Proceedings of the INMM & ESARDA Joint Virtual Annual Meeting
August 23-26 & August 30-September 1, 2021

VERIFICATION OF TRISO FUEL BURNUP USING MACHINE LEARNING ALGORITHMS

Odera Dim
Brookhaven National Laboratory

Carlos Soto
Brookhaven National Laboratory

Yonggang Cui
Brookhaven National Laboratory

Lap-Yan Cheng
Brookhaven National Laboratory

Maia Gemmill
Brookhaven National Laboratory

Thomas Grice
Consultant

Joseph Rivers
Consultant

Warren Stern
Brookhaven National Laboratory

Michael Todosow
Brookhaven National Laboratory

ABSTRACT

Pebble Bed Reactors are fueled with fuel pebbles that are circulated multiple times through the reactor vessel before discharge. During the normal operation of a PBR, ejected pebbles are either returned to the reactor or discharged as spent or damaged fuel depending on the fuel burnup and physical condition of the pebbles. The burnup measurement is usually based on detected radiation signatures of fission products accumulated in the pebble fuel over irradiation in the core. Previous research has shown that height of photopeaks of fission products, such as ^{134}Cs, ^{137}Cs, ^{154}Eu, etc., can be used independently or in combination to infer or predict the level of burnup in the fuel. However, it remains challenging to measure such complex sources due to self-shielding effects, strong radiation background and intervening materials. Another operational challenge is the required high throughput of pebbles undergoing burnup measurement, which necessitates limited measurement time and thus impacts quality of measured gamma-ray spectra. Hence, advanced spectral analysis methods are needed to analyze the noisy gamma spectra and predict the burnup values. This paper proposes to use machine learning (ML) method to interpret gamma-ray spectra and predict the burnup values of the pebbles. ML has achieved widespread success and adoption across a few domains that require pattern recognition and analysis in varied data types. In this work, three proven ML approaches – multilayer perceptron's (MLPs), convolutional neural networks, and transformers – are applied to the task of predicting fuel burnup from measured gamma spectra, and a dataset of simulated spectra is compiled for training and validation of the ML models. This paper discusses the network architecture of these three ML approaches and compare the performance of the simplest of these (MLPs) to a standard linear regression.

INTRODUCTION

Pebble Bed Reactors (PBRs) are fueled with fuel items with roughly the size of a tennis ball that are circulated multiple times through the reactor vessel before discharge. These pebbles are formed from TRIstructural ISOtropic (TRISO) fuel particles, which are composed of a minute fuel kernel of uranium dioxide or uranium oxycarbide surrounded by multiple layers of highly durable and impermeable coating. During the normal operation of a PBR, ejected pebbles are

returned to the reactor or discharged depending on the fuel burnup and physical condition.

The burnup measurement is usually based on detected radiation signatures of fission products accumulated in the TRISO particle kernel over irradiation in the core. Years of research has shown that measurements of fission products, such as ^{134}Cs, ^{137}Cs, ^{154}Eu, etc., can be applied independently or in combination to infer or predict the level of burnup in the fuel (Akyurek, Tucker and Usman 2014). A simple criterion for selecting an isotope for burnup indication is the exhibition of a strong gamma photopeak. However, it remains challenging to measure this complex source due to self-shielding effects, strong radiation background and intervening materials. Another challenge in this measurement is the required high throughput in the burnup measurements. Accommodating this throughput necessitates limited measurement time and thus impacts detection efficiency. A high-performing spectral analysis method is therefore required to identify patterns swiftly and accurately in the time-constrained gamma spectrum measurements. This study uses machine learning (ML) methods to interpret gamma-ray spectra and predict the burnup values of the pebbles. ML has achieved widespread success and adoption across numerous domains that require pattern recognition and analysis in varied data types (Butler 2018) (Carleo 2019). Modern deep learning approaches have supplanted hand-crafted features by learning entirely novel, yet meaningful, features and data representations directly from the raw data via deep neural network architectures; this has led to state-of-the-art and even superhuman performance on a broad range of detection, interpretive, and analytical tasks. In this work, we applied three proven ML approaches – multilayer perceptron (MLPs), convolutional neural networks (CNNs), and transformers – to the task of predicting fuel burnup from measured gamma spectra. To train and test the ML models, we used SERPENT (Leppanen 2015) and GADRAS (Horne, et al. 2014) to model the fuel pebbles and the burnup process and simulate the spectra as the response of a high energy resolution high purity germanium (HPGe) detector. The performance of the ML models was compared to the photopeak-based linear regression method.

MODELING AND SIMULATION

Overall Modeling and Simulation Workflow

Modeling and simulation of burnup measurement in this work was done in two steps. The first step involved a SERPENT Monte Carlo (MC) simulation (Chersola, et al. 2014) to compute the burnup of a modeled pebble in a PBR core. The simulation produced isotopic compositions of a pebble after passing through the PBR core axially. Following that, the transport of gamma photons to the surface of the pebble was simulated to obtain the photon flux information coming out of the pebble. This second simulation was also performed in SERPENT. After that, the gamma source rate at the surface of the pebble and the discrete source lines were written into files as output from SERPENT. The second step of modeling and simulation was to convert these files to a GADRAS readable file format (.GAM files) (Rawool-Sullivan, et al. 2012) and feed them into GADRAS to produce the gamma-ray spectra in N42 or .PCF formats. To expedite the overall process, the burnup and photon transport simulation were combined into a single process that could execute on a cluster machine with 16 nodes. Batch processing of the gamma spectra in GADRAS was done on a Windows 10 desktop computer.

Models of PBR Reactors and Fuel Pebbles

Because this research was not on designing or optimization of the PBR core, modeling focused on single pebble instead of a full PBR core. To simplify and streamline the modelling process, a lattice model approach was adopted. First, the TRISO particles were modeled to establish a baseline for the requirements for a working SERPENT input file that produced reasonable output. The TRISO model was then updated to a pebble model to perform the preliminary burnup simulations and obtain the gamma source rate at the surface of the pebble. The final model was a lattice of 27 pebbles in a 3x3x3 configuration from which the centered pebble was used as the reference pebble. This lattice model was used to reduce the effect of the reflected boundary condition that was directly on the surface of the pebble, which could impact the accuracy of the calculated flux and hence estimation of isotope concentrations. With the lattice configuration, this reflected boundary was moved to the surface binding the 3x3x3 lattice. Figure 1 below shows the three different stages of the model that was used in this work. Table 1 summarizes the main parameters used in modeling the fuel pebbles after consulting the PBR fuel and reactor designers.

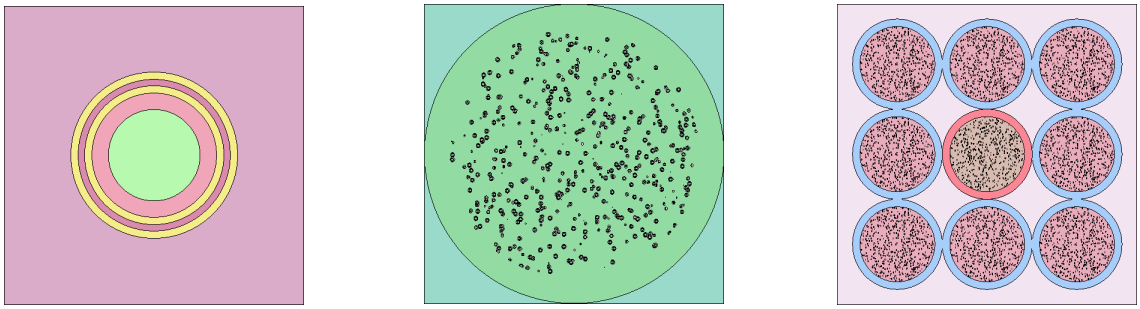

Figure 1. Model of a PBR fuel pebble from a TRISO (Left) to a pebble (Center) and a 3x3x3 Lattice (Right)

Table 1. Parameters used in modeling the fuel pebble

Parameter	Value
Uranium Oxy-Carbide (UCO) Density (atoms/b-cm)	6.9924E-02
Buffer (C) Density (atoms/b-cm)	5.2644E-02
Polycarbonate (PyC)/ Silicon Carbide (SiC) Density (atoms/b-cm)	~9.5262E-02
Number of Pebble/TRISO	27/18857
Pebble/TRISO radius (cm)	3.000/0.0455
Lattice configuration	3 x 3 x 3
Power (MW_{th})	280
Boundary condition	Reflected/Periodic
Pebble/TRISO PF[1]	0.5200/0.1137
Average residence time (days)/Cycles(passes)	522/8
Cooling time before spectral measurement (days)	0, 0.5, 1, 2, 5, 10
Data acquisition time (s)	20, 3600

[1] PF is the Parking Fraction. The TRISO(s) and Pebble(s) have different parking fractions.

Burnup and Photon Transfer Simulations

As mentioned earlier, two SERPENT simulations were used sequentially to produce photon flux on the surface of ejected pebbles. The first simulation was the burnup of a fuel pebble over various duration per cycle. These durations per cycle were computed to span across the average residence time divided by the number of cycles listed in Table 1.

Table 2. Duration per cycle and residence time of fuel pebbles in the simulated PBR core

Duration per cycle (days)	Residence time (days)
60	480
65	520
70	560

In this work, 15 durations per cycle were selected ranging from 30 to 100 days. Table 2 above shows an example of relationship between the duration per cycle and average residence times. Since each of the durations per cycle produces a unique data set, the intervals were chosen to be uniformly spaced at 5 days. The duration of a cycle was further divided into 10 sub-steps to allow proper computation of the isotope concentration over burnup and lower the uncertainty that may result from large time steps. To eliminate any correlation that could occur between any two burnup calculations over residence time, the sub-steps were randomized by 25% across the center value. At the end of each cycle a cooling time was applied to let short-lived isotopes decay. Although typically a pebble bed core would have roughly a cosine axial power shape, this work assumed an average constant power distribution over a cycle for simplicity. The burnup simulation was performed with a million particles, which reduced the statistical uncertainty in the fission reaction rates for neutrons to less than 5% and the uncertainty on the total flux to less than 0.1 %. Figure 2 shows the general structure of the burnup process.

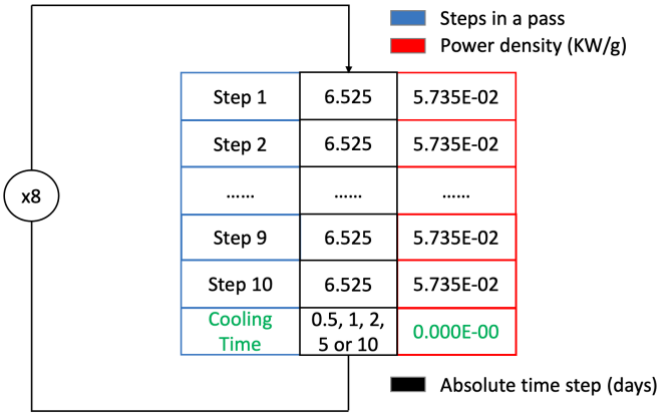

Figure 2. Illustration of the pebble lattice burnup process

The second phase of the burnup simulation ran a transport simulation using the isotope composition and parameters of each of the cycles and sub step. An outwards current tally was imposed on the reference pebble to obtain an estimate of the gamma source rate at the surface of the pebble. The energy grid of the imposed tally ranges from 0 MeV to 3 MeV with 512

uniformly spaced bins. Ten million particles were simulated for each step in the transport simulation that resulted in an uncertainty of less than 15 % in each bin.

Generation of Gamma Spectra

GADRAS is a general-purpose application for the modelling and analysis of radiation detector response, primarily gamma spectrometric instruments and neutron detectors based on proportional counters. To simulate detector response, a HPGe detector with 95% detection efficiency was selected in GADRAS. The detector was calibrated with spectroscopic pairs made up of ^{60}Co, ^{57}Co, ^{241}Am, ^{133}Ba, ^{137}Cs, ^{88}Y and ^{228}Th before spectra were generated. Also, no simulated background was injected into the spectra generated in all datasets. The default deadtime of 10 μs was used for all spectra dataset generated. The deadtime was corrected before the spectra were fed into algorithms.

Dataset for Development and Test of Machine Learning Method

For ML algorithm development and test, multiple gamma spectra datasets were generated for cooling time ranging from 0.5 days to 10 days and data acquisition time of 20 seconds and 3600 seconds. The longest cooling time of 10 days allows short-lived isotopes to decay to negligible level while the longest data acquisition time of 3600 s helps reduce the statistical errors, hence that dataset was used as baseline for this study. On the other hand, short cooling time of 0.5-2 days and data acquisition time of 20 s are assumed by designers in the current reactor designs. The 0.5-, 2- and 5-day cooling dataset contain 120 gamma spectra, the 1-day cooling dataset contains 104 gamma spectra and the 10-day cooling has 112 gamma spectra. The burnup level in each of these datasets ranges from about 2 MWD/kgU to 50 MWD/kgU.

Figure 3 shows examples of simulated spectra. The figure on the left shows the differences between 10 days and 0.5 days cooling time. The latter has significant photopeak's from short-lived fission products and a general higher pebble source activity. The figure on the right shows the spectra of 20-sec and 3600-sec acquisition time from a pebble after 0.5-day cooling time. It is worth noting that short acquisition time (e.g., 20 secs) and cooling time (e.g., 0.5-2 days) are preferred settings from operational point of view but will create very noisy spectra with many photopeak's from short-lived isotopes.

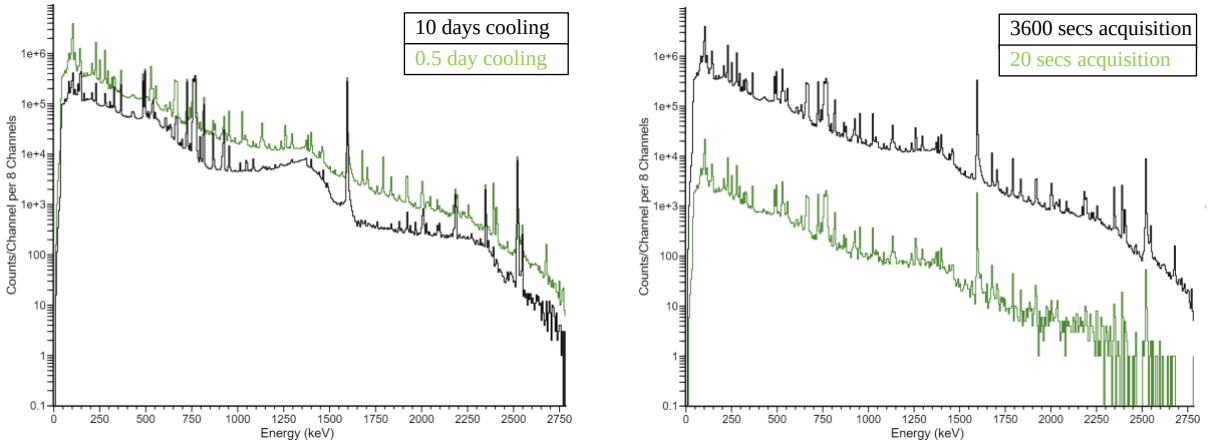

Figure 3. Similar residence and acquisition times (LEFT), Similar residence and cooling times (RIGHT)

ALGORITHMS FOR BURNUP PREDICTION

Linear regression baselines

This works was compared against a standard linear regression as a baseline method: a single-peak regression with the 662 keV photopeak for ^{137}Cs. The correlation between the height of this photopeak and the fuel burnup has been established, in particular for long cooling times. However, this work attempts to determine its effectiveness in predicting burnup values for shortened cooling times and measurement times – situations which introduce noise and quality issues to measured spectra.

ML models

The effectiveness of different machine learning (ML) techniques on improving the prediction accuracy and consistency of fuel burnup from measured spectra was explored. This work focused on three types of ML models at the beginning of this work, Multilayer Perceptron (MLPs), Convolutional Neural Networks (CNN) and Transformers.

The three ML model types differ in how they extract features from input data. Multilayer perceptron (also called fully connected networks, or feedforward neural networks) consist of several linear transformations separated by fixed nonlinear activation functions to support capture of nonlinear relationships, encoded in the weights of multiple network layers. At each layer, all the elements of the intermediate feature representation (starting with the raw spectra values for the first layer) are considered together in a single linear mapping to the next layer's feature representation (after a nonlinear activation), so the MLP model architecture extract global features (Ruck, Rogers and Kabrisky 1990). CNNs, on the other hand, use a number of convolutional kernels at each layer that scan over an entire input sequence, but share weights per layer, and so extract local features [e.g. (Kamuda, et al. 2020)]. In addition to nonlinear activation functions, CNNs employ pooling layers to iteratively increase the receptive field of the convolutional kernels, and so gather local features to build a global feature representation. The Transformer architecture (Vaswani 2017), on the other hand, uses a feature called multi-head attention to create multiple dynamic mappings between all elements of a feature vector, and also incorporates a positional encoding scheme to enable the network to build global feature vectors that simultaneously capture contextualized local information as well. This paper focuses on the performance of MLP models, as our work in evaluating CNN and Transformer architectures for this data is ongoing. All models were created using the PyTorch deep learning library, and hyperparameters were determined empirically on training data.

TEST RESULTS

The predictive accuracy of the baseline regression methods and the trained ML models across eight datasets, covering two measurement times (30 secs and 3600 secs), and four cooling times (0.5, 2, 5, and 10 days) were evaluated. Baseline and ML models were all trained on 80% of the data for each set, and evaluated on the remaining 20% of spectra, to properly measure regression/model generalization performance on unseen data. The full raw spectra were made available to the ML models (4096-dimensional), with differing binning rates allowed as a hyperparameter (along with standard model-specific hyperparameter choices). Data preparation consisted of extracting burnup levels from the SERPENT simulations and transforming the

GADRAS-produced N42 spectra files into single datasets suitable for ML model consumption (e.g. CSV files). Burnup levels are in units of MWD/kgU.

Regression Performance over Reduced Cooling and Acquisition Times

As anticipated, the performance of a standard linear regression over one or multiple photopeaks to predict fuel burnup is significantly affected by reduced cooling and detector acquisition time. Figure 4 shows this effect for linear regression on the 662-keV photopeak of ^{137}Cs. The two-performance metrics used mean average percentage error (MAPE) and coefficient of determination (R^2, which is the square of the Pearson correlation coefficient). In predictive regression analysis, R^2 is the most useful measure of predictive quality, and in particular measures the ability of a regression model to track or correlate well with true values, whereas MAPE measures the expected scale of prediction errors. Performance metrics are shown for an average of 10 randomly shuffled data samples for creating the 80:20 training-testing split.

As shown in the data plots, for sufficiently long acquisition times (1 hour), a simple linear regression can indeed predict fuel burnup reasonably well, with correlation scores over 0.99 for very long cooling times (10 days) and prediction errors of 9%. At shorted cooling times (0.5 days), correlation remains above 0.97 and the error scaling is approximately 17%.

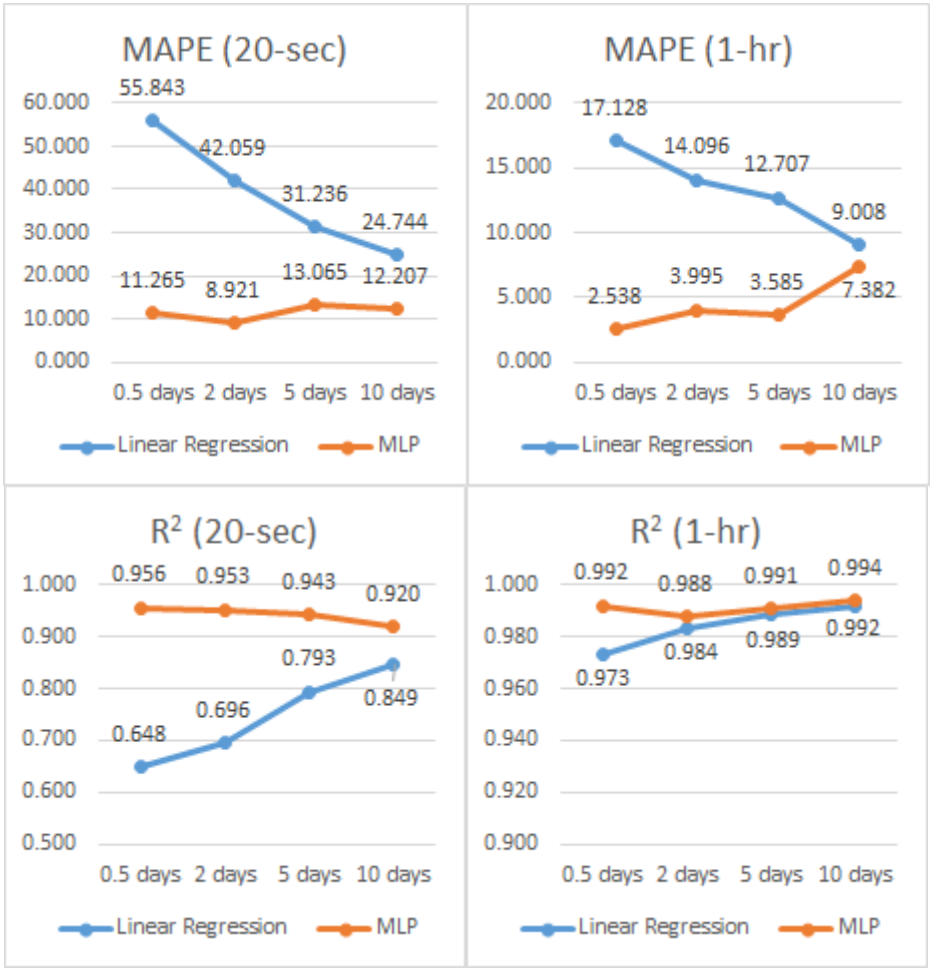

Figure 4. Predictive performance of standard linear regression and trained MLP model over different fuel cooling time and detector acquisition times.

MAPE (mean average precision error) is a measure of error scale (lower is better), while R^2 is a measure of correlation (higher is better).

However, as acquisition time is shorted (20 seconds), linear regression performance falls dramatically (note the difference in plot scales). Even at a long 10-day cooling time, correlation scores fall under 0.85, with errors around 25%, and at 0.5 days of cooling time, correlations fall further to 0.65 and errors climb over 55%. Figure 5 shows the degree to which predictions from this linear regression agree with the true burnup values. Although predictions roughly correlate with the ground truth burnups, the deviations and scale of prediction errors are very apparent. For even shorter cooling times (which may be desirable in PBR reactor designs), this performance can be expected to suffer further.

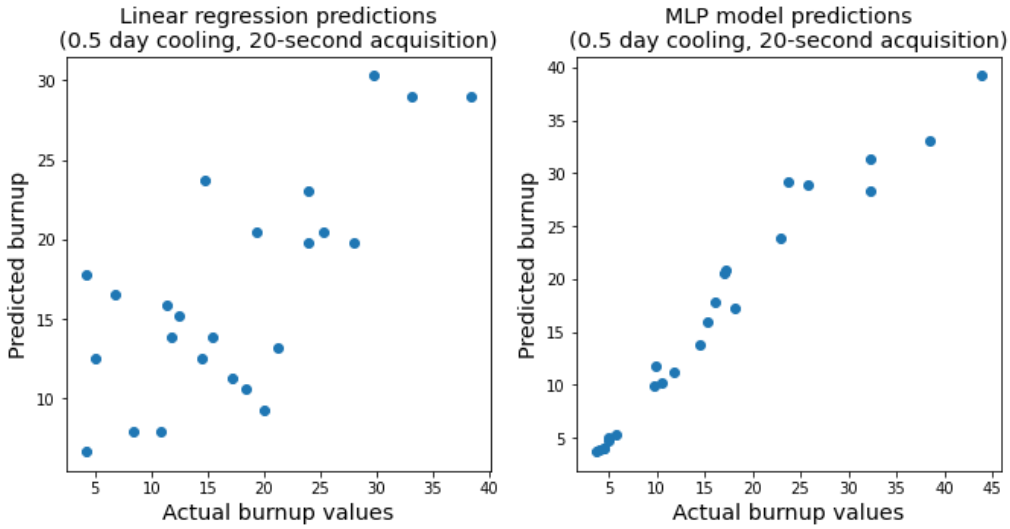

Figure 5. Ground truth vs predicted burnup for linear regression and trained MLP model on 0.5-day cooling and 20-sec acquisition time (the most challenging dataset). A perfect burnup prediction would correspond to all points laying along the main diagonal.

MLP burnup prediction

In stark contrast to the prior linear regression results, the predictive performance of the MLP models that were trained were only minimally affected by changes to fuel cooling and detector acquisition time, over the ranges tested. Figure 4 shows predictive performance of one MLP model trained over the same datasets as the prior linear regression. The model architecture used for this analysis was a 3-layer MLP with hidden layers of size 256 and 32; the standard PyTorch SGD (stochastic gradient descent) algorithm was used for training (for 200 epochs), with initial learning rate of 1e-3 and a 10x reduction every 50th epoch; momentum was enabled and set to 0.9; dropout was disabled. Input spectra were re-binned with a bin width of 32.

Not only does this MLP model significantly outperform the linear regression in all conditions, but the model in fact appears to be performance-saturated for the datasets tested. R^2 correlation between MLP model predictions and true burnup values are effectively level across different cooling times: ~0.99 for 1 hour acquisition times and ~0.95 for 20-seconds. Error rates remain

around 4% for l hour acquisitions and 10% for 20-seconds. In fact, performance in all metrics even appears to improve for shorter (i.e. more challenging) cooling times, particularly for the more challenging 20-second acquisition time datasets. However, this effect is likely an artifact of model parameters not yet being fully optimized; further ML model optimizations should effect additional improvements in all conditions, particularly the 'easier' datasets.

It is noteworthy that even in the most challenging measurement conditions, a relatively simple ML model architecture achieves exceptional agreement in its predictions with the true burnup values, far beyond what is possible with traditional photopeak regression. Figure 5 demonstrates this prediction quality on the most challenging measurement condition, showing the much-improved burnup prediction quality.

CONCLUSIONS AND DISCUSSIONS

Burnup measurement is a critical measure in MC&A. ML-based spectral analysis methods were proposed for estimation of the burnup of a pebble. The preliminary tests in this work showed that both ML-based methods and the photopeak-based linear regression method could achieve high accuracy when the gamma-ray spectra contained negligible background radiation caused by short-lived fission products and minimal statistical errors. However, such ideal condition is not achievable in the actual reactor operation. Under the conditions that the PBR designers are considering today, e.g., 2 days or less cooling time and 20-s acquisition time, the gamma spectra from burnup measurement became noisy. In that case, the proposed ML methods outperformed the conventional linear regression method significantly.

The ultimate goals of this work are to improve MC&A, reduce the operational burden and simplify PBR reactor designs by developing a high-performance ML algorithm and identify the optimal operational condition for deployment. It will allow stakeholders, such as regulators, PBR designers, and operators, to reach an achievable point in MC&A regulation related to spent fuel coming out of the PBR reactors. The preliminary results reported in this paper have demonstrated the promise. The datasets from the modeling and simulation allow further investigation to optimize the ML algorithm and identify the optimal deployment condition. Further results will be reported in future publications. An area of particular focus in our future work will be using the ML models we develop and train to help identify novel spectra features and relationships to aid in the interpretation and uncertainty quantification of ML model results.

ACKNOWLEDGEMENTS

The authors are grateful to Samuel E. Bays and Ryan Stewart from Idaho National Laboratory for their valuable suggestions in modeling pebbles and MC simulation in SERPENT. The authors also thank Per F. Peterson, Margaret Ellenson, Nicolas Zweibaum and Nader Satvat from Kairos Power LLC. for sharing their burnup simulation results for validation of the results in this paper.

The work presented in this paper was funded through the Advanced Reactor Safeguards (ARS) program, Office of Nuclear Energy U.S. Department of Energy. The manuscript has been authored by Brookhaven National Laboratory managed by Brookhaven Science Associates, LLC under Contract No. DE-SC0012704 with the U.S. Department of Energy.

REFERENCES

Akyurek, T, L.P. Tucker, and S Usman. 2014. "Review and charachterisation of the best candidate for isotopes for burnup analysis and monitoring of irradiated fuel." *Annals of Nuclear Energy* 69: 278-291.

Butler, Keith T. et al. 2018. "Machine learning for molecular and materials science." *Nature 559.7715* 547-555.

Carleo, Giuseppe, et al. 2019. "Machine learning and the physical sciences." *Reviews of Modern Physics 91.4* 045002.

Chersola, Davide, Guglielmo Lomonaco, Riccardo Marotta, and Guido Mazzini. 2014. "Comparison between SERPENT and MONTEBURNS codes applied to burnup calculations of GFR-like configuration." *Nuclear Energy and Design* 273: 542-554.

Horne, Steve, Gregory Thoreson, Lisa Theisen, Dean Mitchell, Lee Harding, and Wendy Amai. 2014. "GADRAS-DRF 18.5 User's Manual."

Kamuda, Mark, Jifu Zhao, Huff, and Kathryn. 2020. "A comparison of machine learning methods for automated gamma-ray spectroscopy." *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 954* 161385.

Leppanen, Jaakko. 2015. "Serpent - a Continous-energy Monte Carlo Reactor Physics Burnup Calculation Code."

Rawool-Sullivan, Mohini W., John K. Mattingly, Dean J. Mitchell, and Jesson D. Hutchinson. 2012. "Combined MCNP/GADRAS Simulation of HPGe gamma spectra." *IEEE Nuclear Science Symposium and Medical Imaging Conference Record* (Los Alamos National Laboratory).

Ruck, Dennis W., Steven K. Rogers, and Matthew Kabrisky. 1990. "Feature selection using a multilayer perceptron." *Journal of Neural Network Computing 2.2* 40-48.

Vaswani, Ashish, et al. 2017. "Attention is all you need." *Advances in neural information processing systems* 5998-6008.