AI-BASED DIAGNOSTICS FOR THE CRYOGENIC AND RF SYSTEMS OF THE SPIRAL2 SUPERCONDUCTING LINAC

C. Lassalle[1,2], A. Ghribi[1], M. Di Giacomo[1], F. Bouly[2], P. Bonnay[3]
[1]GANIL, Caen, France
[2]Univ. Grenoble Alpes, CNRS/IN2P3, Grenoble INP, LPSC, Grenoble, France
[3]CEA/IRIG/DSBT/LEA, Grenoble, France

Abstract

The SPIRAL2 superconducting LINAC at GANIL operates 26 quarter-wave resonator cavities whose online diagnostics currently rely on physics-based models limited to single operating points. This paper presents two complementary AI-based diagnostic tools: (i) neural network heat-load virtual observers that estimate the cavity thermal dissipation — a proxy for the intrinsic quality factor Q_0 — from cryogenic process signals, with prediction errors predominantly in $[-2, +1]$ W@4.2 K for loads up to 20 W@4.2 K; and (ii) a machine-learning pipeline meant to detecting anomalies in LLRF data, predicting alarms before they fire, and classifying fault subtypes within the cavity-quench category ($F_1 = 92\%$). This paper presents a state of progress on these two applications.

INTRODUCTION

SPIRAL2 (Système de Production d'Ions Radioactifs Accélérés en Ligne, 2nd generation) is a multi-beam LINAC at GANIL (Caen, France) designed to accelerate protons, deuterons, and heavy ions for nuclear-physics experiments. Its superconducting section comprises 26 quarter-wave resonator (QWR) cavities at 88.05 MHz arranged in two families: 12 type-A cavities ($\beta_{opt} = 0.07$, cryomodules CMA01–CMA12) and 12 type-B cavities ($\beta_{opt} = 0.12$, cryomodules CMB01–CMB07). Each cryomodule houses one or two cavities immersed in a liquid helium bath at 4.2 K; the naming convention is `CMA`*`XX`*`-CAV`*`Y`* and `CMB`*`XX`*`-CAV`*`Y`*, where *XX* is the module index and *Y* the cavity index within the module.

Reliable operation requires continuous monitoring of both cryogenic and radio-frequency (RF) systems. Operational cryogenic diagnostics currently rely on physics-based state observers that are limited to a single linearization point and would require frequent recalibration. On the RF side, fault events are recorded as post-mortem files by the Low-Level RF (LLRF) control system; identifying and anticipating these faults manually is time-consuming. This paper describes two complementary AI applications developed for SPIRAL2. (i) neural network heat-load observers that estimate, in real time, the thermal dissipation in the helium bath as a *virtual* quality-factor Q_0 observable; and (ii) a machine-learning pipeline for the detection, early prediction, and subtype classification of the LLRF anomalies.

Table 1: Input Observable Selections Evaluated for Heat-load Estimation.

Selection	P	L	V_{in}	V_{out}	T_{in}	T_{out}
1	✓	✓	✓	✓	✓	✓
2	✓			✓		
3					✓	✓
4						
5			✓	✓		
6				✓		

HEAT LOAD VIRTUAL OBSERVER

Motivation and Problem Formulation

The dynamic heat load $\dot{Q}$ dissipated by the cavity walls into the helium bath is a key indicator of cavity performance and a proxy for the intrinsic quality factor Q_0. The intrinsic quality factor is defined as

$$Q_0 = \omega_0 \frac{W}{P_{dis}}, \qquad \dot{Q} \equiv P_{dis}, \tag{1}$$

where $\omega_0 = 2\pi \times 88.05$ MHz is the resonant angular frequency, W the stored electromagnetic energy, and $P_{dis} = \dot{Q}$ the power dissipated in the cavity walls and extracted by the helium bath. A continuous, real-time estimate of $\dot{Q}$ is therefore equivalent to a virtual Q_0 observer. Existing state observers (Luenberger, extended Kalman filter) provide estimates but are valid only near a linearisation point and are sensitive to sensor calibration drift. The extended Kalman filter is, in addition, too large to run on industrial PLCs used at GANIL [1]. Neural networks are trained in a supervised fashion using cryogenic process signals as inputs, with the regression target derived from dedicated heat load variation campaigns thanks to embedded resistive heaters in the cryomodules.

Dataset and Observable Selection

Measurements were taken during dedicated heat-load campaigns on cryomodules in 2022 and 2023. Six cryogenic observables are considered (Table 1): helium pressure (P), inlet/outlet temperatures (T_{in}, T_{out}), liquid helium level in the cryostat (L), and inlet/outlet valve positions (V_{in}, V_{out}). Six input configurations (selections 1–6, from all six signals down to T_{out} alone) are evaluated to assess model robustness to missing or unreliable sensors.

Model Architectures

Four neural network architectures have been trained and compared: **MLP** (multi-layer perceptron, 64–32–8–1 dense layers), **1D-CNN** (two Conv1D layers of 128 filters, kernel 2), **LSTM** (50 units, Seq2Vec), **LSTM+Att** (128 units with Luong attention), and a **Stacking** ensemble that combined all four base models. All models are trained with Adam optimiser, MSE loss, early stopping (patience 10) over up to 10 000 epochs with a batch size of 256. A one-hot cryomodule identifier is optionally appended to the inputs to study multi-cryomodule generalization.

Results

Same-year evaluation (2022) Models trained on CMA08 2022 data and tested on held-out 2022 sequences achieve prediction error standard deviations below 0.5 W for selection 1 (all six observables) across most of the cryomodules. On CMA08 data, the Stacking model reaches the lowest bias ($\mu \approx -0.01$ W) and the lowest variance ($\sigma \approx 0.35$ W).

Cross-year evaluation (2022→2023) Selection 1 degrades significantly ($|\epsilon| \lesssim 10$ W) due to calibration drift between the campaigns. Selection 4 (outlet temperature only) maintains satisfactory performance with errors predominantly in $[-2, +1]$ W, demonstrating robustness to changes in sensor calibration. This result motivates the use of reduced observable sets for the deployment.

Multi-cryomodule generalisation LSTM models with a cryomodule identifier generalise across cavities: interquartile ranges of prediction error of 0.32 W on CMA08 and 0.36 W on CMA12, with a slight positive bias on CMA12 ($\mu = 0.15$W).

LLRF ANOMALY DETECTION AND CLASSIFICATION

Data Acquisition and Post-mortem Files

Each SPIRAL2 cavity LLRF system is equipped with a circular buffer that saves 100 096-sample post-mortem files at 88.05 kHz ($\Delta t \approx 11.4$ μs) whenever a hardware alarm is triggered or an operator requests a snapshot. Seven fault types are monitored: electron pick-up threshold exceeded, fast external cut required by the circulator arc detection system, missing RF permission, vacuum threshold exceeded, cavity quench/breakdown, RF protection threshold exceeded and RF signal out of tolerance. Events recorded between 2019 and 2025 from all 26 cavities in 19 cryomodules are used. Filters (LLRF feedback loop to control accelerating voltage in amplitude and phase, beam loading compensation activated, correct loop coefficient KPI ≥ 10, accelerating gradient $\overline{U_{\mathrm{cav}}}|_{t<t_0} \geq 0.1$ MV/m) reduce the raw pool to 3208 (924 normal + 2284 faulty) events. A key finding of this work is that the ALM header field is unreliable as a ground-truth label. Cross-checking it against the seven per-sample binary fault indicators revealed discrepancies in 36% of the files (after filtering). Therefore, the ground-truth labels are derived from the latter half of these waveform-level signals.

Pipeline Overview

The analysis pipeline includes three tasks:

1. **Anomaly detection**: does an event contain an anomaly?
2. **Early detection**: can an alarm be predicted *before* it fires, using only the pre-trigger data?
3. **Subtype classification**: which sub-class does a fault belong to?

Task 1 — Anomaly Detection

Both unsupervised and contrastive approaches are evaluated using the five most informative signals (I/Q components of forward power, amplitude of reflected power, phase difference between cavity and forward signals, I/Q-modulator command), scaled by z-score or min-max.

Autoencoder: LSTM-based encoder–decoder (64–32 units, 16D latent space) trained on normal events; anomalies are flagged by reconstruction error exceeding a threshold, which is optimised to maximise the validation F_1 score.

Contrastive learning: five triplet-loss variants (including batch-hard and proxy-based methods) trained on a 1D-CNN architecture (64–128–64 filters, global pooling, 16D latent space).

10-fold cross-validation with Nadeau–Bengio statistical testing shows that all models (except the autoencoder with z-score local scaling, which destroys amplitude information) achieved statistically comparable F_1 scores. No single approach is significantly superior for anomaly detection on this dataset.

Task 2 — Early Detection

For the first nine of the ten temporal sections from the pre-trigger window (the first half of each post-mortem file), each signal is scaled independently across this full duration prior to sectioning, aiming to classify the event as normal or faulty before the alarm fires. Statistical, spectral, dynamical, and entropy-based features extracted per section using the `tsfresh` library undergo robust scaling to mitigate outliers before being filtered by low-variance removal, Pearson correlation de-duplication, and SelectKBest. Four neural architectures (MLP, 1D-CNN, LSTM, GRU) and two classical models (XGBoost, Random Forest) were compared with both z-score and min-max scaling. Monte Carlo Dropout is used in neural networks to provide epistemic uncertainty quantification, alongside each prediction.

All models achieved F_1 scores in the range 0.78–0.81 (Table 2). The best-performing model is LSTM with z-score scaling ($F_1 = 80.5 \pm 3.0\%$, AUC $= 87.7 \pm 1.6\%$). Although Nadeau–Bengio tests show no statistically significant difference between models in terms of F_1 score, neural networks are ultimately preferred for safety-critical deployment due to their capacity to quantify this epistemic uncertainty.

Table 2: Top-5 Early-detection Models (10-fold CV). Best Model in Bold.

Norm.	Model	F_1 (%)	Prec. (%)	Rec. (%)	AUC (%)
z-score	**LSTM**	**80.5±3.0**	**95.6**	**69.7**	**87.7**
MinMax	XGBoost	80.0±4.0	94.4	69.7	86.2
z-score	Random Forest	79.9±4.1	98.8	67.2	85.6
z-score	GRU	79.6±3.3	94.5	68.8	86.6
z-score	CNN-1D	79.5±3.5	97.4	67.4	87.2

Task 3 — Fault Subtype Classification

Within the "cavity quench / breakdown" fault class, visual inspection of post-mortem waveforms revealed two distinct profiles (Fig. 1): **e-quench** [2] (electronic quench, discharge < 11 μs) and **false alarm** (nominal-duration discharge). ts-fresh features of the accelerating gradient signal are reduced to 2D by t-SNE (using OpenTSNE for transferable transforms) and classified using KNN. On a held-out test set of 11 events, KNN achieves accuracy 90.9%, precision 100.0%, recall 85.7%, F_1 92.3%. The method is encouraging despite the small dataset; the single misclassified event corresponds to an e-quench whose t-SNE projection falls in false alarm clusters.

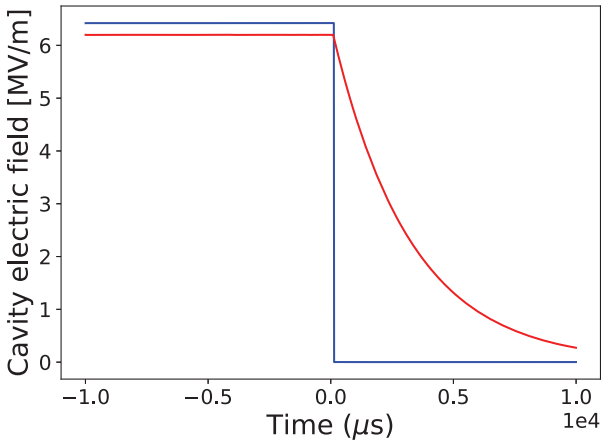

Figure 1: Cavity quench / breakdown profiles.

NEXT STEPS

Cryogenic side Two Q_0 virtual sensors will be planned for shadow-mode deployment in the GANIL control system: one based on Selection 1 (all six observables — wider dynamic range but sensitive to valve-calibration drift) and one based on Selection 4 (outlet temperature only — immune to valve-calibration errors but saturating beyond ≈20 W). Operating both in parallel allows cross-validation during normal exploitation. A further extension to multimodal observers incorporating RF diagnostics, beam-position monitors (BPMs), and beam-loss monitors (BLMs) is envisaged to capture inter-subsystem correlations not visible from cryogenic signals alone.

RF side Manual expert labeling of the post-mortem archive will provide a reliable ground truth for training supervised precursor-detection models — extending the early-detection pipeline beyond the waveform-level fault register. The long-term objective is a semi-continuous online analysis of LLRF waveform buffers during regular SPIRAL2 operation, enabling proactive intervention before beam trips occur.

ACKNOWLEDGEMENTS

This work has been funded by "Region Normandie" as well as the city of Caen, CNRS and CEA.

CONCLUSION

Two complementary AI diagnostic tools were developed for the SPIRAL2 LINAC: (i) neural-network heat-load observers that estimate thermal dissipation from cryogenic process signals, with errors predominantly within $[-2, +1]$ W for heat loads up to 20 W, and with demonstrated cross-year and cross-cryomodule robustness when using the outlet temperature alone; and (ii) a LLRF anomaly pipeline covering unsupervised and contrastive anomaly detection, early alarm prediction (best $F_1 \approx 80.5$ %) with uncertainty quantification via MC Dropout, and subtype discrimination of cavity-quench events ($F_1 = 92$ %). Both tools will be considered for integration into the GANIL control system following the outlined next steps.

REFERENCES

[1] A. Vassal, A. Ghribi, F. Millet, F. Bonne, P. Bonnay, and P.-E. Bernaudin, "Spiral2 cryomodule models: a gateway to process control and machine learning", *Frontiers in Physics*, vol. 10, 2022. doi:10.3389/fphy.2022.875464

[2] T. Powers *et al.*, "Investigations of Arcing Phenomena in the Region Near CEBAF RF Windows at 2K", in *Proc. SRF'93*, Newport News, VA, USA, Oct. 1993, paper SRF93K01. https://jacow.org/SRF93/papers/SRF93K01.pdf