 https://doi.org/10.3011/ESARDA.IJNSNP.2021.1

Investigating the sensitivity to irradiation history when predicting fuel parameters using random forest regression

Erik Branger, Zsolt Elter, Sophie Grape and Markus Preston,
Division of Applied nuclear physics,
Department of Physics and Astronomy
Uppsala University, Sweden
E-mail: erik.branger@physics.uu.se

Abstract

Safeguards verification of spent nuclear fuel assemblies is frequently done by performing non-destructive measurements, which are used to verify the completeness and correctness of operator declarations such as initial enrichment (IE), burnup (BU) and cooling time (CT) of the fuel. However, different irradiation histories may result in the same combination of CT, BU and IE, and such fuels may behave differently despite identically declared values. The goal of this work is to investigate what effect the irradiation history has on the ability to predict the fuel parameters using random forest regression.

Random forest regression models were trained to predict the fuel parameters IE, BU and CT based on combinations of radiation signatures calculated from a previously modelled Pressurised Water Reactor (PWR) spent nuclear fuel library. The radiation signatures studied were the relative gamma-ray activities of Cs137, Cs134 and Eu154, their total gamma-ray activity, the total neutron emission rate and the parametrised early die-away time τ from the Differential Die-away Self Interrogation (DDSI) instrument. The performance of the models were tested on simulations of 2192 PWR fuel assemblies from the Ringhals 3 and 4 nuclear power plants in Sweden, which were simulated based on their documented irradiation histories.

Despite significant differences in irradiation history between the training and testing data sets, the Ringhals assembly parameters could be predicted with similar accuracy as for assemblies in the training set. The relative gamma-ray activities were sufficient to predict the CT with an RMSE of 2 years, and adding a total gamma or total neutron signature allowed the BU to be predicted with an RMSE of 1.4 MWd/kgU. The DDSI early die-away time τ enabled an accurate IE prediction, with an RMSE of 0.16 w%. The differences between irradiation histories introduced a systematic bias where CT was overestimated by about 1 year and the BU by about 1.5 MWd/kgU.

Keywords: Nuclear safeguards, fuel parameter prediction, machine learning, random forest regression, irradiation history

1. Introduction

One of the many tasks undertaken by international nuclear safeguards inspectors is the verification of spent nuclear fuel (SNF) assemblies. Such verifications are done both to verify that the assemblies do indeed contain nuclear material (gross defect verification) and that parts of the fuel assembly have not been diverted (partial defect verification). However, due to the intense radiation emission from fission products and minor actinides, a direct determination of the fissile content is challenging. As a pragmatic solution, most verifications are done using non-destructive assay, aimed at verifying that operator-declared fuel parameters, such as cooling time (CT), burnup (BU) and initial enrichment (IE) are consistent with the measured radiation emissions. Computer codes are then used to estimate the fissile material inventories of the fuel assemblies using CT, BU and IE. These results are combined with other safeguards relevant information to evaluate the completeness and correctness of declarations and compliance with international non-proliferation treaties.

Before SNF assemblies are placed in difficult-to-access storage, such as dry storage or deep geological repository, the completeness and correctness of the operator declarations must be verified to high accuracy and precision, since it may not be possible to re-verify the data after storage. Traditionally, inspectors select an instrument that can measure the fuel assembly inventory at sufficient accuracy and precision, and bring the instrument to the fuel storage to perform a verification campaign. However, for a more thorough analysis, such as before encapsulation, it may be necessary to combine data from multiple instruments that are sensitive to different physical properties, in order to verify fuel parameters and correctness of declarations to the best possible accuracy and precision. Although potential measurement systems which could be used for safeguards measurements have been investigated previously [1], these investigations often focused on what information a single system can provide. This work aims at investigating how to combine different

 Branger, E., Elter, Z., Grape, S., & Preston, M. (2021). Investigating the sensitivity to irradiation history when predicting fuel parameters using random forest regression. *ESARDA Bulletin - The International Journal of Nuclear Safeguards and Non-proliferation*, 62, 2-12.
https://doi.org/10.3011/ESARDA.IJNSNP.2021.1

measurements using machine learning to extract more information from the measurements.

Due to the complex interplay between fuel usage in a reactor and the fission product and minor actinide abundance, machine learning has been investigated in the past few years for interpreting the data in a systematic way, and to extract safeguards-relevant information. Machine learning has been applied to gamma spectroscopy data from spent nuclear fuel [2,3,4], and for predicting fuel parameters based on several types of measurements [5,6]. It has also been used for partial defect detection purposes in spent fuel [7,8,9], for process monitoring at reprocessing facilities [10] and to classify uranium oxide fuels and mixed uranium oxide fuels [11].

This work builds upon the work of [5] and investigates the capability of a random forest (RF) regression model to predict CT, BU and IE for the complete modelled fuel inventory of the Ringhals 3 and 4 PWR reactors in Sweden. The objective is to investigate if a RF regression model trained on SNF having a simplified irradiation history (as done in [5]) generalizes to a realistically modelled SNF inventory, representing a fuel inventory to be placed in a final repository. In order for a regression model to be useful in safeguards, it must be able to reliably predict the CT, BU and IE for a real inventory.

The motivation behind investigating the impact of the irradiation history is that although the radionuclide composition of a SNF is predominantly coupled to the CT, BU and IE of the assembly, the irradiation history may also influence the fission product and minor actinide abundance [12,13]. For fuel assemblies with more than a few years CT, which is the topic of this work, the gamma emission is dominated by Cs134, Cs137 and Eu154 [12]. Due to its half-life of 30.2 years, which is typically longer than the time the fuel assembly is in the reactor, Cs137 is often considered to build up linearly with BU. However, for very long gaps in irradiation, on the order of decades, Cs137 created in the cycles before the gap will have had noticeable time to decay during the gap. Cs134 is created through neutron capture by the direct fission product Cs133, hence its production depends strongly on the neutron flux in the reactor core, and thus the reactor power. Due to its shorter half-life of 2.1 years, even a one-year gap in irradiation will allow Cs134 produced in earlier cycles to decay noticeably. For Eu154, its production path is more complicated, and a fuel depletion calculation using the irradiation history is required for accurate results. With respect to neutron emission, the build-up of the principal neutron-emitting radionuclides depends strongly on the total neutron fluence but is relatively insensitive to the power level. However, the rate of build-up is significantly affected by the initial U235 content and any gaps in the irradiation, though the effect is the most significant at short CTs [12].

2. Methodology

To investigate what effect the irradiation history has on the fuel parameter predictions when using RF regression models, the spent fuel inventories of the Ringhals 3 and 4 PWR reactors were modelled. Section 2.1 explains how the spent fuel modelling was done, section 2.2 describes the non-destructive assay signatures considered, and section 2.3 provides an overview to the RF method that was used to predict the fuel parameters.

Due to the optimized usage of the nuclear fuel at nuclear power plants, many fuel assemblies experience a similar irradiation history. The achievable discharge BU strongly depends on the IE, where an increasing IE enables a higher discharge BU, and possibly more cycles spent in the reactor. Modern fuel assemblies, i.e. those with a short CT, tend to have a higher IE compared to older ones, and thus a higher discharge BU [14]. Since a fully-burned fuel inventory from a commercial nuclear power plant represents a limited set of combinations of CT, BU and IE values, it cannot be used to train a model that should be able to predict all practically achievable fuel parameter values. Furthermore, although the majority of fuel assemblies reach their intended terminal BU upon discharge, some assemblies are discharged earlier, and have values of CT, BU and IE that significantly differ from the majority of the other assemblies. A reliable and robust model should be able to predict the parameters also in such cases.

The RF regression models in this work were trained on the simulated spent fuel library of [15], which covers fuel assemblies with CTs between 0 and 70 years, BUs between 5 and 70 MWd/kgU, and IEs between 1.5% and 6%. Different RF regression models were trained to predict CT, BU and IE. The performances of these models were evaluated on two different test data sets: one test data set from the same fuel library, with fuel assemblies having the same irradiation history as in the training data, and one data set comprising the modelled Ringhals fuel assemblies. By comparing the performance of the RF regression models on these two test sets, it is possible to identify what uncertainties are due to the models themselves and assumptions underlying the fuel depletion calculations, and what additional uncertainty is added in the predictions due to the irradiation history in the Ringhals case. The values of CT, BU and IE from the Ringhals fuel assemblies all fall within the ranges of the parameters in the fuel library used for training, hence the RF models are trained on data that covers all Ringhals combinations of CT, BU and IE. The regression and evaluation strategy used in this work is summarized in Figure 1.

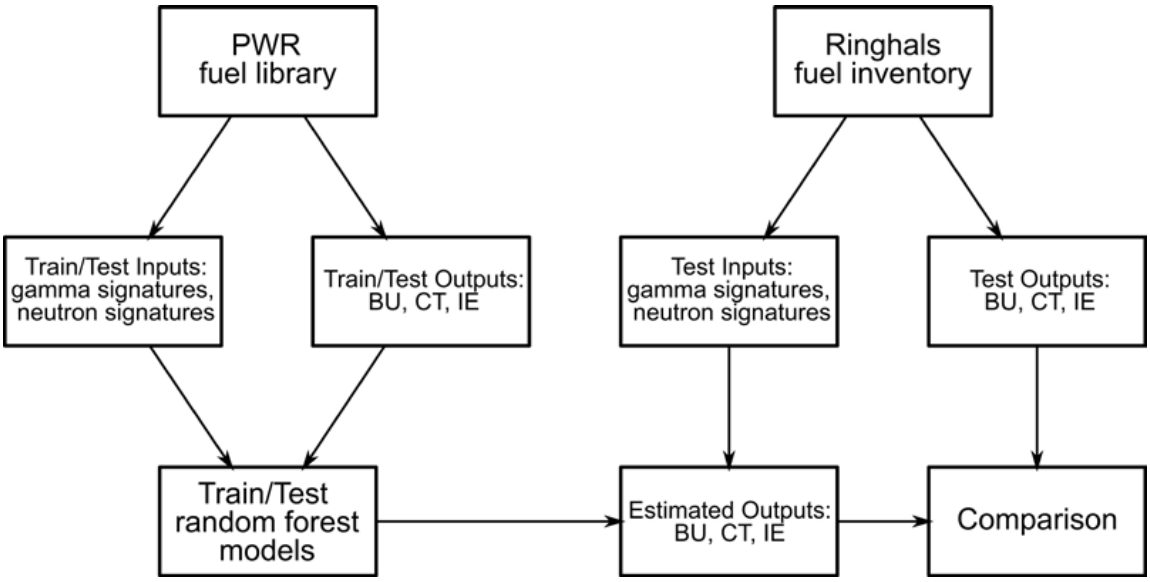

Figure 1: The training of the RF regression models was done using 80% of the fuel library of [15]. The remaining 20% was used to test the performance of the models and investigate what uncertainties are inherent due to the models and underlying assumptions. The trained model performance was also tested on the Ringhals fuel inventory, to assess the performance for fuel assemblies with real irradiation histories.

2.1 Spent fuel modelling

The spent fuel library of [15] contains in total 789 406 uranium dioxide (UO2) fuel samples, and was created using Serpent2 [16]. In the creation of the library, a generic irradiation history was assumed. The fuel was simulated to experience a constant power level, and the desired BU was obtained by increasing the number of irradiation cycles, and adjusting the length of the last irradiation cycle to obtain the desired burnup. The irradiation history in the library also assumed that after a period of 365 days of irradiation, a 30-day cooling period followed. This approximately corresponds to the revision period at the Swedish nuclear power plants.

For the Ringhals 3 and 4 fuel assemblies, the fuel assembly information provided by the operator Vattenfall includes the IE, the start and end dates of each irradiation cycle, and the BU of each cycle. Using this data, the fuel depletion of each fuel assembly was simulated in ORIGEN [17], due to its efficiency since all assemblies had to be simulated individually. The result of the ORIGEN simulations is an estimate of the material composition and neutron emission of each SNF assembly. The radionuclide content was then converted to a corresponding gamma-ray activity using the nuclide half-lives. For this work, it is assumed that the radionuclide gamma-ray activities can be assessed from a gamma-spectroscopic measurement, hence that the RF regression models can be trained using the gamma-ray activities as input features.

The Ringhals data includes the complete SNF inventory of the two reactors from the start of the reactors in 1980 and 1983 until the year 2012, and the fuel radionuclide abundances were calculated to correspond to 1st of July 2020. In this way, the modelled assemblies have a minimum of 8 years of cooling, enabling an investigation of the fuel parameter prediction capability for medium- and long-cooled fuel, which were shown in [5] to be more challenging than fuels with shorter CTs. Additionally, in the context of verification before encapsulation and final storage, many of the fuel assemblies are expected to have relatively long CT to ensure a sufficiently low residual heat. The Ringhals 3 fuel assembly data set contains 1083 assemblies, and the Ringhals 4 set contains 1109 assemblies, for a total set of 2192 assemblies. Due to the potentially sensitive nature of this data set, it cannot be published, and we focus on general results and trends that may be of relevance to PWR reactors in general, while keeping the specifics at a minimum.

Investigating the Ringhals fuel assembly irradiation histories, a few general remarks can be made:

- The PWR fuel library assumes a constant power level, whereas most Ringhals fuel assemblies experienced a roughly 20-40% higher power level in the first one or two cycles, as compared to the remaining cycles. As a consequence, we expect that for the Ringhals assemblies the modelled activity of short-lived radionuclides such as Cs134 is lower at discharge as compared to the fuel

library for the same CT, BU and IE values, since they were produced earlier in the Ringhals case and have had more time to decay before discharge.

- The majority of the Ringhals assemblies experienced a very regular irradiation, without spending cycles outside the reactor. However, a significant fraction of the Ringhals assemblies were reinserted into the reactor for a final low-power irradiation cycle, typically after having spent one or two cycles outside the reactor. For the assemblies that were part of the first core loading, the gap to the final cycle could be significantly longer, up to ten years, and the final cycle could be of comparatively high power. A smaller fraction of the fuel also had a gap in the irradiation after the first or second cycle. This is a noticeable difference to the fuel library, which has no gaps in the irradiation, and this affects the abundance of short-lived fission products at discharge.
- In some of the Ringhals assemblies, burnable absorbers were initially present. Such absorbers are not included in the PWR fuel library. Since the provided operator data on burnable absorbers is incomplete, and the absorbers mainly affects the beginning of the first cycle, they were not modelled. Note also that for depletion calculations, the neutron flux is set to yield the desired power level, which in part compensates for the effect of the burnable absorbers.

A summary of the CT, BU and IE of the Ringhals fuel assemblies is shown in Figure 2.

2.2 Non-destructive assay signals considered

This work considers several of the non-destructive assay signatures used in [5], such as the relative gamma-ray activity of selected abundant radionuclides, the total gamma-ray activity of the selected radionuclides, and the parameterised early differential die-away time τ. Since [5] found that the total Cherenkov light emission carries the same information as the total gamma-ray signature, we use only the total gamma-ray activity, since it does not require an additional measurement instrument. In addition, we include the total neutron emission rate of the fuel as a new signature, to investigate what impact it has on the model capability of predicting the fuel parameters.

Since the fuel assemblies considered here have a minimum CT of eight years, and a maximum CT of almost 40 years, the gamma-ray activity is predominantly caused by Cs134, Cs137 and Eu154, which therefore are the radionuclides considered in this work. These radionuclides are all abundant in SNF and have a long enough half-life to be measurable after more than eight years. For the Ringhals assemblies, these three radionuclides account for more than 99% of the total gamma-ray activity, and other radionuclides can therefore be neglected with a minimal loss of accuracy. For the relative gamma-ray activity, the sum of these three gamma-emitting radionuclide activities were scaled to 1. This corresponds to a measurement where an absolute calibration has not been made, and only the relative intensities of the gamma-ray emissions can be determined. The total gamma activity is the sum of the three gamma-ray activities, and although it does not correspond to the absolute intensity for the fuel assembly, it is proportional to it. This in turn, enables a comparison of the total gamma-ray activities between fuel assemblies.

As in [5], a minimum threshold activity was included, at 0.1 % of the lowest Cs137 activity in the training dataset. The activities of radionuclides below the threshold were set to 0. The value of 0.1% is arbitrary, but reflects the fact that radionuclides with low activities may fall below the threshold of detectability. Based on [5], it is expected that Cs134 (with half-life 2.1 years) is only measurable for a short while after eight years of cooling using this threshold, and Eu154 (with half-life 8.6 years) may be below the threshold for long-cooled, low-BU assemblies. Note however that the real threshold of a measurement depends on both the fuel assembly and the measurement

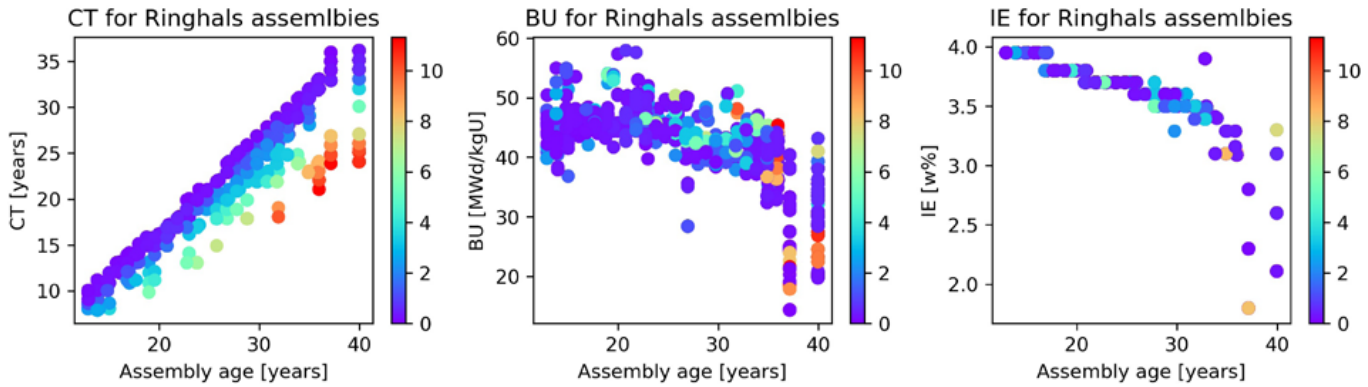

Figure 2: The distribution of CT, BU and IE for the Ringhals 3 and 4 spent fuel assemblies. The colour of the markers indicates the cumulative number of years an assembly spent outside the reactor between consecutive irradiation cycles. The assembly age is the number of years since first irradiation.

setup. Since this work does not consider the measurement setup at all, a global threshold is a simple way to include the main effect of a detectability limit on the subsequent analysis. For shorter CTs, additional radionuclides are expected to contribute and their activity will depend even more strongly on the power level of the final irradiation cycle and hence the irradiation history [12], but such investigations are outside the scope of this work, and such low-CT assemblies are not part of the Ringhals test set.

The total neutron emission rate of the Ringhals assemblies were provided by the ORIGEN simulations, which includes spontaneous fission and (α, n)-reactions. For the fuel database of [15] which was generated using Serpent2, the total neutron emission rate is not provided. The SF neutron emission rate was instead calculated based on the abundance of fissile radionuclides and minor actinides, and neutron emission data for these radionuclides from [12]. The (α, n) neutron contribution was calculated based on the abundance of fissile radionuclides and minor actinides and data from [18]. The total neutron emission rate was calculated as the sum of these two contributions. In general, for most Ringhals fuels the (α, n)-reactions contribute with about 1-5% of the total neutron rate, however for certain long-CT, low-BU assemblies the contribution could reach 20%. In the fuel library, the (α, n) contribution can exceed 50% for very low-BU and long-CT assemblies.

Since the neutron emission rates were calculated using different methods for the two fuel sets, and since the burnup calculations and underlying cross-sections also differ, additional errors and uncertainties arise when comparing the two data sets. A thorough benchmark of the two methods is outside the scope of this paper, but eight Ringhals assemblies with varying irradiation histories were selected and depleted with the real irradiation history using Serpent. The results indicate that the Serpent neutron emission rates were 2-12% higher than the ORIGEN emission rates, primarily due to a higher Cm244 production, which may introduce a bias in the results. However, since the neutron emission rates from the Ringhals fuel assemblies span more than two orders of magnitude, the bias is expected to be modest in the RF regression models.

To assay the fissile content of a SNF assembly, this work includes the signature from the DDSI instrument, which measures neutrons in coincidence to determine a so-called early die-away time τ. This feature is sensitive to the fissile content, making it useful for IE determination using machine learning [5]. To predict the early die-away time for a large number of fuel assemblies, the parameterization function of [19] was used, using an updated set of fit coefficients valid for a larger range of fuel parameters. For fuel assemblies with more than a few years cooling, the updated parameters give comparable results to the original parameters, thus either sets of parameters can be used. To verify the accuracy of the parameterization function, MCNP simulations were run for eight selected Ringhals fuel assemblies with irradiation histories that differ significantly from the one assumed in the parameterization function. The results show that the simulated and the parameterized τ values are within 1.5% of each other, with the exception of one fuel assembly for which τ differs by almost 4%. Hence, we judge the parameterized τ values to be sufficiently accurate for the purpose of this work.

Each signature in the training set, the fuel library test set and the Ringhals test set had a 1% Gaussian noise added to it, to account for measurement uncertainty. It is noted in [5] that such a low uncertainty is not unfeasible for gamma spectroscopy or the DDSI signal τ, however the actual measurement uncertainties will depend on the selected measurement setup, which is not considered in this work. The effect of higher levels of noise is presented in section 3.4.

2.3 Random Forest Regression

The RF method [20] is a further development of the decision tree method, and can be used either for classification or regression, i.e. predicting the value of a continuous variable. The RF method is a supervised learning algorithm, where the model is trained on known input-output pairs. The key improvement of RF over decision trees is that the RF predictor is made up of a number of trees, each being trained on a unique randomly sampled subset of the total training data. The output of the RF regression model is the mean value of the output of all decision trees. By using multiple trees, the RF method becomes less prone to overfitting compared to decision trees, where overfitting means a poor capability to generalize beyond the training data. The RF regression model implementation used in this work is the one from scikit-learn [21].

To control how the RF regression model is trained, two hyper-parameters were in [5] found to have an impact on the model performance. The first is the number of decision trees, and [5] notes that a wide range could be used with similar results. Here, we chose 250 trees as the default number if nothing else is specified. The second parameter is the number of features used in each node of the decision trees to make a decision about which leaf the input data belongs to. Again, [5] notes that several values give similar results, and this work use a default value of 3 in the analysis. This is large enough to be within the previously found minima, and additionally the models trained in this work always contains at least three input features.

Three data sets are typically used in machine learning regression: a training set, a validation set, and a testing set.

The training set is used to provide known input-output pairs to train the model, where the input here is some combination of the gamma-ray activities, the neutron emission rate and τ, and the output is the predicted values of CT, BU and IE. The validation set is used to tune the performance of the model when using different hyper-parameters. The testing set is used to assess the performance of the trained model on new data. Since it was found in [5] that the hyper-parameter optimum is rather broad, tuning the model using the validation set was omitted in this work. However, this work instead considers two test sets: one test set comprising data from the same fuel library used to train the model, and another test set comprising the modelled Ringhals fuels. These two test sets are used to assess what uncertainties arise from the RF regression model itself and assumptions made in the depletion calculations, and what additional uncertainty is introduced by the actual irradiation history introduced in the Ringhals test set. The training set consists of 631 525 unique samples from the fuel library, the fuel library test set consists of the remaining 157 881 fuel library samples, and the Ringhals test set consists of the 2192 samples from both reactors.

As was done in [5], all input features in the training set were rescaled to have a mean value of 0 and a variance of 1 before the analysis. The same scaling was applied to the training set and both test sets, to ensure that all values after rescaling were directly comparable to each other. Since the fuel library and the modelled Ringhals assemblies were modelled using different depletion codes, the ORIGEN data for the Ringhals assemblies were converted from emissions per ton uranium to emissions per cm^3 to match the Serpent data in the fuel library, before the standard scaler was applied.

3. Results

In this section we present prediction results for the two test sets. For a given set of input signatures, three RF regression models were trained: one for predicting CT, one for BU and one for IE. Each model used the same training data and hyper-parameters. Following the work of [5], we first considered different combinations of gamma-ray signatures, and then added neutron signatures to the analysis.

3.1 Analysis using gamma-ray signatures

RF regression models were trained to predict IE, BU and CT for two scenarios: i) using the relative gamma-ray activities of Cs134, Cs137 and Eu154 as input, and ii) using in addition the total gamma-ray activity as input.

Table 1 reports the average and root mean square error (RMSE) of the difference between the predicted and true values of the CT, BU, and IE for the two test sets. The average difference (or error) indicates if there is a systematic bias in the predictions and reflects the accuracy of the models. The RMSE provides an indication of the precision. Figure 3 shows the predicted parameter values versus the true values for scenario ii). The marker colour shows the cumulative outage time between consecutive irradiation cycles, to highlight assemblies that spent cycles outside the reactor before reinsertion. Figure 3 also shows a histogram of the errors in the predicted CT, BU and IE for the fuel library test set and the Ringhals test set.

The overall results in Table 1 match those found in [5]. Using only the relative activities of the selected radionuclides, CT can be predicted with good accuracy and precision. However, fuel assemblies with a gap in their irradiation history are systematically over-predicted, as shown in Figure 3. Table 1 also shows that the RMSE for the CT prediction of the fuel library test set is higher than the RMSE for the Ringhals test set, which can also be seen in the histogram in Figure 3. The cause is that the training data includes low-BU high-CT fuels, where both the Cs134 and Eu154 activities are below the detectability threshold, and only Cs137 remains. Almost all Ringhals assemblies have reached a high BU at discharge, and are more likely to include these two radionuclides with activities above the detection threshold. There are however some long-CT, high-BU assemblies in the Ringhals test set, which are under-predicted, as shown in the CT plot in Figure 3.

The large bias and RMSE in the BU reported in Table 1 reveal that the BU of the Ringhals assemblies cannot be

Features	Data set	CT [days]		BU [MWd/kgU]		IE [w%]	
		Avg. error	RMSE	Avg. error	RMSE	Avg. error	RMSE
Relative radio- nuclide activities	Fuel library	3.3	2159	0.03	12.4	0.03	1.2
	Ringhals	425	608	8.88	12.1	0.35	1.13
Relative radio- nuclide activities and total gamma	Fuel library	1.2	683	0.001	0.90	0.001	0.92
	Ringhals	227	448	0.66	1.39	0.42	0.94

Table 1.Average error and RMSE in the fuel predictions for the fuel library test set and the Ringhals test set, with different input features considered in the analysis. The uncertainties in the values due to randomness in the training of the regression models is around 1% of each value.

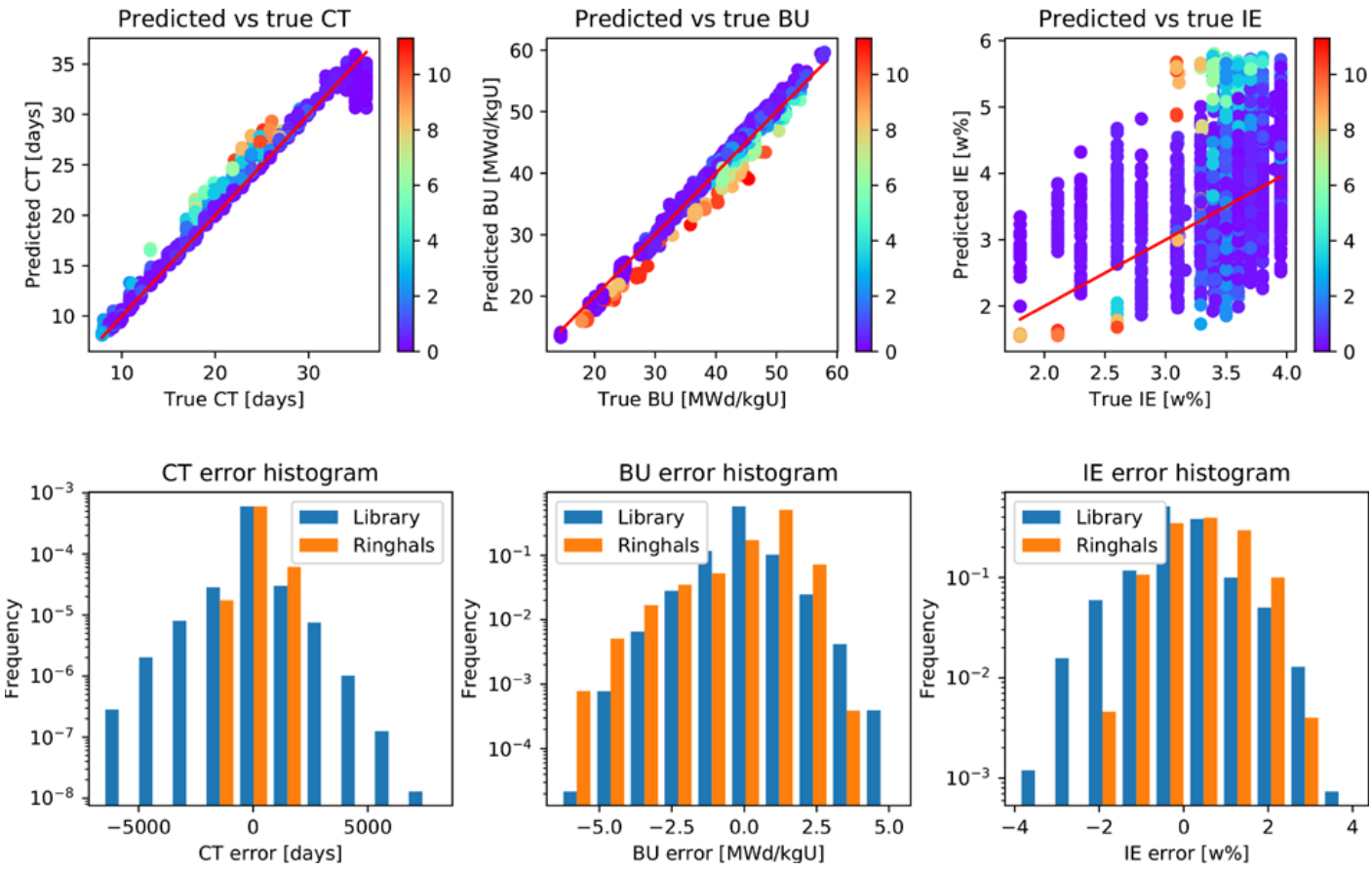

Figure 3. Performance of the RF regression models trained on the relative activities of selected radionuclides, and the total gamma-ray activity. Top row: the true and predicted values of CT (left), BU (centre) and IE (right), for the models. The red line is a guide for the eye, indicating where the regression matches the real values. The colour of the markers shows the cumulative number of years of cooling between consecutive irradiation cycles. Bottom row: Histograms of the errors in predictions of CT (left), BU (centre) and IE (right). The training errors were calculated based on the fuel library test set (labelled "Library"), and the Ringhals test set (labelled "Ringhals").

predicted by only considering relative gamma-ray activities. When the total gamma feature is included, this bias and spread is largely eliminated. The BU plot in Figure 3 shows that some assemblies are under-predicted by around 2-4 MWd/kgU, and these are the ones with a significant gap in their irradiation history. This group of fuels also causes the skewed error distribution seen in the error histogram plot.

Finally, the IE predictions fail since none of the gamma-ray activities depend strongly on it. Additional input features are needed to predict IE, such as neutron-based features, described in the next section.

3.2 Analysis using gamma-ray and neutron signatures

Two different neutron signatures were considered: the gross total neutron emission rate and the parameterized early die-away time τ from the DDSI instrument. RF regression models were trained for two sets of input features: i) using relative radionuclide gamma-ray activities, the total gamma-ray activity and total neutron rate, and ii) using in addition τ.

Table 2 shows the average error and the RMSE for the predictions of CT, BU and IE as a function of input features for the two test data sets. Figure 4 shows plots of the predicted and true parameter values. The colour of the markers shows the cumulative outage time in between irradiation cycles. Figure 4 also shows histograms of the errors in the predictions.

As can be seen when comparing Table 1 and Table 2, including the total neutron emission rate as an input feature improves the CT prediction slightly, the BU prediction remains unchanged, and the IE prediction is significantly improved. For the fuel library test set, the IE predictions are not significantly improved by adding τ when the total neutron emission rate is available. However for the Ringhals test set the precision of the IE predictions are noticeably improved by adding τ, although the accuracy is somewhat worsened due to adding a systematic deviation in the predictions. Hence, for the more realistic

Features	Data set	CT [days]		BU [MWd/kgU]		IE [w%]	
		Avg. error	RMSE	Avg. error	RMSE	Avg. error	RMSE
Relative radio-nuclide activities, total gamma and total neutron	Fuel Library	-0.18	157	0.001	0.53	0.001	0.12
	Ringhals	271	380	0.74	1.40	0.05	0.23
Relative radio-nuclide activities, total gamma, total neutron and τ	Fuel Library	0.16	79	0.001	0.52	0.001	0.11
	Ringhals	270	378	0.72	1.33	0.08	0.16

Table 2. Average error and RMSE in the fuel predictions for the fuel library test set and the Ringhals test set, with different input features considered in the analysis. The uncertainty in the values due to randomness in the training of the regression models is around 1% of each value.

Ringhals test set, τ provides valuable information useful in determining IE.

The plots in Figure 4 show that the bias in the predicted CT for long-cooled Ringhals fuel assemblies, caused by a lack of data at long CT, is removed when additional input features are used. However, the over-prediction of CT and under-prediction of BU for Ringhals assemblies having a gap in their irradiation history remain. This group of assemblies also causes the error histogram plots in Figure 4 to be skewed, where a symmetric distribution would be expected if the errors were random. Additionally, Table 2 shows that the RMSE values for the predictions of the Ringhals test set tend to be a factor 2-5 higher compared to the predictions of the fuel library test set, and this additional uncertainty is also introduced by the irradiation history. Furthermore, the histograms in Figure 4 show that the most likely error in CT is around 1 year, and the most likely error in BU is around 1.5 MWd/kgU. This bias is caused by the differing irradiation histories, and that the final low-power cycles in the Ringhals case resulted in a different abundance of radionuclides at discharge, as compared to a fuel assembly from the library with identical values of CT, BU and IE. Hence, the irradiation history introduces both a bias and an uncertainty in the predictions, when the models are trained on a simplified irradiation history.

3.3 Hyper-parameter selection considerations

Since there are differences in irradiation history as well as in the software used to produce the data sets, it must be verified that the RF regression models are not fitting to features in the training data set that do not generalize to the Ringhals test set. Such a lack of generalization could be due to the choice of hyper-parameters. For the hyper-parameter study, we used all six features (three relative radionuclide activities, their total gamma-ray activity, total neutron emission rate and τ) to predict the fuel assembly parameters. The two hyper-parameters that were investigated here were the number of decision trees used by the RF regression models and the number of features used in each node to make the decision. The default values used in the previous section was 250 trees per RF, and splitting on three features.

For the number of trees, values in the range 50 to 300 in steps of 50 were evaluated. The results show that just as in [5], changing the number of trees has little impact on the performance of the RF regression models. Considering the uncertainty introduced by the randomness in the data sampling, different values of the hyper-parameters do not result in significant differences in fuel parameter prediction capability.

For the number of features used to split each node in the decision tree, values between 1 and 6 were evaluated. For BU and IE, the results are again similar to [5], with little change in performance as a function of the number of features. For CT however, the results improves slightly for the fuel library test set with increasing number of features, but the corresponding performance for the Ringhals test set is a worsened performance with an increasing number of features. This suggests that the RF regression model predicting CT may be fitting to structures in the fuel library that do not generalize to the Ringhals test set, and that splitting using fewer features may be preferred. However this needs to be verified using other test sets before the choice of splitting on one feature can be selected as default.

3.4 Noise considerations

In the previous analysis, a default 1% Gaussian noise was added to all input features in both the training and two test sets, to include the effect of counting statistics and measurement uncertainties. However, depending on the measurement situation, the level of noise may vary. To investigate the impact of also other levels of noise, RF models were trained and then tested when 1%, 5% and 10% noise was applied to all data, with all features used for training. The results are shown in Table 3.

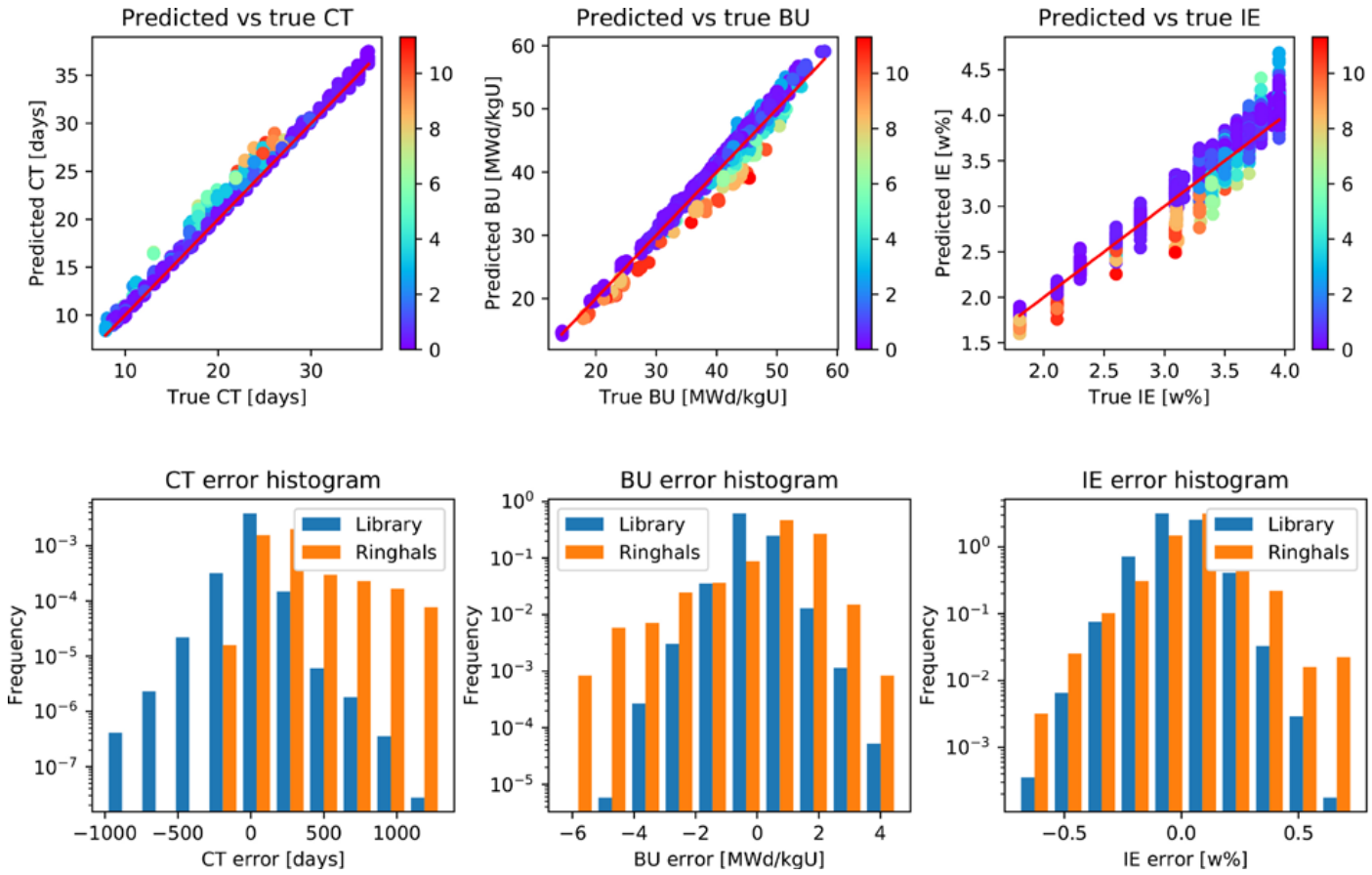

Figure 4. Performance of the RF regression models trained on the relative activities of selected radionuclides, the total gamma-ray activity, the total neutron emission rate and τ. Top row: the true and predicted values of CT (left), BU (centre) and IE (right), for the models. The red line is a guide for the eye, indicating where the regression matches the real values. The colour of the markers shows the cumulative number of years of cooling in between irradiation cycles. Bottom row: Histograms of the errors in predictions of CT (left), BU (centre) and IE (right). The training errors were calculated based on the fuel library test set (labelled "Library"), and the Ringhals test set (labelled "Ringhals").

Overall, the behaviour of the fuel library test set with increasing levels of noise is similar to [5], where the average error changes little, and the RMSE increases with increasing noise. For the Ringhals test set, the RMSE values are higher at lower noise, but the increase as a function of noise is similar to the fuel library test set. Thus, the noise appear to have comparable effects to the precision of the predictions for both test sets. For the Ringhals test set, the average error however increases somewhat with noise. In general, as in [5], the CT prediction is not too sensitive to increased noise, the BU prediction is a bit more sensitive but is manageable also at higher noise levels, but the IE predictions are highly sensitive to the addition of noise. Hence for accurate fuel parameter predictions, effort should be made to have a well-characterized and low-noise DDSI measurement, or a total neutron emission measurement if the DDSI is not used.

4. Conclusions

In this work, we have trained RF regression models to predict the fuel parameters CT, BU and IE of modelled PWR fuel assemblies, based on non-destructive data that could be obtained through gamma and neutron measurements. The models were trained on modelled PWR assemblies from a fuel library with a wide range of CT, BU and IE values, which was created assuming a standardized, simplified irradiation history. The RF regression models were tested on both data from this library, as well as modelled PWR fuels from the Swedish commercial nuclear power reactors Ringhals 3 and 4, to investigate what impact a realistic fuel irradiation history has on the prediction capabilities of RF regression models. In the analyses, input features corresponding to relative radionuclide activities, total gamma-ray activity, total neutron emission rate and the parametrised early die-away time τ from the DDSI instrument were considered.

Based on the results, a gamma-spectroscopic measurement should be sufficient to allow a RF regression model

Noise level	Data set	CT [days]		BU [MWd/kgU]		IE [w%]	
		Avg. error	RMSE	Avg. error	RMSE	Avg. error	RMSE
1%	Fuel Library	0.16	79	0.001	0.52	0.001	0.11
	Ringhals	270	378	0.72	1.33	0.08	0.16
5%	Fuel Library	-0.56	298	-0.001	1.94	0.001	0.36
	Ringhals	311	442	1.40	2.36	0.17	0.39
10%	Fuel Library	-1.32	509	-0.002	3.16	0.001	0.59
	Ringhals	359	541	2.04	3.84	0.30	0.66

Table 3. Average error and RMSE in the fuel predictions for the fuel library test set and the Ringhals test set, with different input features considered in the analysis. The uncertainty in the values due to randomness in the training of the regression models is around 1% of each value

to predict CT well, and a well-calibrated setup that also provides enough information to allow a comparison between fuel assemblies of the absolute total gamma-ray activity will improve the results further. For a good BU prediction, both the relative radionuclide activities and the total gamma-ray activity are required, whereas the total neutron emission rate signature can be added to slightly improve the predictions. For the IE predictions, τ is the input feature considered of highest importance, since it is the only one which probes the fissile content, and hence IE. While the total neutron was sufficient for IE predictions of the fuel library test set, the more realistic Ringhals test set showed that τ provides more information as compared to the total neutron feature. Using the total neutron emission rate instead of τ worsens the results, since the RMSE of the predictions increase. This is because the abundance of neutron-emitting radionuclides does not only depend on the IE but also on the irradiation history, which differs for the training and Ringhals test set.

Due to the differences in irradiation history, the simulated Ringhals fuel assemblies have comparatively lower abundance of radionuclides such as Cs134 at discharge, compared to fuel assemblies from the fuel library with equivalent values of CT, BU and IE. As a consequence, the RF regression models systematically overestimate CT and underestimate BU for the Ringhals assemblies. However, the systematic deviation is rather modest, typically around 2-5% of the values. Fuel assemblies that were outside the reactor for some time before a final low-power cycle are in general predicted to have a CT that is 1-3 years too long, even if the time spent outside the reactor before the final irradiation cycle could be up to ten years. Overall, the predicted fuel parameter values are rather similar to the true values and may therefore be of use to a safeguards inspector despite the fact that fuel irradiation histories may differ from what is assumed in the predictions. As in the previous work, the choice of hyper-parameters has a negligible impact on the performance of the predictions, though the CT predictions do seem to benefit from choosing a low number of features to split nodes on.

5. Acknowledgements

This work was supported by the Swedish Radiation Safety Authority (SSM) under contract SSM2017-5979 and SSM2020-996. The authors are grateful to Klaes-Håkan Bejmer and Andreas Lidén at Vattenfall for providing the irradiation history of the fuel assemblies for the Ringhals 3 and 4 reactors.

6. References

[1] S. Tobin, P. Jansson, Nondestructive assay options for spent fuel encapsulation. SKB technical report, TR-13-30. ISSN 1404-0344. March 2013. https://skb.se/publication/2697744/TR-13-30.pdf

[2] J. Coble, C. Orton, J. Schwantes, Multivariate analysis of gamma spectra to characterize used nuclear fuel, Nucl. Instrum. Methods Phys. Res. A 850, pages 18–24. 2017

[3] C. Hellesen, S. Grape, P. Jansson, S. Jacobsson Svärd, M. Åberg Lindell, P. Andersson, Nuclear spent fuel parameter determination using multivariate analysis of fission product gamma spectra. Annals of Nuclear Energy, volume 110, Pages 886-895. December 2017.

[4] Zs. Elter, L. Caldeira Balkeståhl, S. Grape, C. Hellesen, Nuclear safeguards verification of modelled BWR fuel using a multivariate analysis approach, in: IAEA Safeguards Symposium: Building Future Safeguards Capabilities, Nov 5–8, Vienna, 2018.

[5] S. Grape, E. Branger, Zs. Elter, L. Pöder Balkeståhl. Determination of spent nuclear fuel parameters using modelled signatures from non-destructive assay and Random Forest regression. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, Volume 969, July 2020. DOI: 10.1016/j.nima.2020.163979

[6] A. M. Bachmann, J. B. Coble, S. E. Skutnik, Comparison and uncertainty of multivariate modelling techniques to characterize used nuclear fuel, Nucl. Instrum. Methods Phys. Res. A 991 (2021) 164994.

[7] R. Rossa, A. Borella, Use of machine learning models for the detection of fuel pin replacement in spent fuel assemblies. The ESARDA Bulletin No.58. 2019.

[8] R. Rossa, A. Borella and N. Giani, Comparison of machine learning models for the detection of partial defects in spent nuclear fuel. Annals of Nuclear Energy, volume 147 page 107680. 2020.

[9] Zs. Elter, S. Grape, A methodology to identify partial defects in spent nuclear fuel using gamma spectroscopy data. The ESARDA Bulletin No. 61, 2020.

[10] C.R. Orton, C.G. Fraga, R.N. Christensen, J.M. Schwantes, Proof of concept simulations of the multi-isotope process monitor: An online, nondestructive, near-real-time safeguards monitor for nuclear fuel reprocessing facilities, Nucl. Instrum. Methods Phys. Res. A 629 (1) 209–219. 2011

[11] M. Åberg Lindell, Safeguards Evaluation and Analysis Techniques for Nuclear Material In Generation IV Fuel Cycles (PhD Thesis). Uppsala University, 2018.

[12] D. Reilly, N. Esslin, H. Smith, S. Kreiner. Passive Nondestructive Assay of Nuclear Materials. Technical report. LA-UR-90-732, Los Alamos National Laboratory, Los Alamos, NM, U.S.A. March 1991.

[13] A. Bolind, The use of the BIC set in the characterization of used nuclear fuel assemblies by nondestructive assay. Annals of Nuclear Energy, Volume 66, pages 31-50. 2014. https://doi.org/10.1016/j.anucene.2013.11.010.

[14] U.S. Energy Information Administration, Annual commercial spent fuel discharges and burnup, 1968 - June 30, 2013. Available at https://www.eia.gov/nuclear/spent_fuel/ussnftab3.php Retrieved February 9:th, 2021.

[15] Zs. Elter, L. Pöder Balkeståhl, E. Branger, S. Grape. Pressurized water reactor spent nuclear fuel data library produced with the Serpent2 code. Data in Brief, Volume 33, December 2020. DOI: 10.1016/j.dib.2020.106429

[16] J. Leppänen et. al. The Serpent Monte Carlo code: Status, development and applications in 2013. Annals of Nuclear Energy, volume 82 pages 142-150. 2015.

[17] S. M. Bowman, L. C. Leal, O. W. Hermann, C. V. Parks. ORIGEN-ARP, A Fast and Easy-to-Use Source Term Generation Tool. Journal of Nuclear Science and Technology, vol. 37, no. sup1, pp. 575-579, 2000. DOI: 10.1080/00223131.2000.10874953

[18] Update of the α – n Yields for Reactor Fuel Materials for the Interest of Nuclear Safeguards. S. Simakov, Q. van den Berg, Nuclear data sheets, volume 139, January 2017. https://doi.org/10.1016/j.nds.2017.01.005

[19] L. Caldeira Balkeståhl, Zs. Elter, S. Grape. Parametrization of the differential die-away self-interrogation early die-away time for PWR spent fuel assemblies. Presented at the 41st ESARDA Annual Meeting in Stresa, Italy, May 2019. Published in the ESARDA bulletin, no. 58, June 2019.

[20] L. Breiman, Random forests, Mach. Learn. 45 pages 5–32. 2001

[21] F. Pedregosa et al. Scikit-learn: Machine Learning in Python. Journal of Machine learning Research, volume 12, pages 2825-2830. 2011.