BPM system. The diagram shows the flow from online acquisition and validation to anomaly detection, storage, offline aggregation, and reporting.

Online Monitoring Pipeline

The online stage is event-driven and runs on a dedicated Python node. For each acquisition produced by the BPM FESA devices, UCAP retrieves the raw data and metadata before applying the validation, preprocessing, and anomaly detection steps described above.

The results of the online analysis are stored both in the Next CERN Accelerator Logging Service (NXCALS) [12] and in HDF5 files [13]. The HDF5 files are mainly used for debugging and also contain the source data and the original acquisition metadata.

Daily Reports

OAF collects UCAP results for daily analysis, aggregating quality-check failures and anomaly detections into summary tables and statistics at different levels of detail.

OAF also supports reporting, including reports that can be sent by email. These can rank BPMs from most to least problematic and include indicators such as average failure rates, giving operators and experts both an overall view of BPM system health and enough detail for troubleshooting.

RESULTS

We analysed the monitoring outputs from 16 February 2026 to 30 March 2026, covering the LHC BPM recommissioning after the Year-End Technical Stop (YETS). Figure 4 shows the ten BPMs with the highest recurrence scores in the horizontal plane of Beam 1. The rows are ordered by a recurrence score defined as the product of the number of flagged days and the number of different methods that flagged the BPM at least once. The colour scale gives the daily flagged fraction, i.e. the fraction of that day's acquisitions in which the BPM was flagged by at least one method. The anomaly burden is clearly higher during the early recommissioning period and decreases after the first Stable Beams period, although a small subset of BPMs remains persistently problematic; BPM_A.7R3.B1 is the clearest case, being flagged on 30 days and in 74.0% of its acquisitions.

Across the same Beam 1 horizontal dataset, the largest anomaly yield is produced by the zero-value checks (6.20% of BPM observations), followed by spike detection (3.93%) and the SVD-based detector (3.78%), while Isolation Forest (0.33%) and especially the tune-based check (0.12%) are much more selective. These values are method-level yields and are not additive, since the same BPM may be flagged by several tests. A high tune-based rate would be particularly concerning.

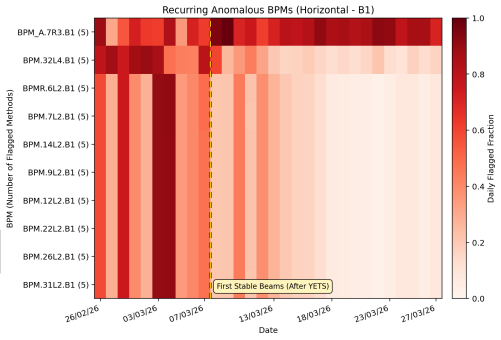

Figure 4: Ten BPMs with the highest recurrence scores in the horizontal plane of Beam 1 during recommissioning after the YETS. Colour shows the daily flagged fraction; rows are ordered by recurrence score. All shown BPMs were flagged by all five methods.

CONCLUSIONS AND OUTLOOK

We have developed and deployed a fully automated monitoring framework for the LHC BPM system. It provides diagnostics, daily aggregation, and reporting, enabling earlier and more consistent identification of malfunctioning BPMs. Initial results from the 2026 recommissioning period show that the framework can reliably identify persistently problematic devices.

Future work will focus on upgrading the OAF daily reports with a proper time-series database, allowing anomaly evolution and long-term trends to be tracked more systematically and making operational follow-up easier. The framework will also be extended with a timing-alignment check for both single-bunch and multi-bunch acquisitions, verifying that the BPM signals are correctly synchronised with the accelerator timing reference despite cable and electronics delays.

REFERENCES

[1] O. Brüning *et al.*, "LHC design report, vol. 1: The LHC main ring", CERN, Geneva, Switzerland, Rep. CERN-2004-003, 2004. doi:10.5170/CERN-2004-003-V-1

[2] E. Calvo Giraldo *et al.*, "The LHC beam position system: Performance during 2010 and outlook for 2011", in *Proc. DIPAC'11*, Hamburg, Germany, May 2011, paper TUPD12, pp. 323–325. https://jacow.org/DIPAC2011/papers/TUPD12.pdf

[3] A. Guerrero, J-J. Gras, J-L. Nougaret, M. Ludwig, M. Arruat, and S. Jackson, "CERN front-end software architecture for accelerator controls", in *Proc. ICALEPCS'03*, Gyeongju, Korea, Oct. 2003, paper WE612, pp. 342–344. https://jacow.org/ica03/papers/WE612.pdf

[4] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 4th ed. Baltimore, MD, USA: Johns Hopkins University Press, 2013, ISBN 978-1-4214-0794-4.

[5] E. Fol, R. Tomás, J. Coello de Portugal, and G. Franchetti, "Detection of faulty beam position monitors using unsupervised learning", *Phys. Rev. Accel. Beams*, vol. 23, no. 10, p. 102805, 2020. doi:10.1103/PhysRevAccelBeams.23.102805

[6] R. J. Bell and P. Dean, "Atomic vibrations in vitreous silica", *Discuss. Faraday Soc.*, vol. 50, pp. 55–61, 1970. doi:10.1039/DF9705000055

[7] F. T. Liu, K. M. Ting, and Z.-H. Zhou, "Isolation forest", in *Proc. 8th IEEE Int. Conf. on Data Mining (ICDM'08)*, Pisa, Italy, Dec. 2008, pp. 413–422. doi:10.1109/ICDM.2008.17

[8] E. Jacobsen and P. Kootsookos, "Fast, accurate frequency estimators", *IEEE Signal Process. Mag.*, vol. 24, no. 3, pp. 123–125, May 2007. doi:10.1109/MSP.2007.361611

[9] L. Cseppentő and M. Büttner, "UCAP: A framework for accelerator controls data processing @ CERN", in *Proc. ICALEPCS'21*, Shanghai, China, Oct. 2021, pp. 230–235. doi:10.18429/JACoW-ICALEPCS2021-MOPV039

[10] B. Kolad, J-J. Gras, S. Jackson, and S. B. Pedersen, "The CERN Beam Instrumentation Group Offline Analysis Framework", in *Proc. IBIC'16*, Barcelona, Spain, Sep. 2016, pp. 449–452. doi:10.18429/JACoW-IBIC2016-TUPG45

[11] A. Samantas, M. Gonzalez-Berges, J-J. Gras, and S. Zanzottera, "Evolution of the CERN Beam Instrumentation Offline Analysis Framework (OAF)", in *Proc. ICALEPCS'21*, Shanghai, China, Oct. 2021, pp. 965–969. doi:10.18429/JACoW-ICALEPCS2021-THPV042

[12] J. P. Wozniak and C. Roderick, "NXCALS - architecture and challenges of the Next CERN Accelerator Logging Service", in *Proc. ICALEPCS'19*, New York, NY, USA, Oct. 2019, pp. 1465–1469. doi:10.18429/JACoW-ICALEPCS2019-WEPHA163

[13] The HDF Group, "Hierarchical Data Format, version 5", https://www.hdfgroup.org/solutions/hdf5/.