Data-driven gradient optimization for field emission management in a superconducting radio-frequency linac

S. Goldenberg,[1,*] K. Ahammed,[2] A. Carpenter,[1] J. Li,[2] R. Suleiman,[1] and C. Tennant[1]

[1]*Thomas Jefferson National Accelerator Facility, Newport News, Virginia 23606, USA*
[2]*Old Dominion University, Norfolk, Virginia 23529, USA*

(Received 17 November 2024; accepted 11 March 2025; published 21 April 2025)

Field emission can cause significant problems in superconducting radio-frequency linear accelerators (linacs). When cavity gradients are pushed higher, radiation levels within the linacs may rise exponentially, causing degradation of many nearby systems. This research aims to utilize machine learning with uncertainty quantification to predict radiation levels at multiple locations throughout the linacs and ultimately optimize cavity gradients to reduce field emission-induced radiation while maintaining the total linac energy gain necessary for the experimental physics program. The optimized solutions show over 40% reductions for both neutron and gamma radiation from the standard operational settings.

DOI: 10.1103/PhysRevAccelBeams.28.044603

I. INTRODUCTION

Jefferson Lab's Continuous Electron Beam Accelerator Facility (CEBAF) [1] relies on two superconducting radio-frequency linear accelerators (SRF linacs) to deliver high-energy electron beams to nuclear physics experiments in the four experimental halls [2]. An integral part of these linacs is cryomodules which contain multiple SRF cavities. These SRF cavities provide the main accelerating gradients to the electron beam and currently produce the 12 GeV beam necessary for scientific discovery.

A. Field emission

When SRF cavities are operated at high radio-frequency (RF) gradients, it is possible for electrons to be emitted from the cavity walls, known as field emission (FE) [3,4]. The basic physics of FE in SRF cavities is generally understood [5–7], and modern cavity surface processing and assembly techniques control FE well prior to installation [8]. However, during routine operation, degradation of the FE onset (i.e., the gradient at which the first FE-induced gammas are detected) has been observed. It is believed this degradation is caused by particulates and hydrocarbons entering the cavity, contaminating its surface, and introducing new field emitters [9].

Field emission is one of the most detrimental problems for CEBAF linacs [10–13]. When field emitted electrons hit the cavity walls, cryogenic heat loads increase. If those electrons are accelerated beyond the source cavity, their energies can exceed neutron production and material activation thresholds, producing radiation and damaging beamline components. Prior to the 12 GeV upgrade to CEBAF, activation levels detected were typically low enough to not require "radiation area" postings. These postings are required for areas where the whole-body radiation dose is >0.005 rem/h. However, since the energy upgrade, CEBAF has suffered from significant FE-induced radiation. With RF on, dose rates observed at 30 cm from the beamline are as high as 10 and 100 rem/h for neutron and gamma radiation, respectively. This level of radiation causes significant damage to beamline components, including vacuum valves, magnets, and cables for beam position monitors and ion pumps. Replacing these components can require significant resources. Worse, portions of both linacs are considered "radiation areas" for days or even weeks into scheduled downtime, limiting maintenance activities to repair and replace components, and increasing radiation doses for personnel.

Most vacuum valves in both linacs are now damaged causing them to leak. Radiation damaged valves, where the valves' seals become brittle might be one of the most detrimental problems preventing higher cavity gradients. Particulates from the damaged seals propagate through the beamline, entering the cryomodules. This contamination introduces FE even at very low gradients and requires the removal and refurbishment of the cryomodules. The leaky valves prevent vacuum isolation of the cryomodules, complicating the maintenance activities and the removal and installation of cryomodules. Pending available funds, CEBAF has started on a multi-year project to replace Viton vacuum valves in the linacs with all-metal valves [14].

*Contact author: sgolden@jlab.org

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

The refurbishment is further complicated by the activation of many of the cryomodule and beamline components due to FE.

Determining FE-source cavities is challenging for a number of reasons. At CEBAF, existing radiation monitors were designed to detect high bursts of gamma radiation due to beam loss events and quickly shut down the beam. Until recently, there was no readily available way to continuously monitor the relatively low level of radiation produced by FE in real time within the accelerator. Newly designed neutron and gamma radiation dose rate detectors (NDX) [15,16] were successfully installed and provide real-time radiation measurements in 13 locations along both linacs at CEBAF. However, detecting FE at the cavity level remains a challenge as each set of 13 detectors covers a total of 200 cavities. Additionally, field emitted electrons may be captured and accelerated from the source cavity, causing radiation far away from the source. Tests at CEBAF have shown that FE can be accelerated in either direction in the linac, with one instance passing through 14 cryomodules (>100 m) [13].

Figure 1 shows one of the NDX detectors positioned between two cryomodules at a distance of about 0.5 m from the beamline. At the center of the detector, two identical ionization chambers [17] filled with different helium gas isotopes (^{3}He and ^{4}He) serve as the sensors in the detector. The chambers' output currents are measured by highly sensitive electrometers [18]. This configuration enables accurate measurement of the two different radiation dose rates.

A diagram of the south linac including the locations of the 25 cryomodules (containing eight cavities each) and 13 detectors is given in Fig. 2. These detector locations correspond to cryomodules with expected FE activity. Cryomodule styles are typically denoted based on the

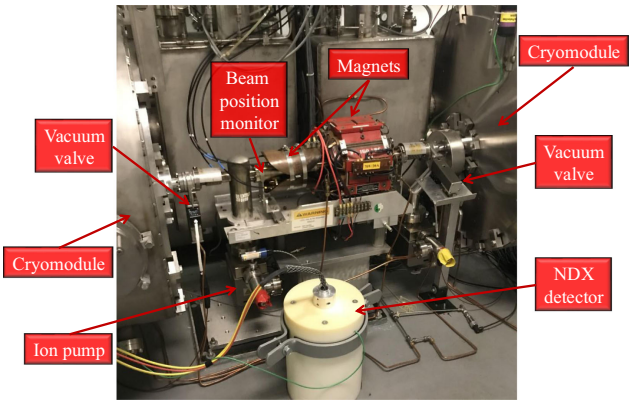


FIG. 1. NDX 2L25 detector positioned next to the beamline connecting two cryomodules at a distance of about 0.5 m. Neighboring cryomodules are separated by a 1.2 m long beamline. The white polyethylene outer layer of the detector is yellowing due to FE radiation.

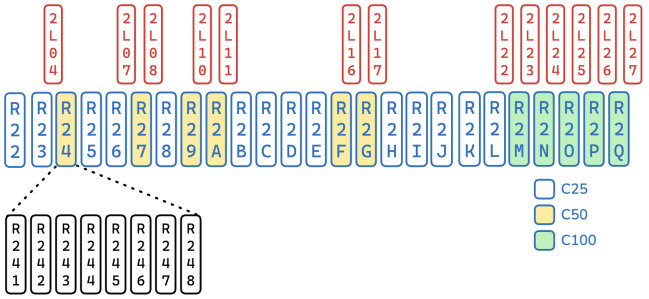


FIG. 2. Layout of the CEBAF south linac. Cryomodules are outlined in blue, detectors in red. There are eight individual cavities per cryomodule (shown in black). The south linac contains three cryomodule types indicated by their background color: C25, C50, and C100, where the numbers indicate the nominal energy gain in MeV expected from each type.

expected total energy gain. A C25 cryomodule, for example, would be expected to produce around 25 MeV of energy gain. In CEBAF's south linac, detectors have been placed near the C50 and C100 cryomodules, which produce their highest gradients and are therefore more likely to produce FE.

B. Methodology

If an FE-source cavity can be determined, radiation levels can be minimized by lowering its gradient and redistributing the gradient to other FE-free cavities. A proof-of-principle test of this gradient redistribution for reducing FE radiation was demonstrated in the injector by slowly increasing the gradient for each cavity individually to determine its contribution to the radiation levels [13]. Individual testing like this is not feasible as a long-term solution, as it requires a significant investment in time. Additionally, it is very likely for field emitters to be enhanced and new field emitters to be excited, causing previous FE onset values to change.

These changing onset values, as well as changes in field emission intensity, make the task of modeling radiation responses particularly challenging. Furthermore, a cavity's contribution to radiation levels is rarely isolated, as interactions with nearby cavity gradients can alter radiation response by accelerating FE upstream or downstream. These complex interactions and the high dimensional input space of the problem make it an ideal problem for machine learning (ML). Our goal is to produce an ML surrogate model that predicts radiation levels based on cavity gradients. This model will power an offline optimization process for redistributing gradients to reduce radiation levels while maintaining linac energy gain. This will be the first time ML is used to mitigate the problem of field emission in an SRF linac with a great number of cavities. Other ML efforts at CEBAF to improve SRF operation are described in [19–21], while work at other SRF facilities can be found in [22–24].

II. DATA

Our models utilize two sources of data, the RF control system and the NDX system. From the RF system, we use the measured cavity gradient (MV/m) for model input. Other miscellaneous values that describe the broader machine state are also used to validate our data collection. The cavity gradients are near instantaneous measurements and updated at a rate of 1 Hz. While CEBAF has multiple styles of cryomodules and RF controls, the data we use are common across all SRF cavities within a linac.

NDX detectors measure both neutron and gamma radiation dose rates over a specified time horizon. We choose to measure dose rates over one-second windows as it matches the RF gradient update rate, provides a reasonable trade-off between response time and measurement variability, and matches the standard configuration used during beam operations. This system provides the 26 radiation dose rates (neutron and gamma for each of the 13 detectors) used as targets when training our models.

Since our goal is to minimize radiation during experimental operations, we require that our data be as reflective of those machine states as possible. This means we want all RF cavities to be properly phased, to have gradient settings similar to their operational gradients, and with beam transport magnets set to match operational optics for those gradients. However, this does not require all cavities to be operational and running at their design limits. With 200 SRF cavities per linac, beam operations routinely occur with some degraded or disabled cavities. Additionally, some cavities are intentionally run below their design limits in order to reduce the frequency of certain RF trip types. This however gives us headroom for rebalancing gradient throughout a linac both during data collection and beam operations.

A. Data collection

For this work, we developed software that actively controls a linac to investigate the radiation response to changing SRF cavity gradients. Raw data was collected by first configuring a linac to operational standards with RF on and beam absent. Collecting data without the beam should not bias our results, as under normal operation, the electron beam does not influence radiation levels in the linac unless there is a beam loss event. Beam loss events do not produce the constant radiation responses that require mitigation through gradient optimization, as they result in a large spike in radiation levels for a very short time, generally less than 50 μs, and cause the beam to shut down.

Our software scans a variety of gradient settings for individual cavities or for combinations of cavities in the neighborhood of their typical set points. Individual cavity scans produce results that are easy for human analysis and can clearly identify which cavities are active field emitters. However, individual cavity scan data lacks information on radiation responses when many cavities are changed at once. Additionally, individually scanning 200 separate cavities is a time-intensive endeavor. On the other hand, combination scans demonstrated two significant advantages in our testing. First, combination scan data more closely approximates solutions produced by our optimizers, as they alter multiple cavity gradients to reduce radiation and rebalance energy gain throughout the linac. Second, due to sparse opportunities for data collection, it was important to efficiently sample a wide variety of gradient distributions. Adequate coverage of simultaneous gradient changes is critical for training models that perform well in the optimization phase.

Data collection is challenging in practice due to two main factors. First, our sampling techniques often exercised the RF systems harder than steady-state beam operations, and second, the 200 cavity gradients created an enormous sample space, which required being strategic with our data collection to make the best use of the allocated time. Ensuring sufficient coverage in this environment within our 0.5–4 h window required both forethought and flexibility.

Most data collection used combination scans with gradient offsets ranging from approximately -4 to $+0.5$ MV/m. The exact distributions varied, but roughly 20% of cavities experienced large gradient reductions, and 10% of cavities remained completely unchanged. Some scans only introduced changes to every fifth cryomodule to reduce correlations between neighboring cavities. Smaller individual cavity scans of suspected field emitters would be performed if time permitted.

Early data collection efforts made all gradient changes simultaneously and in parallel, which caused unwanted correlations in our training data. This was improved for our final datasets by introducing a random delay before changing a cavity gradient. These delays were sampled from a uniform $\mathcal{U}(0, x)$ distribution, where x ranged from 30 to 60 s depending on the scan. Changes to individual cavity gradients were effected slowly, between 0.1 and 0.4 MV/m/s, in order to maintain RF stability and track changes in radiation levels. Cavities would take additional pauses as needed to accommodate supporting systems such as RF tuners or the cryogenic system.

After data collection, the raw data was processed by removing known issues. This included removing samples where cavities had experienced an RF fault. Finally, NDX detectors provide average dose rates for the preceding one-second window, which lag gradient readings and require adjustment.

In total, we collected data on three occasions. Information about these datasets is included in Table I. Some sessions were devoted to data collection, while others involved short tests of individual settings (labeled scan and demo, respectively). Neutron radiation readings appear relatively stable over this time frame; however, gamma radiation has a more noticeable increase in the final datasets.

TABLE I. Statistical and categorical information about our data. Neutron and gamma radiation readings (in rem/h) were totaled over all 13 detectors and averaged over the first ten samples of that day's data collection to get initial readings.

Date	Type	Samples	Neutron	Gamma
May 9	Scan	5880	17.38	228.12
May 14	Scan	4134	16.44	219.63
May 19	Scan	1733	18.82	283.82
May 19	Demo	1200		

B. Data drift

The field emission and radiation environment in CEBAF is prone to both sudden and gradual change. This is due to changes in the configuration of CEBAF, as well as changes in radiation response to a given configuration. During an experimental run, SRF cavities routinely experience problems that reduce their gradient or require the cavity to be completely bypassed. In most cases, these problems are addressed in the next maintenance period. In the interim, the lost gradient must be redistributed throughout the linac to maintain the linac energy gain required by experimenters. Given the complexity of modern particle accelerators, a linac likely has many other unidentified sources of drift. We use principal component analysis (PCA) to reduce the gradient sample space for the south linac's 200 cavities down to two dimensions. These two dimensions clearly show the aggregate change of cavity gradients across the different datasets we use (see Fig. 3). The behavior of field emitters also changes routinely. We have observed both sudden turn-on events, where a field emitter rapidly strengthens producing much higher radiation than before, and more gradual changes, where field emitters slowly change their average radiation level over the course of hours or days without major changes to cavity gradients. For examples of these behaviors relative to detector resolution, see Fig. 4.

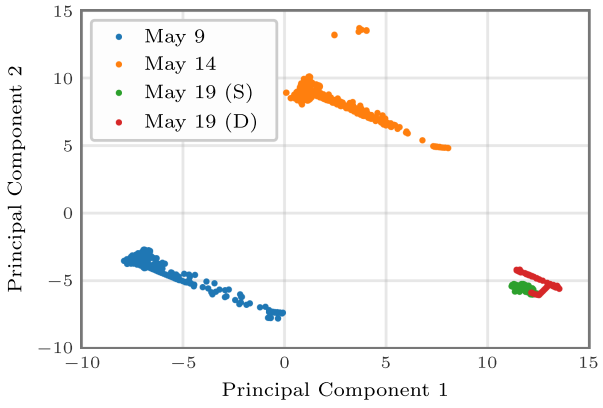


FIG. 3. Example of data drift. PCA-based dimensionality reduction of measured cavity gradients from the different datasets we collected. (S) and (D) denote the scan and demonstration datasets, respectively.

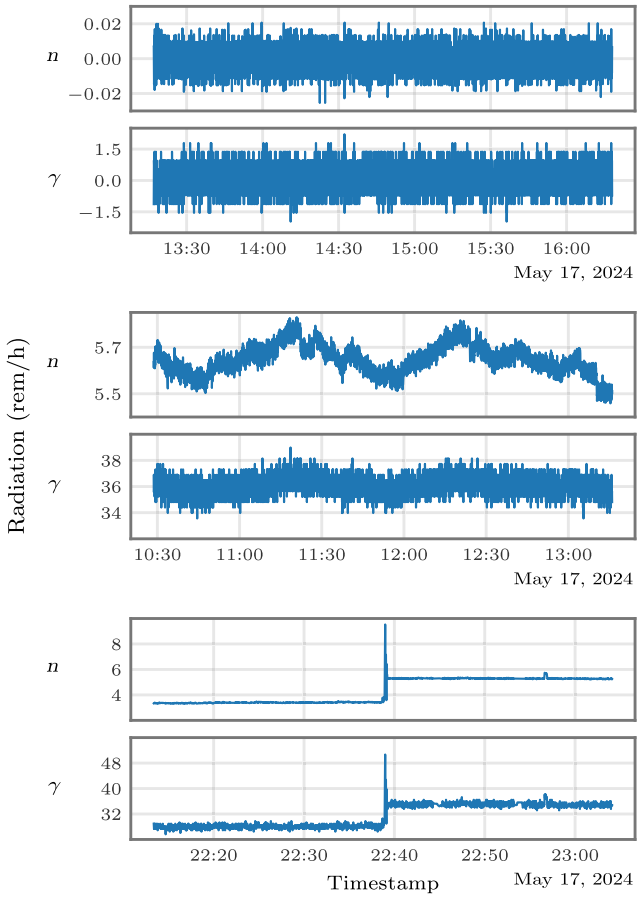


FIG. 4. Examples of 2L25 detector resolution and variance over time without changes to gradient settings. Top: baseline detector resolution with RF turned off showing noise with a standard deviation of 0.006 rem/h (neutron) and 0.548 rem/h (gamma). Middle: variability in radiation with RF on due to gradual changes in field emission over a few hours. Bottom: sudden turn-on event due to new or strengthened field emission.

This paper primarily focuses on the modeling of radiation over short time scales without the use of continual learning strategies. However, even on the time scale of a week, we observe the radiation signature changing at individual detectors.

III. SURROGATE MODELING

The ability to perform offline optimization necessitates the ability to predict radiation levels throughout a linac given a set of cavity gradients. Change is a constant in CEBAF's linacs; our solution must also allow for simple, periodic retraining. Methods that provide uncertainty quantification are desirable to guide an optimizer away from underperforming regions of the model based on noisy or poorly sampled regions of data. There are a variety of methods that could accomplish these goals, including quantile regression methods [25–27], Bayesian methods and their approximations like MC-dropout [28], Gaussian

process methods like SNGP [29], and ensemble methods like deep ensembles [30]. A more complete review of uncertainty quantified deep learning can be found in [31].

As our first attempt to improve FE-induced radiation, we focused our surrogate modeling efforts on the south linac using a simultaneous quantile regression (SQR) model [32]. Our work focused on this approach due to its inference speed, support of uncertainty quantification, ease-of-update, and ability to support high dimensional inputs.

A. Architecture

To reduce the possibility of data and model correlations affecting performance, we developed a single model per detector where the input for each model consists of 16 gradients corresponding to the cavities in the two adjacent (upstream and downstream) cryomodules. While this could cause problems for instances where FE travels further from the source cavity, the majority of the radiation response is often seen locally and therefore will be captured by this reduced model. We found this approach was more likely to match the FE behavior when compared to a single model using all 200 gradient measurements.

Each of our neural network models utilizes SQR to produce predictions as well as an estimate of model uncertainty. SQR is an extension of quantile regression, which attempts to model functions that represent specific quantiles by minimizing the quantile loss. Given a chosen quantile $0 < \tau < 1$, observations y, and predictions $\hat{y}$, the quantile loss over n samples is defined by

$$\frac{1}{n}\sum_{i=0}^{n}(y_i - \hat{y}_i)(\tau - 1_{y_i<\hat{y}_i}), \tag{1}$$

where $1_{y_i<\hat{y}_i}$ is the indicator function for $y_i < \hat{y}_i$. This loss is not symmetric (i.e., positive and negative losses are treated differently) in order to push the convergence of the model away from the median and toward the desired quantile. When $\tau = 0.5$, using the quantile loss is equivalent to half the mean absolute error: $\sum_{i=0}^{n} \|y_i - \hat{y}_i\|/n$.

Compared to standard quantile regression, SQR estimates the entire conditional distribution through a single model rather than separate models or outputs for each desired quantile. This can help reduce a problem commonly known as the quantile crossing problem, where $\tau_1 < \tau_2$, but $f_{\tau_1}(x) > f_{\tau_2}(x)$, where $f_\tau(x)$ is the model prediction for input x and quantile τ. To do this, a single model is trained using the quantile loss by including the desired quantile as part of the input to the model. Therefore, the model learns a single conditional function $f(x, \tau)$ instead of multiple separate ones, which is believed to be more likely to produce a monotonic function with respect to τ [32].

To be specific, when training the SQR model, τ is chosen for each sample in each batch of data as a uniform random variable between 0 and 1, and the loss is calculated using those values of τ. For this work, we use $\tau = 0.5$ (i.e., $f[x, 0.5]$) for predictions, as it estimates the median radiation response. Uncertainty estimates were produced by choosing $\tau = 0.16$ and 0.84, which represents ± 1 standard deviations of a standard Gaussian random variable. Then, we compute $\sigma = |f(x, 0.84) - f(x, 0.16)|/2$ for the estimated model uncertainty with gradient settings x. By computing σ this way, we provide a single value that represents the model uncertainty to pass to our optimizer and the uncertainty toolbox [33] for calibration metrics. One forward pass is required for each value of τ, totaling three forward passes. However, since τ is an input to the model, these forward passes can be batched together to improve prediction speed.

Our model architecture is shallow but wide, encompassing 16 inputs, a single dense hidden layer with 550 units, a dropout layer with a rate of 0.4, and a dense output layer with 2 units: one for neutron and the other for gamma radiation. Both dense layers were regularized using L2 regularization with a factor of 4×10^{-5} and a `leaky_relu` activation defined by

$$f(x) = \begin{cases} x, \text{if } x \geq 0 \\ \alpha x, \text{otherwise} \end{cases}, \tag{2}$$

where $\alpha = 0.3$. We explored using additional cryomodules but found no significant changes to model performance. These initial hyperparameters were chosen based on a less formal preliminary study in December 2023. While it is possible that our model could be improved with further hyperparameter tuning, the performance was sufficient to achieve good performance on our May training sets without overfitting.

B. Performance

We performed eightfold cross validation for data collected on May 9. First, the 5880 data points were split into 56 batches containing 105 s of data each. These 56 batches were then randomly divided into eightfolds, with seven batches for testing. Each fold is then trained on the remaining 49 batches for a total of 5145 training and 735 testing samples. This was done to ensure that each fold is tested with at least a few samples with novel gradient settings, as a purely sequential splitting created folds with either very low or very high variability in gradients. We present the mean and standard deviation of the root-mean-squared error (RMSE) over the eightfolds for each sensor in Table II. Additionally, we show the standard deviation of each detector signal over the first 80 samples to show an estimate of the sensor noise when gradients are stable.

Table II shows that our models are accurate for detectors near C25s and C50s (2L04–2L17) and do not overfit to the detector noise (ϵ), while model predictions for detectors near the C100s are less accurate. This makes sense as

TABLE II. Mean and standard deviation of test RMSE (denoted μ and σ, respectively) over eightfolds. We also calculate the noise of each detector with RF on (denoted ϵ_n and ϵ_γ) based on the first 80 samples of the May 9 data collection when gradients were stable. We see that the models for all detectors before 2L22 have average RMSE values that match the data noise. The models for detectors near the C100s, which experience the largest changes in radiation, are less accurate.

Detector	ϵ_n	μ_n	σ_n	ϵ_γ	μ_γ	σ_γ
2L04	0.008	0.008	0.000	0.464	0.545	0.036
2L07	0.007	0.007	0.000	0.496	0.597	0.049
2L08	0.011	0.014	0.002	0.601	0.941	0.211
2L10	0.012	0.014	0.000	0.487	0.487	0.014
2L11	0.014	0.014	0.000	1.924	2.008	0.035
2L16	0.016	0.017	0.000	0.871	0.886	0.029
2L17	0.017	0.017	0.001	2.242	2.235	0.098
2L22	0.008	0.052	0.030	0.502	1.649	0.623
2L23	0.014	0.276	0.138	1.313	4.669	1.941
2L24	0.013	0.259	0.091	0.507	3.123	0.932
2L25	0.033	0.310	0.131	0.506	1.921	0.814
2L26	0.018	0.375	0.103	0.808	4.857	1.494
2L27	0.010	0.100	0.038	0.638	1.181	0.259

little-to-no radiation was detected in the 2L04–2L17 regions. The average RMSE for detectors that experience meaningful radiation is within 0.4 rem/h for neutron radiation and 5 rem/h for gamma radiation. Additionally, the standard deviation of RMSE over the eightfolds is much higher for these detectors, indicating that the test sets for certain folds may contain novel gradient responses that are hard to predict.

To better understand these results, we calculated the standard deviation over the full May 9 collection for each detector. We found the standard deviation for 2L04–2L17 neutron radiation over all of May 9 is less than 0.017 rem/h, which matches the maximum ϵ for those detectors in Table II. This indicates that these detectors are dominated by sensor noise rather than variation as a result of gradient changes. However, the standard deviation for detectors 2L22–2L27 is more than an order of magnitude higher than the calculated noise floor (maximum 0.613 rem/h). A similar phenomenon is observed for gamma radiation. Therefore, we truncate our tables to focus on detectors 2L22–2L27, which show nontrivial gradient responses. For completeness, the tables for detectors 2L04–2L17 are presented in the Appendix.

Before running a full test on the accelerator, we were able to collect data on May 14 to validate our model trained on all the data of May. RMSE results and Pearson correlation coefficients for this model and dataset are presented in Table III. We see large increases in average error for all but 2L23; however, all detectors except 2L24 have good correlation values. To emphasize this, we plot Fig. 5, which shows the radiation delta from the beginning of data collection on May 14. This plot removes

TABLE III. May 14 RMSE and correlation for the neutron (n) and gamma (γ) predictions from our model trained on the full May 9 data. Other than 2L23, all detectors show a noticeable increase in RMSE when compared to Table II.

	RMSE		Corr	
Detector	n	γ	n	γ
2L22	0.169	2.567	0.850	0.913
2L23	0.267	2.862	0.938	0.946
2L24	0.574	6.727	0.303	0.396
2L25	1.210	6.571	0.766	0.726
2L26	1.518	18.070	0.926	0.924
2L27	0.245	2.225	0.962	0.928

any systematic shift in the predictions that could be accounted for prior to data collection and helps to visualize the correlation between the observations and predictions. We also show an uncertainty band covering $\pm 3\sigma$ and see that most observations are contained within the band. One notable portion of these graphs where the observations fall outside of the uncertainty band is between 10:20 and 10:35, where there was a large predicted reduction in radiation that did not occur in 2L24. However, that drop in radiation exists in 2L25 and is slightly underpredict for neutron radiation, suggesting that the behavior of a field emitter may have become more directional or moved away from 2L24, pushing radiation further down the linac.

C. Uncertainty quantification

To determine the calibration of our models, we utilize the expected calibration error (ECE) calculated using the uncertainty toolbox [33]. This error is calculated by comparing normalized residuals, $(y_i - \hat{y}_i)/\sigma_i$, with the inverse cumulative distribution function of a standard normal $\mathcal{N}(0, 1)$ distribution. Specifically, ECE calculates the average absolute error between the observed proportion of normalized residuals and the expected proportion predicted by a normal distribution for 100 different equally spaced proportions between 0 and 1.

The results of the May 9 model on the training data as well as the May 14 validation set are shown in Table IV. It is clear that the model is relatively well calibrated on the training data, but its performance deteriorates for the C100 detectors, which aligns with the increases in RMSE in Table III. After further analysis, we noticed the standard deviation of our uncertainty estimates remained very similar for both datasets, and the mean uncertainty estimates only increased by a small margin. This leads us to believe that the model has done a good job modeling the detector noise; however, when predictions are less accurate (like those on the May 14 data), the model becomes less well calibrated. In general, these results reinforce the idea that providing accurate uncertainties for out-of-distribution (OOD) data is especially challenging. Additionally, if the

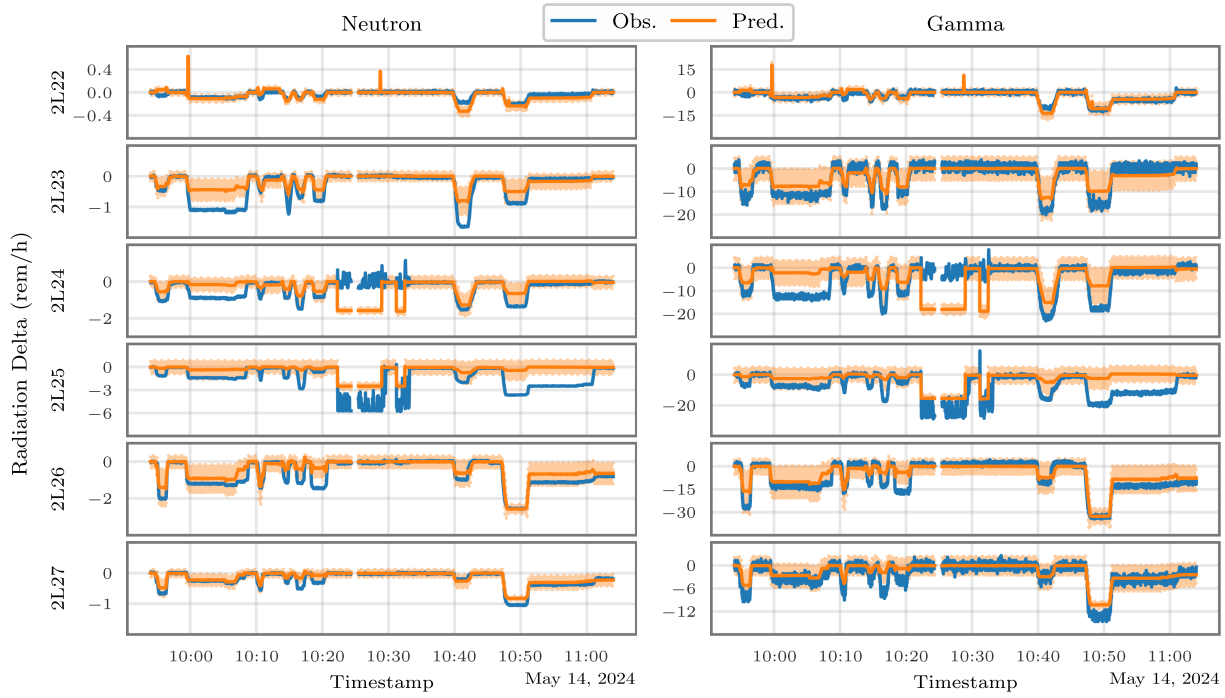


FIG. 5. Radiation delta from the initial timestep for observations (in blue) and model responses (in orange) on the May 14 dataset. Uncertainties are shown as an orange shaded region equivalent to $\pm 3\sigma$. By plotting the radiation delta rather than the exact observations and predictions, we remove any systematic shift and highlight true missed predictions. While most model responses are well correlated, we see a possible change in behavior for 2L24 (near 10:25).

relationship between radiation and gradients changes, (i.e., concept drift exists), it may be nearly impossible for a model that only receives cavity gradients to predict higher uncertainty for the drifted data. Unfortunately, operations staff identified a newly active field emitter affecting the detector at 2L25 between May 9 and May 14. Additionally, a cavity was bypassed shortly after taking data on May 9, and rebalancing the gradient significantly lowered the radiation for the detector at 2L26. It is clear that some level of both concept and data drift are occurring during this work, which complicates the use of UQ.

TABLE IV. ECE values using our model trained on the full May 9 data. We provide results for both the May 9 training data and the May 14 validation data. We see that the model is generally well-calibrated in training, but performs worse in validation for the C100 detectors.

	Training		Validation	
Detector	n	γ	n	γ
2L04	0.026	0.059	0.025	0.036
2L07	0.034	0.061	0.016	0.055
2L08	0.032	0.057	0.039	0.052
2L10	0.013	0.084	0.011	0.074
2L11	0.032	0.033	0.043	0.033
2L16	0.024	0.031	0.015	0.019
2L17	0.041	0.029	0.035	0.025
2L22	0.066	0.028	0.478	0.407
2L23	0.119	0.028	0.442	0.175
2L24	0.129	0.067	0.325	0.334
2L25	0.140	0.062	0.333	0.391
2L26	0.142	0.101	0.490	0.490
2L27	0.153	0.059	0.486	0.372

IV. OPTIMIZATION

The core of our problem is the need to reduce radiation caused by SRF cavity field emission while maintaining the linac energy gain required by the experimental program. Linac energy gain can be directly calculated for a given gradient distribution, but predicting the associated radiation levels can only be achieved by use of our surrogate models. While not all cavities function as inputs to a surrogate model, all do contribute to the linac energy gain and, consequently, must be included in the optimization problem.

The need to minimize both neutron and gamma radiation while respecting bounds on linac energy gain defines a multiobjective optimization problem with constraints. We also desired flexibility to expand our problem definition to include metrics such as uncertainty quantification or other CEBAF characteristics, which may not be readily differentiable. This is important for future deployment, as radiation production is only one of many problems operations must address when selecting a linac gradient

distribution. These requirements led us to use an evolutionary algorithm, as there are many variants readily available for solving this class of problem. We chose the nondominated sorting genetic algorithm II (NSGA-II) [34] as the main optimization algorithm due to its widespread familiarity within the accelerator physics community at large, particularly at Jefferson Lab [35–37].

Each optimization problem was defined around an initial linac configuration that represented the current operational configuration. Bounds on individual cavity gradients were determined as the minimal range that met both real-world constraints and a $[-3, +0.5]$ range around the cavity's initial gradient setting. This bounding helps to limit the optimizer from exploring far outside the training distribution, respect real-world cavity limits, and keep solutions close to the initial configuration, which addresses many more considerations than simply energy gain and radiation.

Minimizing uncertainty estimates and maximizing energy gain are also included as objectives as they help steer the optimizer in useful directions. We prefer regions of data where the observed variation in radiation was smaller, and early studies indicated that prioritizing lower uncertainty values could help prevent the exploitation of model weaknesses. Linac energy gain is already constrained to an acceptable region for beam transport, but we did not want only solutions clustered at the edge of minimum viability. Therefore, we include energy gain as both a constraint and an optimization target.

In order to provide a mathematical description of this problem, we first define a few key quantities. Given a gradient distribution $G = [g_1 g_2, \cdots g_{200}]$, and a vector of cavity lengths L, the linac energy gain is calculated as their dot product,

$$E(G) = G \cdot L.$$

Different cavity styles have different active lengths; however, all are either 0.5 or 0.7 m. As stated earlier, this energy gain must be maintained within an error tolerance ϵ of the experimentally required value E^*. For our problem, the linac energy gain tolerance is 0.5 MeV. The predicted neutron radiation and uncertainty for detector $d \in D$ are given as $\hat{y}_t(G, d)$ and $\hat{\sigma}_t(G, d)$, respectively, with radiation type $t \in [n, \gamma]$. Finally, we define

$$R_t(G) = \sum_{d \in D} \hat{y}_t(G, d),$$
$$U_t(G) = \sum_{d \in D} \hat{\sigma}_t(G, d).$$

Given an upper and lower bound for each gradient (g_i^u and g_i^l), our optimization problem is given by

$$\begin{aligned} \underset{G}{\text{minimize}} \quad & R_n(G), R_\gamma(G), U_n(G), U_\gamma(G), -E(G) \\ \text{subject to} \quad & |E(G) - E^*| \le \epsilon \\ & g_i^l \le g_i \le g_i^u, \quad \forall\ g_i \in G. \end{aligned} \tag{3}$$

While the problem is formulated in reducing absolute radiation amounts, the optimization problem should be fairly tolerant to baseline changes in radiation at individual detectors. This is because finding the minimum of the function $R_t(G)$ is equivalent to finding the minimum of $R_t(G) + c$, where c is some constant. We have noticed a historical pattern of increased or decreased baseline radiation detection over time that does not appear tied to new field emitters. While having a surrogate model that can make accurate absolute radiation predictions is ideal, even a model that only captures relative changes in radiation should be sufficient to direct the optimization process to cavity gradients that reduce radiation.

The initial population is generated by a two-phase greedy algorithm. The first phase searches for small gradient changes that reduce the total radiation dose rate. Each cavity is checked for its radiation reduction potential, and the cavity that corresponds to the largest radiation drop has its gradient reduced for the next iteration of the search. The second phase increases cavity gradients in a similar fashion while restoring linac energy gain and increasing the neutron radiation as little as possible. This search respects the individual cavity constraints at all steps, but individual iterations may violate the total energy constraint. The solutions chosen at each iteration of both phases are retained and scaled to meet the energy constraint while respecting individual cavity bounds. These solutions form the initial population for NSGA-II. This initialization strategy yielded a significant decrease in time to convergence, as it both provided a warm start to the multiobjective problem and an initial population that did not violate any constraints.

Convergence criteria are chosen to be a set number of generations due to its simplicity and flexibility. Approximate convergence would be achieved on the radiation dose rate objectives after 1000–10,000 generations. This corresponds to just a few minutes of compute time on a standard personal computer. Minor improvements across all fronts could be achieved if the optimization were run for hours longer. However, solutions are often needed in near real time as the operational environment is dynamic. For example, a derated RF cavity immediately requires a new gradient distribution to compensate for its lost energy gain. This convergence criteria give end users control over this time versus optimality trade-off. If additional computation time appears needed, the user can simply continue the optimization longer. While crude, this approach to convergence provided a

simple means to produce radiation reductions in both research and control room situations.

The NSGA-II algorithm produces a family of solutions at various positions along the Pareto front. We leverage this in our optimization software to allow operations staff to review the various proposed solutions, and select the one(s) that seem most likely to reduce radiation and respect other accelerator constraints. Ultimately, operations staff could simply review the suggested gradient changes and only apply some subset that they believe to be the most workable. The flexibility allows for our optimization process to benefit operations in a variety of workflows.

V. RESULTS

In order to provide new gradient settings that reduce FE, we first model the linac radiation response and noise (i.e., uncertainty) with a neural network based on simultaneous quantile regression. Then we utilize this surrogate model to optimize radiation with a genetic algorithm, subject to a number of constraints, including total energy and individual cavity gradient bounds. This two stage approach allows analysis of the neural network outputs separately from the optimization process and should produce effective gradient settings given an accurate surrogate model.

We developed Python-based software, called FEMinimizer, that provides an interactive GUI for investigation and selection of an optimized gradient distribution. Machine learning models are integrated and run using the ONNX runtime framework [38], and optimization algorithms are directly implemented (e.g., greedy) or using the pymoo package (e.g., NSGA-II) [39].

The application allows a user to select a historical point in time and a linac for which to generate an optimization configuration, the cavity gradient offset ranges used in optimization, and the algorithm used to perform the optimization. It also features the ability to save, view, or load an optimization configuration. Optimizations can be run iteratively for a user-specified number of generations or iterations. The set of solutions is displayed as a series of plots showing each pair of objectives. For NSGA-II, these are typically called Pareto fronts. Individual solutions can be selected based on their location in the Pareto fronts, which loads a more detailed view of that solution. Once selected, commands to apply that gradient distribution can be generated through the application.

In order to demonstrate the effectiveness of this integrated approach to FE management, we collected data on May 9 and May 14 to support a beam study on May 19. As mentioned above, models were trained and validated on May 9 data, tested on May 14 data, and then the trained models were loaded into the FEMinimizer software to generate a population of gradient distributions. Below, we describe the results of applying one of these solutions to CEBAF.

A. Demonstration results

Since the model was still highly correlated to radiation, we conducted a real-world test using our NSGA-II optimizer at the next opportunity on May 19. The optimizer took in the current gradient settings of the south linac, and produced a set of new gradients to optimize for the objectives listed in Sec. IV. In order to see how these new settings impacted radiation levels, we applied the largest changes to the linac one-at-a-time, followed by applying all of the smaller changes at once. The selected changes in gradient settings are shown in Fig. 6. Predictably, a few C100 cavity gradients were significantly reduced by the maximum -3 MV/m, while many C25 cavity gradients were increased to compensate for the lost energy.

Accuracy and correlation results are shown in Table VI with a full plot of the radiation responses in Fig. 7. Like before, we see larger RMSE errors for most detectors when compared to May 14 (except 2L25 and 2L27) and correlations are still quite good. This short test run shows a number of interesting features. While the model does not track predictions exactly, the major source of error appears to be from an upward shift in radiation.

To further verify our results, we show the observed and predicted total neutron and gamma radiation at the start and end of our demonstration, as well as the delta and

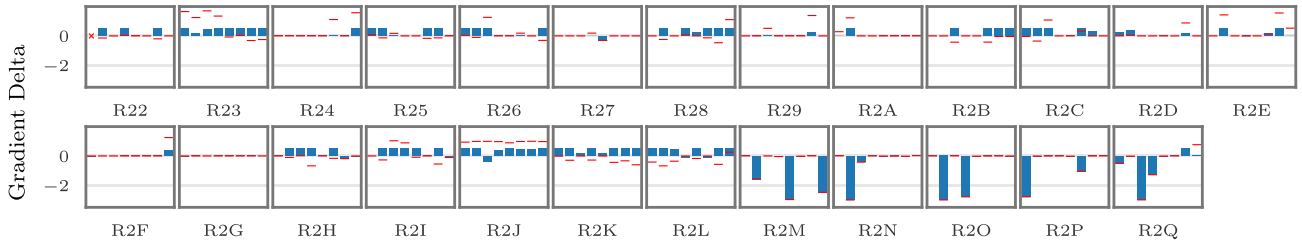


FIG. 6. Delta between gradient settings chosen by our NSGA optimizer and the baseline settings (in blue). Since a few settings were modified between the baseline and the beginning of the demonstration, the true demonstration deltas (for Fig. 7) are shown as red lines. R221 was turned off after our baseline was taken, thus this bar is replaced by an "x". The demonstration delta for R2Q7 was $+3.8$ (not shown).

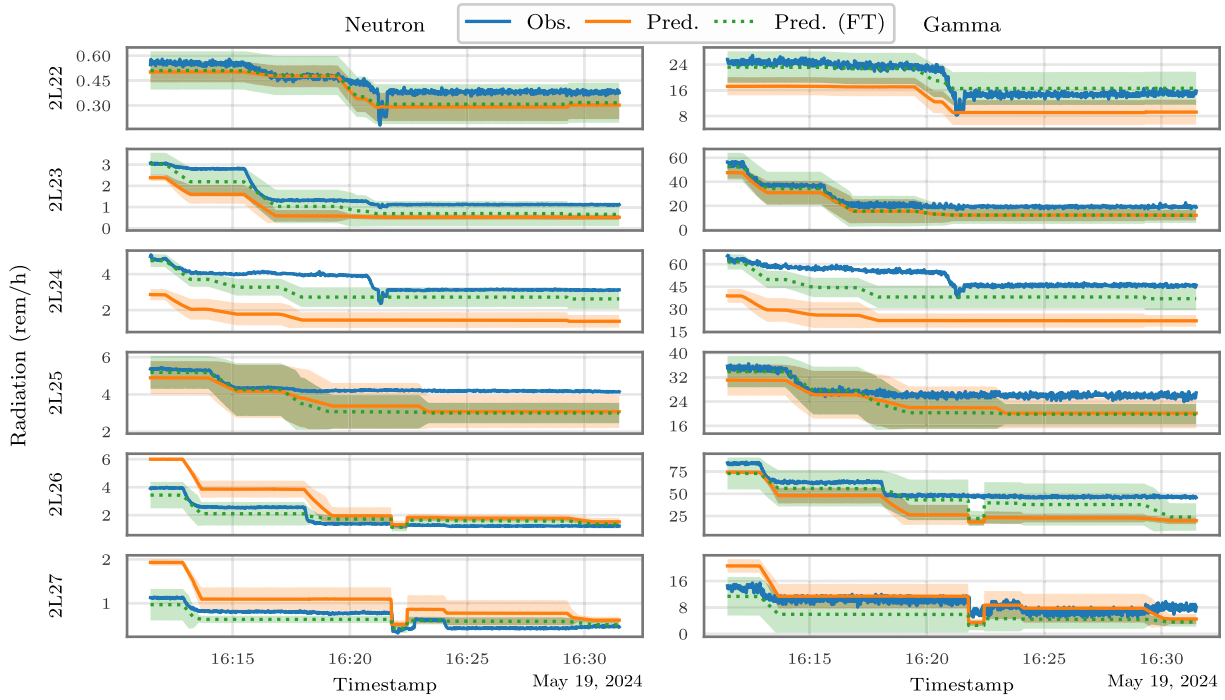


FIG. 7. Prediction quality on May 19 demonstration data for our model trained on May 9 and then fine-tuned on the May 19 data. Observations are in blue, predictions are in orange, and the fine-tuned predictions are dotted green. Uncertainties are shown as shaded regions equivalent to $\pm 3\sigma$. The model is well correlated, and fine-tuning it helps to move predictions toward the observations for most detectors.

percentage difference between them in Table V. At the initial gradient settings ("start"), the predicted total neutron radiation matches the observed radiation well, unlike the total gamma radiation levels, which are significantly under predicted. The discrepancy in gamma radiation could possibly be explained by Table I, which shows total gamma radiation significantly increased between our May 9 training data and the May 19 demonstration, compared to a much smaller increase in total neutron radiation. It should also be noted that this discrepancy is likely due to changing conditions in the linac, rather than poor model performance, since the correlation in the total radiation is 0.991 and 0.988 for neutron and gamma, respectively. We also see that the model is generally optimistic and estimates a 60% reduction in radiation, exceeding the observed 45% and 41% for neutron and gamma, respectively. Even so, these reductions are significant and represent tangible real-world improvements to radiation management at CEBAF.

TABLE V. Observed and predicted total radiation for our demonstration on May 19. The model predictions are slightly optimistic, but reasonable overall.

	y_n	$\hat{y}_n$	y_γ	$\hat{y}_\gamma$
Start	18.99	18.63	300.24	240.76
End	10.48	7.46	176.25	98.30
Delta	8.51	11.16	123.99	142.46
Percent	44.82	59.93	41.30	59.17

B. Fine-tuning

Since there was an observable shift in performance between May 9 and our demonstration on May 19, we decided to test the efficacy of fine-tuning our model on a small amount of scan data taken on the demonstration day. To do this, we trained our model for 100 epochs on this smaller, but more recent, dataset. As presented in Fig. 3, this dataset contains many novel gradient settings, including multiple cavities with changes in baseline gradient settings that were larger than 2 MV/m. In Fig. 7, we see that the majority of improvement comes from a simple shift in model responses to account for possible detector drifts and gradient changes. Since the dataset we tuned the model on was small, and the space of possible gradient settings is very large, it is likely that the fine-tuning dataset does not contain enough information to significantly alter model responses.

The accuracy and correlation results are shown in Table VI and Fig. 7 alongside the original model results. While the shape of model responses is similar after retraining, we see that it significantly improved 75% of our RMSE metrics, including a reduction of more than 16 and 10 rem/h for 2L24 γ and 2L26 γ, respectively. Only minor degradations in the RMSE performance metrics are observed for both the 2L25 detectors and the 2L27 γ detector. Additionally, as seen in Fig. 7, many uncertainty predictions increased, which is likely due to a combination of factors, including retained information from the earlier training data,

TABLE VI. Results of our original model and a fine-tuned model (subscripted ft) on the May 19 demonstration data. Highlighted cells indicate better performance. The majority of RMSE results significantly improve when fine-tuning, and the detectors that see worse RMSE scores only decrease by a small amount.

Metric	Detector	n	n_{ft}	γ	γ_{ft}
RMSE	2L22	0.071	**0.059**	6.482	**1.959**
	2L23	0.746	**0.417**	6.321	**5.802**
	2L24	1.978	**0.664**	26.932	**10.539**
	2L25	**0.840**	0.950	**4.514**	4.926
	2L26	1.048	**0.378**	20.655	**10.581**
	2L27	0.372	**0.148**	**2.631**	3.704
Corr	2L22	0.928	**0.932**	**0.969**	0.968
	2L23	0.976	**0.981**	0.987	**0.987**
	2L24	0.790	**0.799**	0.765	**0.792**
	2L25	0.884	**0.907**	0.889	**0.902**
	2L26	**0.986**	0.944	**0.977**	0.895
	2L27	**0.924**	0.802	**0.843**	0.830

and the broader distribution of residuals from our limited fine-tuning. In general, correlations were also improved, but none of the changes were significant as the model was already highly correlated with the true observations.

C. Trip rate modeling

As alluded to in Sec. IV, operational gradient distributions are optimized against many additional considerations. A primary concern of the current gradient distribution methods is minimizing the number of RF cavity arc trips that occur within the linac [40,41]. An individual cavity's arc trip rate is well modeled using a log-linear statistical relationship with its gradient, and these model parameters are stored in a historical database similar to other CEBAF parameters. The estimated trip rate across the linac is simply the sum of the trip rates of the modeled cavities. A critical follow-up question to the May 19 demonstration is how redistributing the gradient away from active field emitters impacts the overall trip rate.

In order to investigate this topic, we modify the optimization problem from Sec. IV by adding the minimization of the linac trip rate as an objective, including a constraint bounding the maximum linac trip rate at ten trips/hour, and relaxing the individual cavity gradient change constraints from $+0.5$ to $+1.0$ MV/m. We found that limiting gradient increases to only 0.5 MV/m had a significant impact on trip rates, as the optimizer was forced to increase gradients on less stable cavities. This is consistent with the fact that trip rates are modeled using an exponential response to cavity gradients, and some cavities are much more prone to trips than others.

Using this new optimization problem, we generated a population of gradient distributions and highlighted a specific reoptimized gradient distribution that yields a trip rate similar to the baseline (Table VII). This new solution suggests a neutron and gamma radiation reduction of 37% and 34%, respectively. Unfortunately, it is impossible to know what the real radiation would be for this new solution. Given the observed tendency of our models to overestimate radiation reduction for large gradient changes, we suspect that these estimates may also be optimistic. However, that suspicion is tempered by this new solution's more conservative gradient reductions. That makes it reasonable to believe that this solution would produce substantial radiation reductions, even if not at the previously observed levels of 45% and 41% for neutron and gamma, respectively. It is important to note that even if this estimate is true, the inclusion of additional operational considerations could increase radiation and/or trips by further constraining the problem.

TABLE VII. Comparison of model estimated radiation and trip rates between the demonstration starting point (baseline), the optimized gradients from the demonstration (optimized), and gradients reoptimized for both trip rate and radiation levels (reoptimized).

	Baseline	Optimized	Reoptimized
Trip rate (trip/h)	4.27	6.92	4.27
Neutron (rem/h)	18.63	7.46	11.73
Gamma (rem/h)	240.76	98.30	159.12

VI. DISCUSSION AND CONCLUSION

Our final demonstration on May 19 gives a proof of concept about how this approach could benefit operations through reduced radiation. Operations staff had been mitigating radiation throughout the experimental run prior to our demonstration, therefore the 40%–45% reduction in radiation represents a real, tangible decrease in radiation over what would have been seen in operations. Field emission has been a constant problem, and our approach represents a viable method to significantly reduce its observed effects through passive mitigation.

As we note throughout this work, CEBAF operations contend with many considerations beyond field emission when setting cavity gradients. Operations must also address cavity trip rates, electrical power limits, RF klystron power limits, and energy lock gradient reserves, among other concerns. Integrating field emission mitigation into linac gradient distribution requires a considerable effort and system expert discretion. Trip rates represent one of the most pressing concerns when optimizing gradient distributions, and our brief investigation suggests that significant radiation reduction can be achieved even when accounting for cavity trips.

Our model architecture provides for reasonable surrogate models. It avoids many extraneous correlations

present in our relatively small training datasets and can be readily fine-tuned on new data. We expect that further refining data collection practices would improve the already sufficient model performance. While our model provides uncertainty quantification, this aspect suffers from drift in the configuration of CEBAF and the presence of at least one new active field emitter. As such, it appears that detecting OOD conditions is an open question in this challenging scenario. However, our model appears to do a reasonable job of representing the uncertainty introduced by random fluctuations in detector noise. Future improvements in OOD detection could help dissuade optimization processes from exploiting model weaknesses due to the limited training data and the large input space. Another avenue for improvement would be to model the entire linac with a single model. Most of our models were simple by design to avoid spurious correlations that often appeared in our data. Future work could investigate other techniques, such as self-attention to better identify which cavities are field emitters, physics-informed models to ensure model outputs better represent the real world, and methods for uncertainty quantification that improve OOD identification.

For the model to remain reflective of radiation response through the span of an experimental run, we would likely need to institute continual learning techniques. Here, we take the first successful step in this direction, by showing the ability for fine-tuning on a small training set to reduce prediction errors. These data were collected in 30 minutes and could reasonably be performed by operations semi-routinely. An initial step in continual learning could involve retraining only the final biases for models with good correlation but poor errors due to shifts.

Another avenue for maintaining model accuracy over time is leveraging passive data from CEBAF. RF cavities routinely trip and recover giving potentially useful information on how radiation is impacted by changing gradient. Additionally, cavity gradients are routinely redistributed throughout the linacs in order to maintain a constant linac energy gain. Using these passive data sources may reduce our reliance on invasive data collection that interferes with experimental runs. We briefly explored this topic in [42]; however, that work focused on only a single section of a linac and was hampered by data limitations. This work did suggest that RF trip data could be useful in improving the model performance in certain cases, making further study warranted.

Many tools were created to enable this work, and most represent a solid start on production software needed to implement this in routine operations. While data collection still requires some amount of manual intervention, we have developed a software framework for controlling RF throughout a linac and implementing a variety of sampling strategies. The optimization software already provides a polished user interface and is integrated with CEBAF systems for accurately reflecting CEBAF configuration at arbitrary points in time. This significantly lowers the barrier to entry for production use of the techniques showcased in this work.

ACKNOWLEDGMENTS

We would like to acknowledge Jay Benesch, Pavel Degtiarenko, and Michael McCaughan for their many informative conversations, and CEBAF operations and engineering staff more broadly for their support throughout this endeavor. This work was supported by the U.S. Department of Energy, Office of Science, Office of Nuclear Physics under Contract No. DE-AC05-06OR23177.

APPENDIX: ADDITIONAL RESULTS

Here, we present additional results for the non-C100 detectors. Compared to the C100 detectors, these detectors exhibit consistently low error but poor correlations. Since these detectors have almost no true signal and are dominated by noise, any model with good correlation results would be overfitting to the true data. The only exception is the results for 2L08 radiation which sometimes contains a change in radiation signal that can be modeled.

Table VIII presents the validation results of our model on the May 14 dataset. Table IX shows the same metrics

TABLE VIII RMSE and correlation on the May 14 dataset for our model trained on the full May 9 data.

	RMSE		Corr	
Detector	n	γ	n	γ
2L04	0.008	0.576	0.098	0.365
2L07	0.007	0.706	0.028	0.371
2L08	0.017	1.142	0.721	0.874
2L10	0.014	0.511	0.027	−0.023
2L11	0.014	1.998	0.014	−0.057
2L16	0.017	0.940	0.024	0.263
2L17	0.017	2.240	0.186	0.081

TABLE IX RMSE and correlation on the May 19 dataset for our model trained on the full May 9 data.

	RMSE		Corr	
Detector	n	γ	n	γ
2L04	0.008	0.513	0.003	0.069
2L07	0.008	0.671	−0.098	−0.353
2L08	0.011	2.907	0.039	0.793
2L10	0.014	0.492	0.053	−0.042
2L11	0.014	1.967	−0.071	0.058
2L16	0.016	0.861	−0.017	0.007
2L17	0.016	2.255	−0.034	0.010

TABLE X RMSE and correlation results on the May 19 demonstration set using a fine-tuned May 9 model. The fine-tuning utilized data collected on May 19, prior to the demonstration.

	RMSE		Corr	
Detector	n	γ	n	γ
2L04	0.008	0.514	0.003	0.069
2L07	0.008	0.588	−0.093	−0.352
2L08	0.012	1.091	0.039	0.798
2L10	0.014	0.494	0.051	−0.043
2L11	0.014	1.962	−0.070	0.056
2L16	0.016	0.857	−0.016	0.007
2L17	0.016	2.255	0.032	−0.013

(RMSE and Pearson's correlation) for the May 19 demonstration. Table X shows the results of the fine-tuned model on the May 19 demonstration dataset.

[1] P. A. Adderley, S. Ahmed, T. Allison, R. Bachimanchi, K. Baggett, M. BastaniNejad, B. Bevins, M. Bevins, M. Bickley, R. M. Bodenstein *et al.*, The Continuous Electron Beam Accelerator Facility at 12 GeV, Phys. Rev. Accel. Beams **27**, 084802 (2024).

[2] J. Arrington, M. Battaglieri, A. Boehnlein, S. Bogacz, W. Brooks, E. Chudakov, I. Clöt, R. Ent, H. Gao *et al.*, Physics with CEBAF at 12 GeV and future opportunities, Prog. Part. Nucl. Phys. **127**, 103985 (2022).

[3] R. H. Fowler and L. Nordheim, Electron emission in intense electric fields, Proc. R. Soc. A **119**, 173 (1928).

[4] H. Padamsee, J. Knobloch, and T. Hays, *RF Superconductivity for Accelerators* (Wiley, Hoboken, NJ, 2008).

[5] H. A. Schwettman, J. P. Turneaure, and R. F. Waites, Evidence for surface-state-enhanced field emission in RF superconducting cavities, J. Appl. Phys. **45**, 914 (1974).

[6] A. Navitski, S. Lagotzky, D. Reschke, X. Singer, and G. Müller, Field emitter activation on cleaned crystalline niobium surfaces relevant for superconducting RF technology, Phys. Rev. ST Accel. Beams **16**, 112001 (2013).

[7] D. Reschke, V. Gubarev, J. Schaffran, L. Steder, N. Walker, M. Wenskat, and L. Monaco, Performance in the vertical test of the 832 nine-cell 1.3 GHz cavities for the European X-ray Free Electron Laser, Phys. Rev. Accel. Beams **20**, 042004 (2017).

[8] M. Omet, H. Araki, T. Dohmae, H. Ito, R. Katayama, K. Umemori, and Y. Yamamoto, Efforts to suppress field emission in SRF cavities at KEK, in *Proceedings of the 21st International Conference on RF Superconductivity, SRF2023, Grand Rapids, MI* (JACoW, Geneva, Switzerland, 2023), pp. 167–171, 10.18429/JACoW-SRF2023-MOPMB033.

[9] M. Martinello, P. Berrutti, B. Giaccone, S. Belomestnykh, M. Checchin, G. V. Eremeev, A. Grassellino, T. Khabibouilline, A. Netepenko, R. Pilipenko, A. Romanenko, S. Posen, G. Wu, D. Gonnella, M. Ross, J. T. Maniscalco, and T. Powers, Plasma processing for in-situ field emission mitigation of superconducting radio-frequency (SRF) cryomodules, arXiv:2203.12442.

[10] M. Drury, A. Burrill, G. K. Davis, J. Hogan, L. King, F. Marhauser, H. Park, J. Preble, C. E. Reece, R. Rimmer, A. Reilly, H. Wang, and M. Wiseman, Performance of first C100 cryomodules for the CEBAF 12 GeV upgrade project, in *Proceedings of the 26th International Linear Accelerator Conference, Tel Aviv, Israel,* (2013), MOPB030, https://inspirehep.net/literature/1364048.

[11] A. Freyberger, Commissioning and operation of 12 GeV CEBAF, in *Proceedings of the 6th International Particle Accelerator Conference (IPAC'15), Richmond, VA* (JACoW, Geneva, Switzerland, 2015), pp. 1–5, 10.18429/JACoW-IPAC2015-MOXGB2.

[12] R. Geng, Root causes of field emitters in SRF cavities placed in CEBAF tunnel, in *Proceedings of the 7th International Particle Accelerator Conference (IPAC'16), Busan, Korea* (JACoW, Geneva, Switzerland, 2016), pp. 3198–3201, 10.18429/JACoW-IPAC2016-THOBB03.

[13] R. Geng, A. Fisher, A. Freyberger, R. Legg, and R. Suleiman, Field emission in superconducting accelerators: Instrumented measurements for its understanding and mitigation, in *Proceedings of the 6th International Beam Instrumentation Conference (IBIC'17), Grand Rapids, MI* (JACoW, Geneva, Switzerland, 2018), pp. 470–477, 10.18429/JACoW-IBIC2017-TH1AB1.

[14] Vacuum valves, VAT Group AG (2024), https://www.vatvalve.com/page/vacuum-valves [accessed October 01, 2024].

[15] P. V. Degtiarenko, NDX: Neutron dose rate meters with extended capabilities, in *Proceedings of the 2019 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), Manchester, UK* (2019), pp. 1–5, 10.1109/NSS/MIC42101.2019.9059727.

[16] P. V. Degtiarenko, Neutron detector and dose rate meter using beryllium-loaded materials, US Patent 10,281,600 (2019), https://www.osti.gov/doepatents/biblio/1568305.

[17] Ionization Chambers, LND, Inc.: Designers and Manufacturers of Nuclear Radiation Detector (2024), https://www.lndinc.com/ [accessed October 01, 2024].

[18] I400 Four Channel Digital Electrometer, Pyramid Technical Consultants, Inc. (2024), https://ptcusa.com/ [accessed October 01, 2024].

[19] M. M. Rahman, A. Carpenter, K. Iftekharuddin, and C. Tennant, Accelerating cavity fault prediction using deep learning at Jefferson Laboratory, Mach. Learn. **5**, 035078 (2024).

[20] H. Ferguson, J. Li, A. Carpenter, C. Tennant, D. Thomas, and D. Turner, Detecting anomalous SRF cavity behavior with unsupervised learning, Phys. Rev. Accel. Beams **28**, 034602 (2025).

[21] C. Tennant, A. Carpenter, T. Powers, A. Shabalina Solopova, L. Vidyaratne, and K. Iftekharuddin, Superconducting radio-frequency cavity fault classification using machine learning at Jefferson Laboratory, Phys. Rev. Accel. Beams **23**, 114601 (2020).

[22] A. Eichler, J. Branlard, and J. H. K. Timm, Anomaly detection at the European x-ray free electron laser using a parity-space-based method, Phys. Rev. Accel. Beams **26**, 012801 (2023).

[23] L. Boukela, A. Eichler, J. Branlard, and N. Z. Jomhari, A two-stage machine learning-aided approach for quench identification at the European XFEL, arXiv:2407.08408.

[24] F. Wang, Enhancing SRF cavity stability and minimizing detuning with data-driven resonance control based on dynamic mode decomposition, AIP Adv. **13**, 075104 (2023).

[25] R. Koenker, *Quantile Regression*, Econometric Society Monographs (Cambridge University Press, Cambridge, England, 2005), 10.1017/CBO9780511754098.

[26] G. Shen, Y. Jiao, Y. Lin, J. L. Horowitz, and J. Huang, Deep quantile regression: Mitigating the curse of dimensionality through composition, arXiv:2107.04907.

[27] I. Chronopoulos, A. Raftapostolos, and G. Kapetanios, Forecasting value-at-risk using deep neural network quantile regression, Journal of financial economics **22**, 636 (2023).

[28] Y. Gal and Z. Ghahramani, Dropout as a bayesian approximation: Representing model uncertainty in deep learning, in *Proceedings of the 33rd International Conference on Machine Learning*, edited by M. F. Balcan and K. W. Weinberger (PMLR, New York, NY, 2016), Vol. 48, pp. 1050–1059, https://proceedings.mlr.press/v48/gal16.html.

[29] J. Liu, Z. Lin, S. Padhy, D. Tran, T. Bedrax Weiss, and B. Lakshminarayanan, Simple and principled uncertainty estimation with deterministic deep learning via distance awareness, Adv. Neural Inf. Process. Syst. **33**, 7498 (2020), https://proceedings.neurips.cc/paper/2020/hash/543e83748234f7cbab21aa0ade66565f-Abstract.html.

[30] B. Lakshminarayanan, A. Pritzel, and C. Blundell, Simple and scalable predictive uncertainty estimation using deep ensembles, arXiv:1612.01474.

[31] M. Abdar, F. Pourpanah, S. Hussain, D. Rezazadegan, L. Liu, M. Ghavamzadeh, P. Fieguth, X. Cao, A. Khosravi, U. R. Acharya, V. Makarenkov, and S. Nahavandi, A review of uncertainty quantification in deep learning: Techniques, applications and challenges, Inf. Fusion **76**, 243 (2021).

[32] N. Tagasovska and D. Lopez-Paz, Single-model uncertainties for deep learning, in *Advances in Neural Information Processing Systems*, edited by H. Wallach, H. Larochelle, A. Beygelzimer, R. Garnett, F. d'Alche-Buc, E. Fax, and R. Garnett (Curran Associates, Inc., Red Hook, NY, 2019), Vol. 32, https://proceedings.neurips.cc/paper_files/paper/2019/file/73c03186765e199c116224b68adc5fa0-Paper.pdf.

[33] Y. Chung, I. Char, H. Guo, J. Schneider, and W. Neiswanger, Uncertainty toolbox: An open-source library for assessing, visualizing, and improving uncertainty quantification, arXiv:2109.10254.

[34] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, A fast and elitist multiobjective genetic algorithm: NSGA-II, IEEE Trans. Evol. Comput. **6**, 182 (2002).

[35] A. Hofler, B. Terzić, M. Kramer, A. Zvezdin, V. Morozov, Y. Roblin, F. Lin, and C. Jarvis, Innovative applications of genetic algorithms to problems in accelerator physics, Phys. Rev. ST Accel. Beams **16**, 010101 (2013).

[36] Y. Ge, Y. Zhong, N. Yuan, Y. Sun, Z. Yang, W. Ma, L. Zou, I. Murata, and L. Lu, Optimization of moderator materials by NSGA II based on macroscopic cross-sections: Applications in accelerator neutron sources, J. Instrum. **18**, P08004 (2023).

[37] S. B. van der Geer and M. J. de Loos, Multi-objective genetic optimization with the general particle tracer (GPT) code, in *Proceedings of the 6th International Particle Accelerator Conference, IPAC'15, Richmond, VA* (JACoW, Geneva, Switzerland, 2015), pp. 492–494, 10.18429/JACoW-IPAC2015-MOPJE076.

[38] ONNX Runtime developers, ONNX runtime (2021), https://onnxruntime.ai/.

[39] J. Blank and K. Deb, Pymoo: Multi-objective optimization in python, IEEE Access **8**, 89497 (2020).

[40] J. Benesch, A longitudinal study of field emission in CEBAF's SRF cavities 1995–2015, arXiv:1502.06877.

[41] C. E. Reece, Continuous wave superconducting radio frequency electron linac for nuclear physics research, Phys. Rev. Accel. Beams **19**, 124801 (2016).

[42] K. Ahammed, J. Li, A. Carpenter, C. Tennant, and R. Suleiman, Transfer learning for field emission mitigation in CEBAF SRF cavities, in *Proceedings of the 15th International Particle Accelerator Conferemce, IPAC'24, Nashville, TN* (JACoW, Geneva, Switzerland, 2024), pp. 160–163, 10.18429/JACoW-IPAC2024-MOPC44.