BESIII
A mature e^+e^- experiment

JUNO
A next-generation neutrino experiment

Toward Online Full-Histogram Anomaly Detection

for HEP Monitoring Systems

From ML-Based Feasibility Study to Online Workflow Validation

Shuihan Zhang | Institute of High Energy Physics, CAS, China

25th IEEE Real Time Conference
May 25 – 29, 2026

Monitoring Challenges in BESIII and JUNO

Massive real-time histogram streams make online increasing challenging

BESIII

Operating since 2009
BEPCII collider, 2–4.6 GeV

>100 histogram pages
(~1,000 histograms) each shift

Manual inspection
Some histograms not shown
due to human limits

**Relying on human eyes:
easy to miss subtle anomalies**

Massive Histogram Streams

Online Monitoring
& Human Inspection

Operator Burden
Increasing

Need Scalable and
Intelligent Anomaly
Detection

JUNO

20 kt liquid scintillator
~20,000 PMT channels

~40 GB/s
triggered waveform data

Second-level
online response

**High data rate & channel multiplicity:
require real-time, automated monitoring**

Common Need: scalable, intelligent, and reliable **real-time anomaly detection** for stable operation

Why Real-Time Monitoring Matters

Offline analysis is often too late for detector operation.

Offline Limitations

Delayed Response
Problems found hours later.

Potential Data Loss
Anomalies may corrupt data before discovery.

Inefficient Intervention
Late alerts increase workload and investigation time.

Difficult Localization
Harder to trace root causes after conditions change.

Real-Time Advantages

Low-Latency Feedback
Detect issues in seconds–minutes.

Operational Interpretability
Clear, actionable information for operators.

Scalable Monitoring
Handle massive histogram streams from multiple subsystems.

Continuous Protection
Enable timely intervention and stable data taking.

Real-time monitoring delivers the right information to the right people at the right time.

Essential for data quality and reliable detector operation.

Research Strategy

From Feasibility Study to Online Workflow Validation

A progressive approach toward deployable, interpretable, and scalable online monitoring.

Phase 1

JUNO-SPMT Feasibility Study

Can ML methods detect histogram anomalies?

- Large-scale PMT data
- Algorithm evaluation
- Feasibility verification

Phase 2

BESIII Online Workflow Validation

Can the workflow operate reliably online?

- Real-time data access
- Expert-in-the-loop workflow
- Operational validation

Phase 3

Transferable Monitoring Framework

Can the framework scale to future experiments?

- Modular architecture
- Scalability & adaptability
- Reusability & transferability

The goal evolved from algorithm validation to **deployable online monitoring.**

- From model performance to operational value
- From offline analysis to real-time action
- From single-experiment solutions to a generalized framework
- From research prototype to production system

Phase 1 — Feasibility Study on JUNO-SPMT

Why we start Phase 1 and which methods we explore

Motivation

Rule-based monitoring has limitations

Simple rules miss complex and unknown issues.

Manual monitoring is not scalable

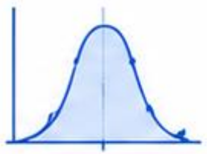
~20,000 histograms to check, human resources are limited.

Goal

Assess whether **ML methods** can detect histogram anomalies **reliably** in **online conditions**.

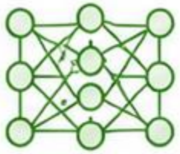
Methods Explored

χ^2 Test



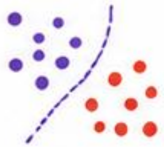
- Traditional statistical method
- Simple and interpretable

Autoencoder (AE)



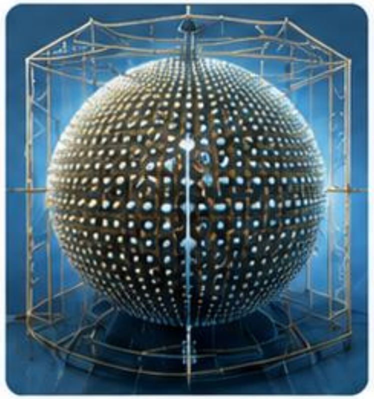
- Learn normal patterns in data
- Detect anomalies via reconstruction error

Supervised Classification



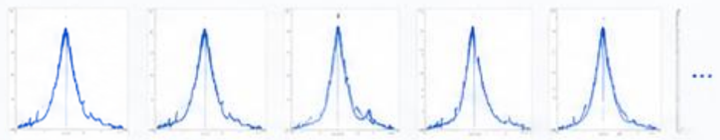
- Learn from labeled normal/abnormal samples
- High precision for known fault types

JUNO-SPMT Online Test Environment



- 25,600 SPMT channels
- ~20,000 histograms
- Online test-like conditions

Histogram Streams



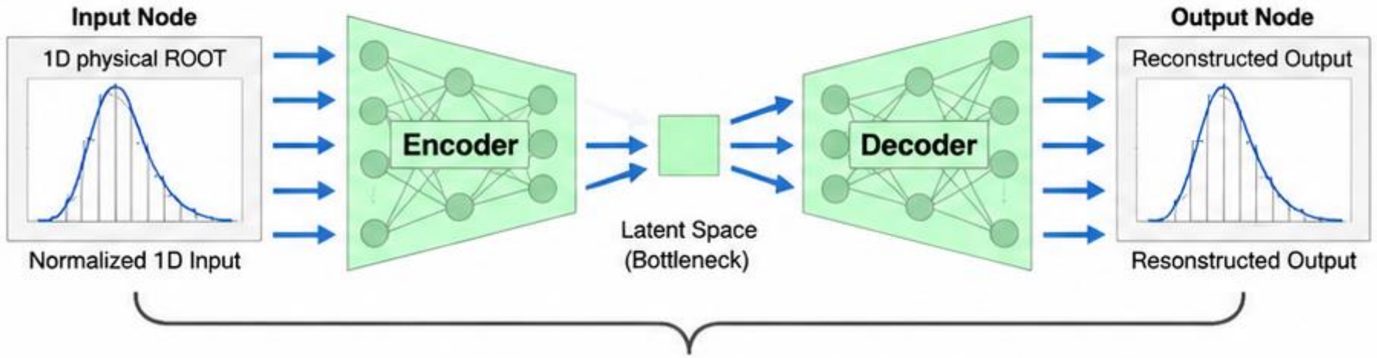
High rate, high volume, and complex detector behaviors

Phase 1 goal: assess feasibility of different ML methods on real JUNO-SPMT online data.

We compare χ^2 test, Autoencoder, and Supervised classification to identify a promising direction for real-time anomaly detection.

Phase 1 — AE-based Real-Time Anomaly Detection

By perfectly modeling the normal distribution, any deviation is mathematically isolated as an anomaly

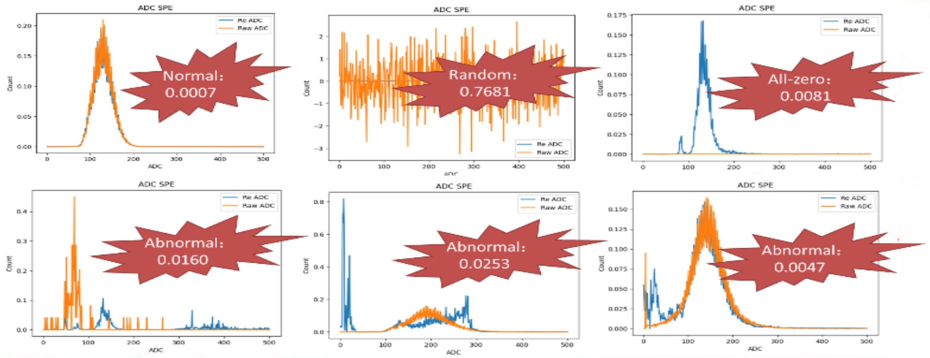


|| **Input - Output** ||^2 (Resonstruction Error)

- **Path A** | Error < Threshold (λ) → Normal
- **Path B** | Error > Threshold (λ) → Anomaly

Test Environment:
Large-scale multichannel test utilizing JUNO-SPMT 25,600 channels and ~20,000 samples

Training & Evaluation:
Grid search for optimal hyperparameters and model architecture



Model Deployment

Feasibility Proven:

- Deployed during JUNO-SPMT **online** tests
- Reduced manual review burden by **85%** & successfully identified **97%** of abnormal histograms

Real Online Deployment
Deployed in JUNO-SPMT online test environment

Manual Review Reduced by 85%
Significantly lower operator workload

97% Anomalies Detected
High detection efficiency in real conditions

Data-driven AE method is both effective and deployable under realistic online monitoring conditions

Phase 1 — Lessons Learned from Algorithm Evaluation

Key problem: High offline accuracy ≠ deployable online monitoring

Algorithm Comparison (per sample, single-threaded)

Model	Time / sample (single-threaded)	Precision	Recall	Strength	Limitation
χ^2	0.035 ms	33%	94%	Simple, efficient	Limited for unknown anomalies
AE	0.0012 ms	64%	99%	Flexible for unknown faults	Threshold-sensitive, needs normal data
Classification	0.0006 ms	99%	99%	High precision for known faults	Requires high-quality labels

No single algorithm is sufficient for real online operation.

A complete online workflow requires the combination of multiple strategies, interpretable rules, and human expertise.

Machine Configuration & Experiment Conditions

- Intel Xeon Gold 6442Y, 2 sockets × 24 cores each, 96 logical cores
- Single-threaded execution
- Dataset: 10,000 SPMT samples

Conclusion: Algorithm performance is necessary but not sufficient. The real challenge is building a reliable, interpretable, and scalable online workflow.

Why BESIII for Online Workflow Validation?

From algorithm evaluation to operational validation in a real detector environment.

System Characteristics

Mature Online Operation
- Long-term stable running
- 24/7 data taking
- Well-understood performance

Established Infrastructure
- Complete monitoring system
- Standardized data flow
- Unified interfaces

Rich Historical Database
- Large data volume
- Various running conditions
- Valuable references

Operational Validation Platform

BESIII is used as a workflow validation platform rather than an isolated ML benchmark.

Algorithm Study → Workflow Validation → Operational Reliability

Why BESIII is Suitable

Expert Knowledge

Realistic Online Environment

Continuous Detector Evolution

The focus shifts from model performance to operational reliability.

- Not just: Can the model detect anomalies?
- But: Can the workflow run reliably online?
- And: Can operators trust and act on results?
- Ultimately: Can the system support long-term operation?

Phase 2 — BESIII Online Workflow Validation

Can the monitoring workflow operate reliably in real online conditions?

Online Monitoring Context

Histogram Production

- ~10^4 online histograms
- Multi-subsystem monitoring
- Periodic updates

Reference Histograms & Metadata

- Dynamic references
- Run & detector metadata

Hierarchical Detection Workflow

Rule-based Screening

- Fast pre-selection
- Existing operational rules
- Low-latency filtering

Operational safety layer

Statistical Methods

- χ^2 test
- Already operational
- Stable & interpretable

ML-based Methods

- AE
- Classification
- Extensible models

Under evaluation

Alert & Expert Validation

- Human-in-the-loop
- Alert verification
- Operational feedback

Why Hybrid Detection?

- False Alarms
- Unstable Models
- Maintenance
- Shifter Trust

Phase 2 validates the end-to-end monitoring workflow in real online conditions.

- Multiple ML methods deployed in parallel
- Real-time data flow and decision making
- Expert feedback and alert validation
- System stability and robustness testing
- From algorithm evaluation to operational validation

Practical Deployment Challenges & Strategy

BESIII | JUNO

Why not deploy aggressive ML directly?

High False Alarms

High false alarm rate can overwhelm operators and mask real issues.

Unstable Thresholds

Thresholds are sensitive to data distribution and hard to maintain.

Detector Drift

Detector conditions change over time, causing model performance degradation.

Operator Trust

Black-box decisions are hard to trust without interpretability and control.

Latency Constraints

Real-time systems require low latency and predictable performance.

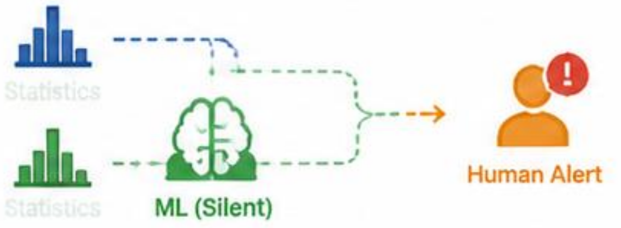
Progressive Deployment Strategy: Ensuring Stable and Safe Online Use

1 Conservative Phase

- Statistical methods + Human confirmation
- Maintain χ^2 safety barrier, collect initial labels

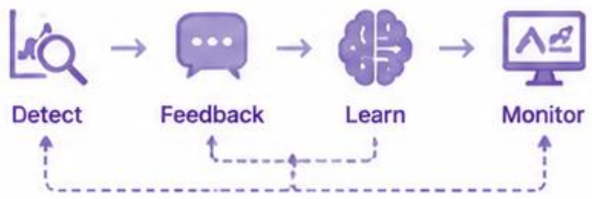
2 Transition Phase

- Statistical + ML dual-track
- ML alerts human if anomalies missed by statistics



3 Mature Phase

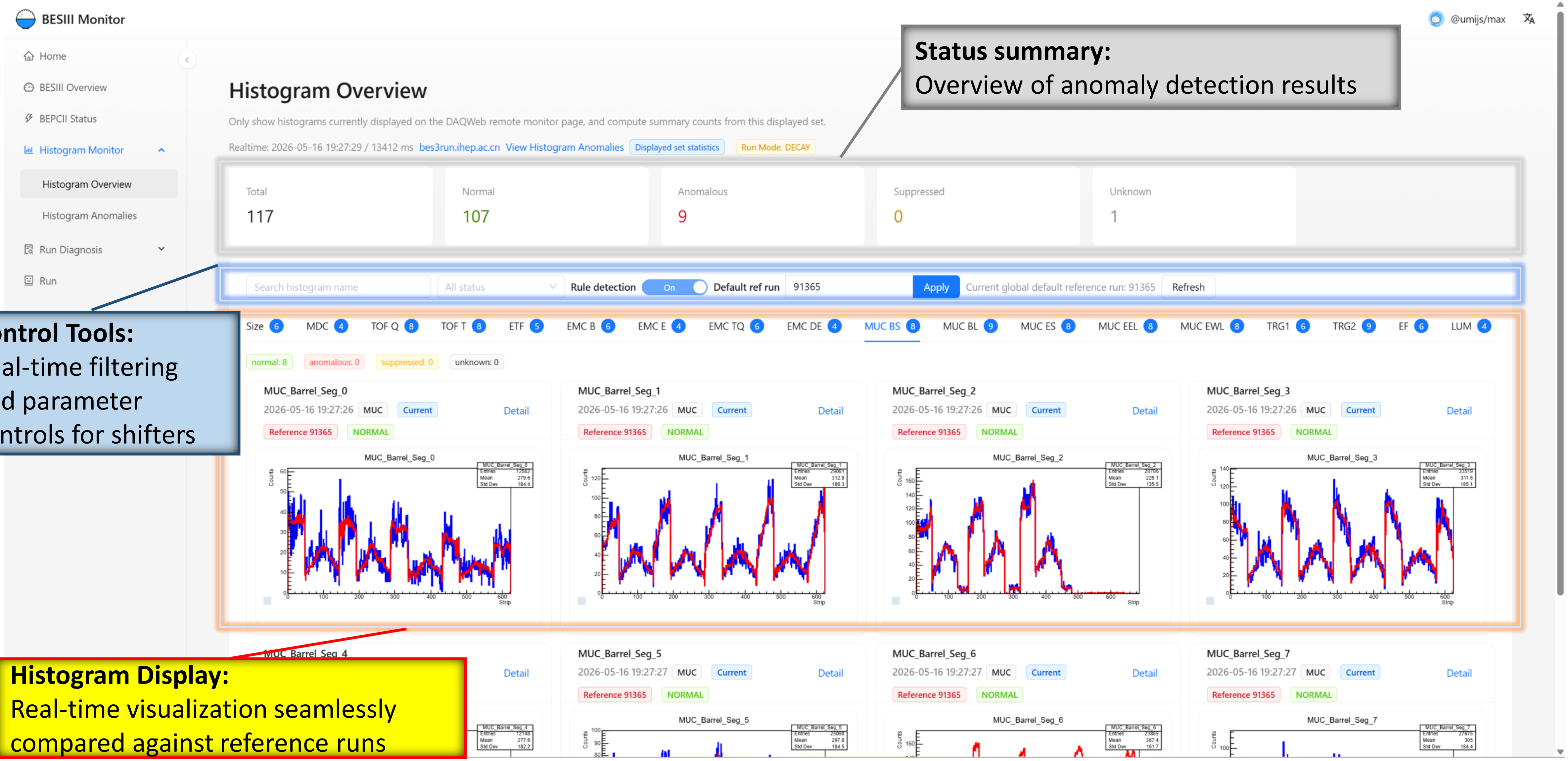
- Full automated "Detect-Feedback-Learn" loop
- Real-time online monitoring



Safe deployment is more important than aggressive automation.

Current Status —— Histogram Overview Interface

The overview page supports real-time monitoring, filtering, and reference-based comparison



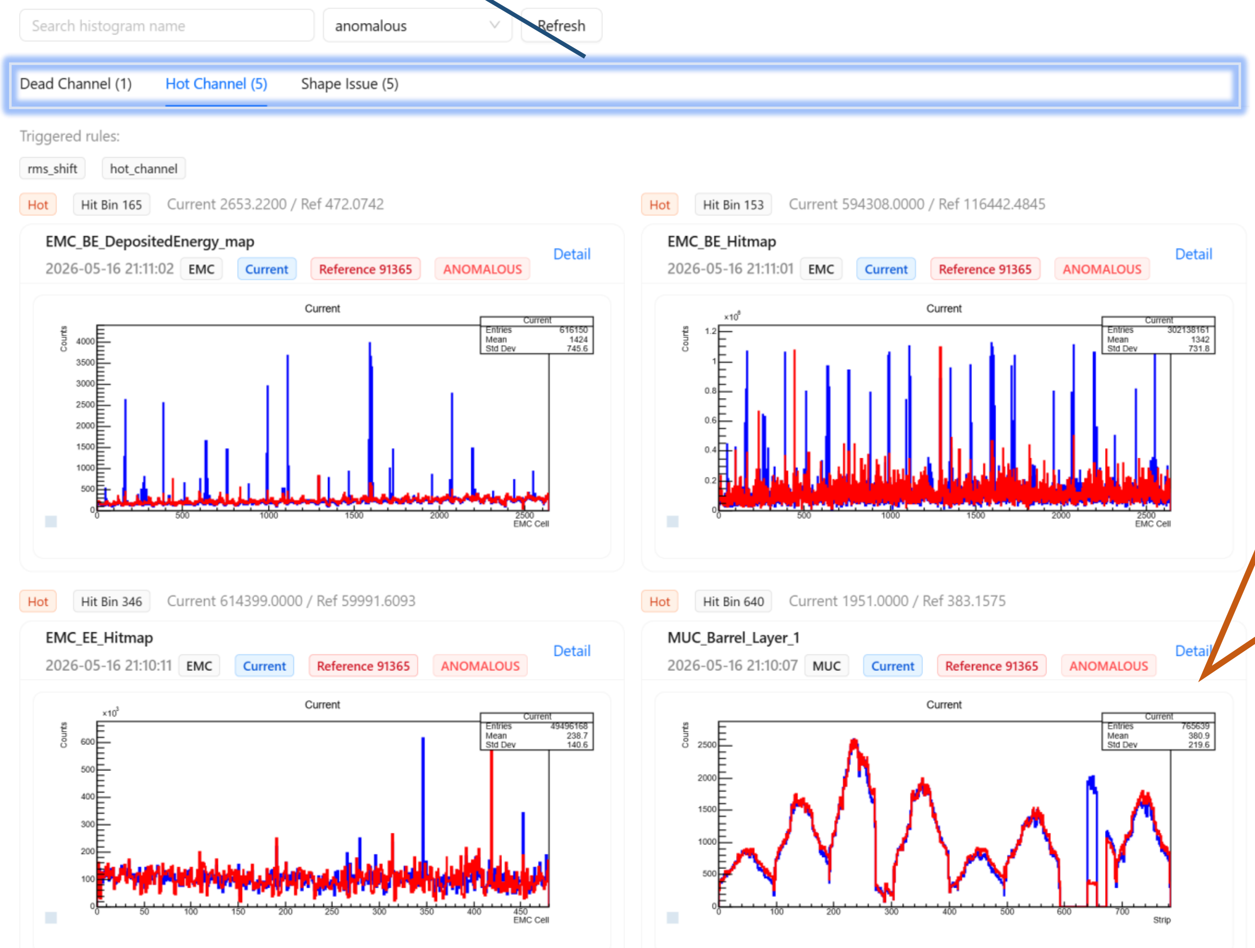
Status summary:
Overview of anomaly detection results

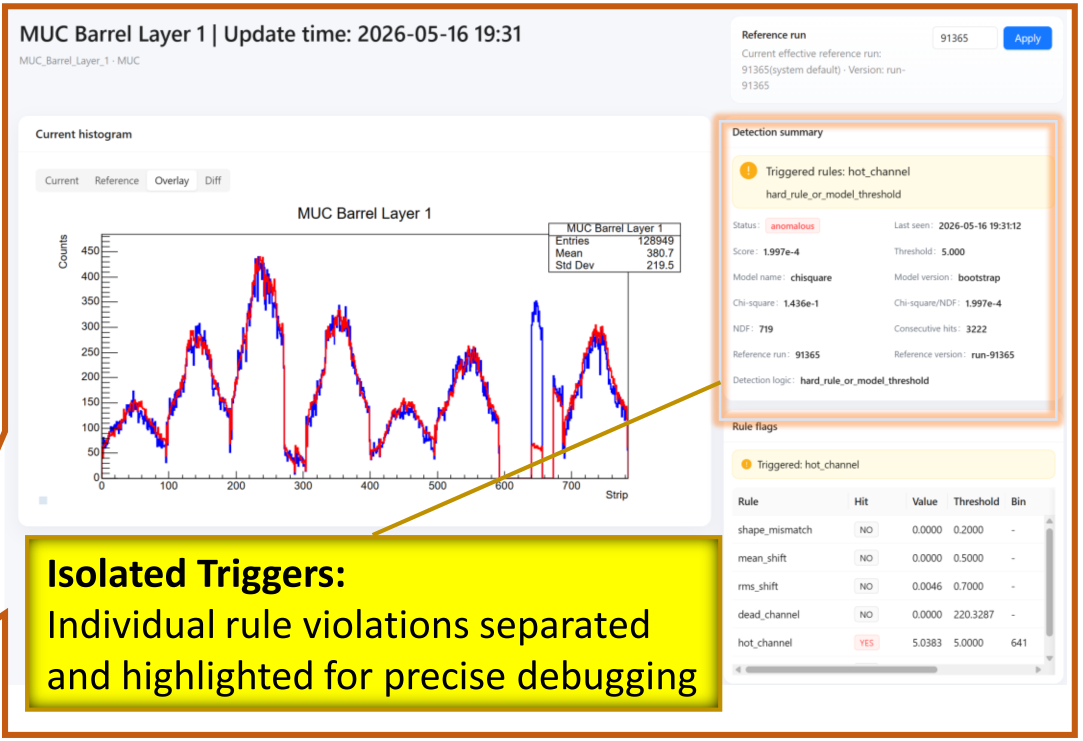
Control Tools:
Real-time filtering and parameter controls for shifters

Histogram Display:
Real-time visualization seamlessly compared against reference runs

Current Status —— Anomaly Display Interface

Categorized Grouping:
Results logically grouped by anomaly type for rapid interpretation





Isolated Triggers:
Individual rule violations separated and highlighted for precise debugging

Designed entirely to accelerate expert diagnosis and reduce cognitive load

Closed-Loop Evolution Under Detector Drift

How can the monitoring workflow remain reliable under long-term detector evolution?

1 Online Monitoring

- Real-time anomaly detection
- Continuous detector operation
- Low-latency monitoring workflow

2 Expert Validation

- Operators verify alarms
- False positives identified
- Expert-in-the-loop validation

3 Verified Anomaly Collection

- Store confirmed anomaly cases
- Build operational knowledge base
- Accumulate validated monitoring data

4 Threshold / Reference Update

- Detector conditions evolve over time
- Static thresholds become unreliable
- Reference histograms require updates

5 Model Recalibration

- Periodic workflow adjustment
- Incremental model recalibration
- Maintain long-term monitoring stability

Continuous adaptation under detector evolution

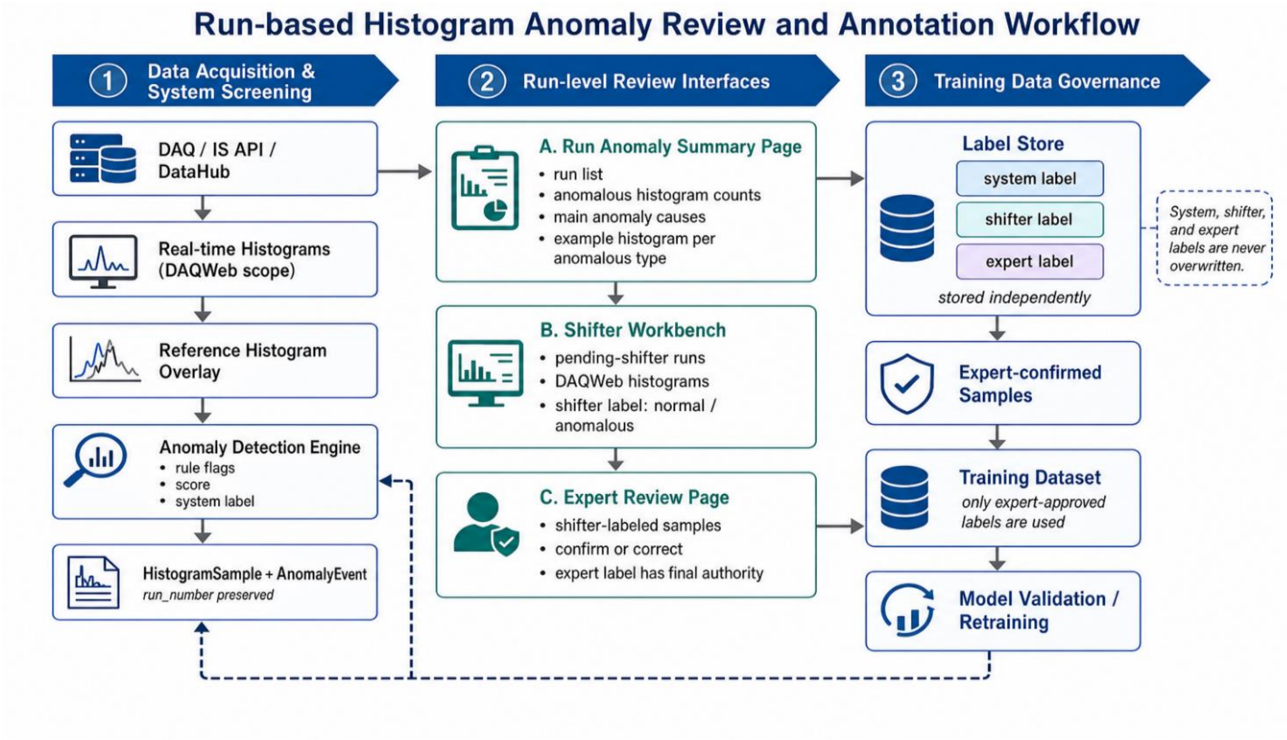
Static monitoring strategies degrade under evolving detector conditions.

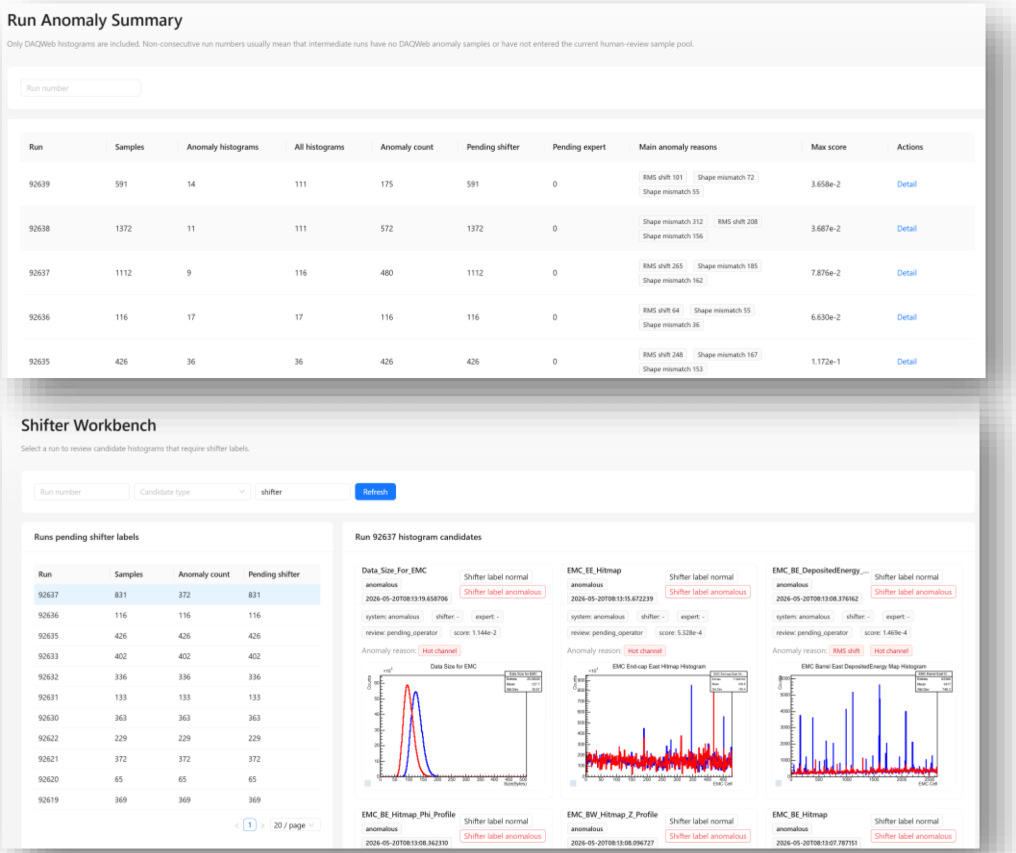
- Expert-in-the-loop validation
- Operational knowledge accumulation
- Adaptive thresholds & references
- Long-term stability & reliability

Current Status — Expert Feedback Integration

Closing the loop: from online detection to model improvement

Online detection → **Expert review** → **Sample accumulation** → **Model/Reference Update**





An efficient and scalable approach to accumulate **high-quality** training datasets with **minimal expert effort**, enabling **continuous model improvement** and **reliable long-term online operation**.

From Feasibility Study to Deployable Workflow

*Bridging validated methods into a reliable, **expert-driven online monitoring system***

What We Have Achieved

- Full-histogram monitoring framework ✓
- χ^2 integrated into BESIII workflow ✓
- Real-time visualization interface ✓
- Expert feedback integration ✓
- Reference-based comparison ✓

Current Operational Status

The system is running a **hybrid detection mechanism** based on rule-based screening and χ^2 statistical models, while gradually accumulating **expert-labeled** high-quality samples.

Online Data Streams → Hybrid Detection (Rule-based + χ^2) → Expert Validation (Human-in-the-loop) → Labeled Samples Accumulation

Hybrid Detection Mechanism

Rule-based Screening + χ^2 Statistical Model

- Fast pre-selection
- Statistical validation
- Low-latency & interpretable

Expert-Labeled Dataset

- Experts label real anomalies
- Build high-quality dataset
- Support future model updates
- Ensure long-term operation

Key Challenges Ahead

- Detector drift adaptation
- Threshold robustness
- False alarm management
- Scalability under long-term operation
- Operator trust and usability

The primary challenge is no longer **anomaly detection itself**, but **reliable long-term online operation.**

Operational Reliability | Continuous Adaptation | Expert-in-the-loop Validation | Toward Deployable Online Monitoring

Summary and Outlook

Building toward a transferable intelligent monitoring framework

Completed

100% COMPLETE

Feasibility proven:
Verified the core capability of data-driven anomaly detection using ML-based methods

- AE-based anomaly detection validated
- Deployed in online test environment
- 95%+ detection performance achieved
- Significant reduction in manual workload
- Operational feasibility demonstrated

In Progress

50% IN PROGRESS

BESIII workflow under validation:
Active development of an online full-histogram anomaly detection system prioritizing workflow usability and robustness

- Hybrid (Rule-based + χ^2) mechanism online
- Real-time visualization & expert interfaces
- Reference-based comparison
- Expert feedback and sample accumulation
- Workflow integration and optimization

Future Outlook

0% FUTURE

Transferable framework next:
Build a robust and deployable online monitoring framework that can later be extended to JUNO and other future experiments

- Long-term stability and robustness
- Scalability to larger datasets & experiments
- Adaptive to detector evolution and drift
- Improved operator trust and usability
- Generalization to multi-experiment scenarios

Thanks for your attention