 https://doi.org/10.3011/ESARDA.IJNSNP.2019.3

Use of machine learning models for the detection of fuel pin replacement in spent fuel assemblies

Riccardo Rossa, Alessandro Borella
SCK•CEN Belgian Nuclear Research Centre
Boeretang 200 – 2400 Mol, Belgium
E-mail: rrossa@sckcen.be

Abstract:

The nuclear material contained in the spent fuel assemblies represents the majority of the material verified during the safeguards inspections, and the replacement of spent fuel pins from an assembly is one of the possible scenarios to divert nuclear material.

Due to the high number of fuel pins contained in a fuel assembly (e.g. 264 pins in a PWR 17x17 geometry), a practically infinite number of diversion scenarios can be considered by a potential proliferator. In this framework, Monte Carlo simulations were used to model some of the possible diversion scenarios and to develop a database of detector responses corresponding to different non-destructive assay (NDA) techniques. In addition, the database contains the detector responses obtained with complete fuel assemblies with different initial enrichment, burnup, and cooling time.

Given the large size of the database and the multiple detector responses resulting from the NDA techniques, the use of machine learning is proposed for the data analysis. In this work we focus on the classification problem with the aim of classifying the diversion scenarios based on the percentage of replaced pins. Several machine learning models were developed for this problem using decision trees, discriminant analysis, support vector machine, and nearest neighbors algorithms. The accuracy of the models was calculated as the number of correct classifications in the whole dataset.

The results from the study show that the selection of the detector type used as input in the machine learning model has a strong impact on the accuracy of the developed model. In general the use of gamma-ray detectors leads to higher accuracies compared to the use of neutron detector responses. In addition, several machine learning models achieved a complete correct classification.

Keywords: Machine learning, fuel diversion, Monte Carlo, spent fuel, non-destructive assays

1. Introduction

As defined in the INFCIRC/153 [1] the technical objective of safeguards is the timely detection of diversion of significant quantities of nuclear material. The nuclear material contained in the spent fuel assemblies represents the majority of the material verified during the safeguards inspections [2], and the replacement of spent fuel pins from an assembly is one of the possible scenarios to divert nuclear material.

The capabilities to detect a subset of missing or replaced spent fuel pins, the so-called partial defect testing, were assessed in the past for the Fork detector [3]. In addition, several non-destructive assay (NDA) techniques are proposed to improve the current capabilities for partial defect testing [4], [5]. Among others NDA techniques, the Self-Indication Neutron Resonance Densitometry (SINRD) and the Partial Defect Tester (PDET) have been investigated in the past years at SCK•CEN [6].

Given the large number of diversion scenarios that can be developed and the multiple detector responses resulting from the two NDA techniques, the use of machine learning [7] is proposed for the data analysis as alternative to a previous approach chosen in recent work [8]. Due also to the continuous increase in the computer power, machine learning is extensively used in many fields where large amount of data is available [9], [10]. Within SCK•CEN, research on machine learning applied to the safeguards field focused so far on the use of artificial neural networks (ANN) for the determination of initial enrichment, burnup, and cooling time of spent fuel assemblies [11].

In this contribution the NDA technique chosen for the study is described in Section 2, whereas the overview of the Monte Carlo simulations is presented in Section 3, and the description of the machine learning models is included in Section 4. The results from the data analysis are discussed in Section 5, followed by the conclusion and outlook for future work in Section 6.

2. Description of the NDA technique

2.1 Models geometry

The NDA technique chosen for this study stems from the Self-Indication Neutron Resonance Densitometry (SINRD) and the Partial Defect Tester (PDET).

The SINRD technique is a passive NDA technique that was originally developed by LANL based on the passive

Rossa, R., & Borella, A. (2019). Use of machine learning models for the detection of fuel pin replacement in spent fuel assemblies. *ESARDA Bulletin - The International Journal of Nuclear Safeguards and Non-proliferation*, 58, 22-34. https://doi.org/10.3011/ESARDA.IJNSNP.2019.3

neutron emission from spent fuel due to spontaneous fissions and (α,n) reactions [12]. The principle of the SINRD technique is to measure the attenuation of the neutron flux in the 0.3 eV energy region to obtain a direct estimation of the ^{239}Pu content in the spent fuel. The technique has been studied for the measurement of spent fuel underwater at LANL [13], [14], [15], whereas a measurement approach in air has been investigated at SCK•CEN [16], [6]. Previous research [6] indicated that the use of ^{239}Pu fission chambers increased the sensitivity to the ^{239}Pu content in the spent fuel. In addition, the fast neutron flux was measured with ^{238}U fission chambers according to the approach developed at SCK•CEN.

The Partial Defect Tester (PDET) is a NDA technique that measures the passive neutron and gamma emission from the fuel assembly by inserting a set of small detectors in the guide tubes of a PWR fuel assembly [17], [18], [19]. The PDET was originally proposed by LLNL with the aim of detecting partial defects, and a PDET prototype has been built and tested in a measurement campaign at the Swedish Interim Storage Facility CLAB in January 2015 [20]. The total neutron flux was measured with ^{235}U fission chambers, whereas the gamma-ray flux was measured with ionization chambers.

The NDA detector setup adopted in this study is shown in Figure 1 for both measurements in air and in fresh water. A thick slab of polyethylene surrounds the fuel assembly during the measurement in air (left side of Figure 1) to ensure neutron moderation. The PWR 17x17 fuel assembly geometry was considered in all simulations.

The detector types considered for the SINRD technique and those used by PDET are combined in the NDA technique proposed for this study, and the details of the detector responses are described in Section 2.2. The detectors positions include the guide tubes (red and yellow positions in Figure 1) as in the approach proposed for the PDET detector, but also additional detectors are placed around the fuel assembly (green positions).

2.2 Detector responses

The detector responses were calculated from the results of the Monte Carlo simulations and following the approach proposed in [6]. As shown in Figure 1, for both measurement in air and in fresh water the detector positions and the detector types are identical. The calculated detector responses include:

- Thermal neutrons (TH): bare ^{235}U fission chamber;
- Fast neutrons (FAST): bare ^{238}U fission chamber;
- Resonance region neutrons (RES): difference between the neutron counts with a ^{239}Pu fission chamber covered by Gd foil and a ^{239}Pu fission chamber covered by Cd foil;
- Gamma-rays (P): ionization chamber.

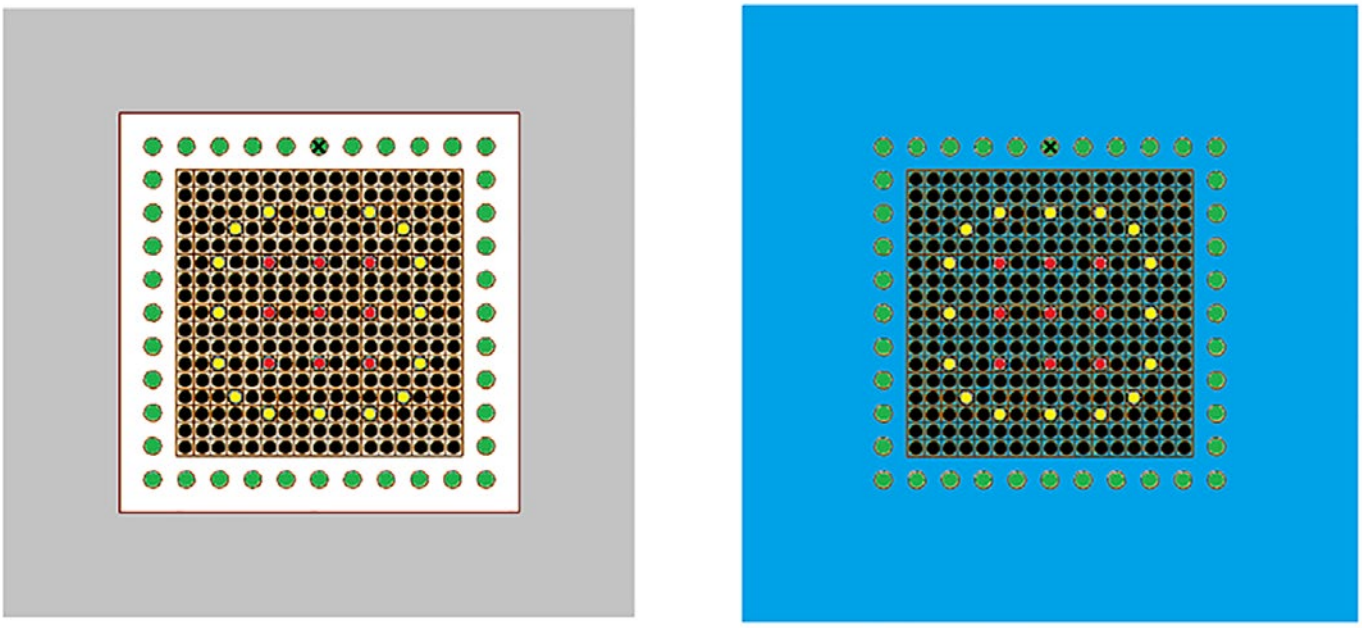

Figure 1: Monte Carlo models of the detector setups chosen in this study. The PWR 17x17 fuel assembly is represented with the fuel pins shown in black. The positions of the detectors are depicted in red, yellow, and green. The detector responses are normalized to the value obtained for the detector position marked with a cross. The picture on the left shows the setup for the measurement in air, with the polyethylene slab in grey surrounding the fuel assembly, and the picture on the right shows the setup for the measurement in fresh water (depicted in blue).

3. Overview of the Monte Carlo simulations

3.1 Complete fuel assemblies

The same set of Monte Carlo simulations was performed both for the NDA technique in air and in fresh water. From each simulation the detector responses were normalized to the value obtained for the detector marked with a cross in Figure 1. The average value was then calculated for the nine central guide tube positions, the sixteen peripheral guide tube positions, and the forty detector positions outside the fuel assembly. The choice to normalize the detector responses followed the approach for the PDET detector where the detector responses are normalized to one detector position [18]. In this way the range of values of the normalized detector responses is greatly reduced compared to the range of the un-normalized detector responses. Future work will consider as input the detector responses before the normalization and will estimate the influence of this step in the accuracy of the machine learning models.

A first set of simulations was performed with complete fuel assemblies, i.e. assemblies with all the fuel pins with equal material composition and source strength, considering different values of:

- Initial enrichment: 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0%;
- Burnup: 5, 10, 15, 20, 30, 40, 60 GWd/t_{HM};
- Cooling time: 1, 5, 10, 50 years.

A total of 196 simulations resulted from all combinations of these parameters. The aim of these simulations is to assess the influence of the fuel irradiation history on the calculated

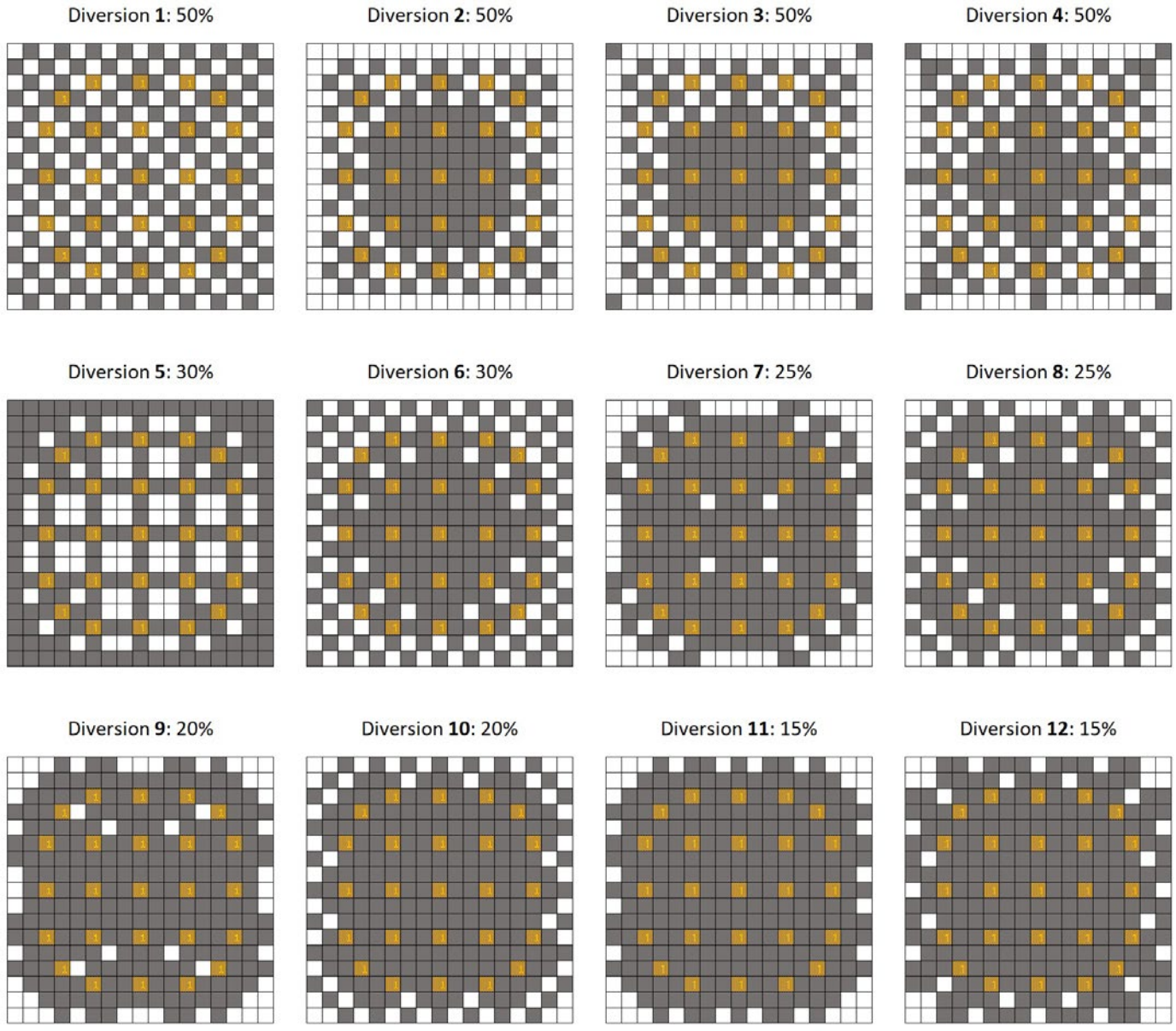

Figure 2: Overview of the diversions scenarios considered in this study. The fuel pins are marked in grey, the dummy pins in white, and the guide tube positions in yellow. The percentage of dummy pins is mentioned for each scenario.

detector responses. The fuel composition and source strength were taken from the SCK•CEN reference spent fuel library [21]. This library did not contain fuel assemblies with burnable poisons; therefore, the impact of this design characteristic could not be estimated at this stage.

3.2 Diversion scenarios

A second set of simulations considered twelve diversion scenarios, where some of the spent fuel pins were replaced by dummies made of stainless steel. The remaining spent fuel pins had equal material composition and source strength as in the case for complete fuel assemblies, whereas no source term was included in the dummy pins. The scenario with pin replacement is expected to be more difficult to detect compared to the case of diversion without replacement.

The diversion scenarios are shown in Figure 2 and they cover cases with replacement between 50% and 15% of the total number of fuel pins. The fuel pins are depicted in grey, the dummy pins in white, and the guide tube positions in yellow. Most of the replacement occurs on the outer region of the fuel assembly, but also a chess-board pattern (Diversion 1) and diversion from the inner section of the fuel assembly (Diversion 5) are included.

All simulations concerning the diversion scenarios considered fuel with a cooling time of 5 years, but the influence of the irradiation history was taken into account simulating fuel with different values of:

- Initial enrichment: 2.0, 3.5, 5.0%;
- Burnup: 10, 30, 60 GWd/t_{HM}.

A total of 108 simulations were carried out in this set of simulations.

4. Machine learning models used for the data analysis

4.1 Introduction to machine learning

Machine learning is used nowadays for a broad range of applications such as speech recognition, financial fraud detection, and cancer prognosis. [22], [23], [24], [25], [26]

The machine learning models can be divided into two broad categories of supervised and unsupervised learning [27]. In the case of supervised learning the observations in the dataset have associated output values, whereas in the case of unsupervised learning the input data do not have corresponding output values.

A machine learning model for supervised learning is first developed during the training phase on a set of known input and output data. Once the model is trained, it is used to predict (prediction phase) new input data for which the output data is unknown. The generic workflow for supervised learning is shown in Figure 3. The machine learning model developed by supervised learning uses regression or classification techniques depending on the type of output data. Regression techniques are used to predict output data that can assume continuous values (e.g. changes in temperature or pressure, fluctuations in housing prices), whereas classification techniques are used to classify input data into categories that can assume only a limited set of values (e.g. type of fruit, benign/malign tumor). [9]

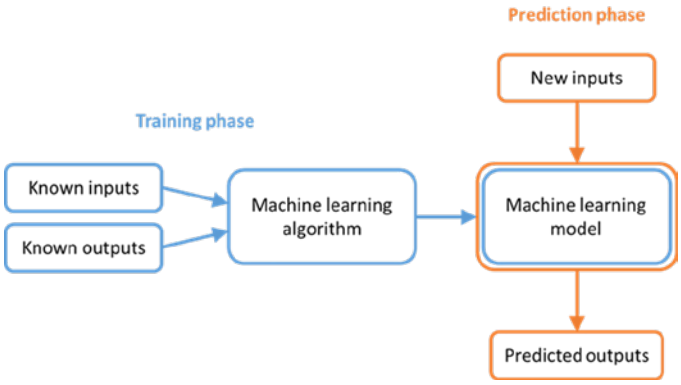

Figure 3: Generic workflow in case of supervised learning. The development of the machine learning model from known inputs and outputs is indicated as training phase, whereas the prediction of outputs from new inputs is referred to as prediction phase.

Independently from the machine learning algorithm chosen, the data needed to develop a machine learning model is generally organized in a database. According to the machine learning terminology the records in the database are called observations and the input variables are called features or predictors. In case of supervised learning the output variables are called responses. Specific to classification techniques, the responses can assume only a finite set of values (either in numerical or text format) called classes.

The detector responses calculated with the Monte Carlo simulations described in Section 3 were organized in a database where the normalized average detector responses were the features and the percentage of replaced pins represented the response. Six classes were defined for the response, representing 50, 30, 25, 20, 15, and 0% of replaced pins. Therefore, each observation in the database consisted in 12 features and 1 response. An extract of the database for the NDA technique with the fuel assembly in air is shown in Table 1; the same database structure was used for the data with the fuel assembly in fresh water. Since each observation in the training database contained the corresponding response class, the detection of fuel pins diversion was treated as a supervised machine learning problem to be solved with classification techniques.

The accuracy of the models was calculated as the number of correct classifications in the whole dataset, and this metric was used to compare the different machine learning models developed.

Features												Resp.
Central det. positions				Peripheral det. positions				External det. positions				
TH	FA	RES	P	TH	FA	RES	P	TH	FA	RES	P	
0.35	1.85	0.40	2.09	0.52	1.64	0.56	1.91	1.10	0.86	1.04	0.85	0
0.34	1.85	0.35	2.09	0.51	1.64	0.51	1.91	1.10	0.86	1.05	0.85	0
0.34	1.86	0.32	2.09	0.51	1.64	0.48	1.91	1.10	0.86	1.06	0.85	0
0.34	1.86	0.30	2.09	0.51	1.64	0.47	1.91	1.10	0.86	1.06	0.85	0
0.33	1.87	0.28	2.08	0.51	1.65	0.45	1.91	1.10	0.86	1.06	0.85	0

Table 1: Extract of the training database for the NDA technique with the fuel assembly in air.

4.2 Parameters chosen for the machine learning models

The machine learning models for this study were developed using the Classification Learner App that is part of the MATLAB Statistics and Machine Learning Toolbox [28].

The toolbox offers the choice of several machine learning algorithms that can be used for supervised and unsupervised learning problems. The Classification Learner App has a graphical user interface (GUI) that allows the user to select in the first window the training database, the variables to be used as features, and those to be used as responses.

The selection of the validation scheme used to assess the accuracy of the developed model is the next step in the GUI. The default MATLAB 5-fold cross-validation approach was chosen, where the training database is divided into five equal-sized subsections. As noted in [28], the validation scheme is used only for the estimation of the model accuracy; the final model is always trained using the complete training database.

The next section in the GUI is the selection of the model type and model parameters, and the training of the model. Four main families of machine learning models were used for the data analysis: decision trees, discriminant analysis, support vector machines, and nearest neighbors classifiers. The principles of the models are described extensively in literature [28], [22], [9], [29], [30], [31]. The parameters used for each developed model are listed in Tables 2-5.

Model name	Maximum number of splits	Split criterion	Surrogate decision splits
Simple tree	5	Gini's diversity index	Off
Medium tree	10	Gini's diversity index	Off
Complex tree	30	Gini's diversity index	Off

Table 2: Parameters chosen for the decision trees models.

Model name	Function used to separate classes	Covariance matrix
Linear discriminant	Linear	Diagonal
Quadratic discriminant	Quadratic	Diagonal

Table 3: Parameters chosen for the discriminant analysis models.

Model name	Kernel function	Box constraint level	Kernel scale mode	Manual kernel scale	Multiclass method	Standardize data
Linear SVM	Linear	1	Auto	------	1-vs-1	Yes
Quadratic SVM	Quadratic	1	Auto	------	1-vs-1	Yes
Cubic SVM	Cubic	1	Auto	-----	1-vs-1	Yes
Coarse Gaussian SVM	Gaussian	1	Manual	4	1-vs-1	Yes
Medium Gaussian SVM	Gaussian	1	Manual	1	1-vs-1	Yes
Fine Gaussian SVM	Gaussian	1	Manual	0.25	1-vs-1	Yes

Table 4: Parameters chosen for the support vector machine models. In the first three models the kernel scale mode was set to Auto, so the option Manual kernel scale was not used.

Model name	Number of neighbors	Distance metric	Distance weight	Standardize data
Coarse kNN	100	Euclidean	Equal	Yes
Medium kNN	10	Euclidean	Equal	Yes
Fine kNN	1	Euclidean	Equal	Yes
Cosine kNN	10	Cosine	Equal	Yes
Cubic kNN	10	Minkowski (cubic)	Equal	Yes
Weighted kNN	10	Euclidean	Squared inverse	Yes

Table 5: Parameters chosen for the nearest neighbors models.

5. Results

5.1 NDA technique with fuel assembly in air

The classifier models described in Section 4.2 were applied to the detector responses for the NDA technique with the fuel assembly in air. The accuracies of all models developed using as features the detector responses from the external detector positions are shown in Table 6, whereas the results obtained using all input features are shown in Table 7. These results were selected because they represent the lowest and highest accuracies among the models developed.

The rows of Tables 6-7 indicate the names of the machine learning models, whereas the columns indicate the normalized detector responses used for the analysis. As described in Section 2.2 the detector responses refer to neutron detectors sensitive to the thermal (TH), resonance (RES), or fast (FA) energy regions, and to the total gamma-ray emission (P). One or more detector responses were considered in the analysis and are included in the table. In Table 6 only one feature per detector response was used, namely the average detector response from the detectors in the external positions. In Table 7 three separate features were used for each detector response, corresponding to the average values from the detectors in the central, peripheral, and external positions, respectively. The values included in the table are the accuracy of each model, which is defined as the percentage of observations with correct classification. It was not possible at this stage to estimate the uncertainty of the calculated accuracy, but this topic will be addressed in future work.

The results of Table 6 indicate that the selection of the features used as input variable in the model is important to obtain a reliable classification, and in general the use of the gamma-ray detector response leads to higher accuracy of the model compared to other detector types. This result is in line with previous research [8]. However, the addition of multiple features does not strongly improve the accuracy of the model in most of the cases.

Once the features used in the model are chosen, similar accuracies were obtained for most of the machine learning models applied in this study. However, “Linear discriminant”, “Coarse kNN”, and “Cosine kNN” models showed several cases where the accuracy was lower than 75%. Complete correct classifications were reached for the “Complex tree”, “Fine Gaussian SVM”, “Fine kNN”, and “Weighted kNN” models when the gamma-ray detector response (P) was used as feature alone or in combination with the fast neutron detector response (FA).

The accuracies calculated for the machine learning models using the detector responses from all positions (i.e. central, peripheral, and external) are included in Table 7. Most of the conclusions drawn from the results in Table 6 are also applicable for Table 7, but in general the use of the detector responses from all available positions lead to an increase in the model accuracy. The largest accuracies are obtained using the responses of detectors sensitive to fast neutrons or gamma-rays. Accuracies lower than 75% were obtained for all cases using “Fine Gaussian SVM” and “Coarse kNN” models except when the feature used was only the fast neutron detector response (FA) or the gamma-ray detector response (P). Complete correct classifications were reached for several models, usually when the gamma-ray detector response (P) was used as feature either alone or in combination with other features.

5.2 NDA technique with the fuel assembly in fresh water

Machine learning models based on the parameters described in Section 4.2 were also developed from the detector responses of the NDA technique with the fuel assembly in fresh water. The accuracy for each model was computed and compared to the accuracy of the corresponding model developed for the NDA technique with the fuel assembly in air. Tables 8-9 show the accuracy calculated using the detector responses from the external positions or from all positions, respectively.

The results in the tables show that in most of the cases the accuracy calculated for the fuel assembly kept either in air or in fresh water is within ±5%. Therefore, the comments reported in Section 5.1 are also valid in this Section. Focusing on the cases where the difference in accuracy is larger than 5%, the accuracy for the NDA technique with the fuel assembly in fresh water is generally higher than the corresponding value obtained with the fuel assembly in air. This is observed in Table 8 when the responses of detectors sensitive to thermal neutrons (TH) or resonance region neutrons (RES) are used as features alone or in combination. On the contrary, a decrease between 5 and 10% was observed for several models when the responses of detectors sensitive to resonance region neutrons (RES) or fast neutrons (FA) are both used as features. Complete correct classifications were achieved using the “Fine Gaussian SVM” model with the (FA,P) features, the “Fine kNN” model using the (FA,P), (TH,FA,P), and (RES,FA,P) features, and the “Weighted kNN” model using the (FA,P) feature.

It is worth to note that for the NDA technique with the fuel assembly in fresh water no detector response alone is able to reach a complete correct classification using the detector positions located outside the fuel assembly. This is in contrast to the results obtained for the fuel assembly stored in air, where the gamma-ray detector response reached a complete correct classification with several machine learning models. The results obtained for the NDA technique with the fuel assembly in air are remarkable in the sense that they indicate that a passive gamma measurement in air has the potential for a complete correct

classification. In addition, if neutron detectors such as fission chambers are not needed, it would significantly simplify the design of a measurement device for spent fuel assay.

However, a safeguards verification underwater, like the one based on the PDET approach, is probably more realistic than a measurement in air, like the one based on the SINRD approach.

Hence future work will be targeted at improving the accuracy of the current models to reach a complete correct classification also for the NDA technique with the fuel assembly in fresh water.

Considering the accuracies calculated using all detector positions in Table 9, the largest difference between the NDA technique with the fuel assembly either in air or in fresh water were obtained using as input features the responses of detectors sensitive to thermal neutrons (TH) or resonance neutrons (RES) either alone or in combinations. An increase between 5 and 10% of the accuracy calculated with the fuel assembly in fresh water was obtained for decision tree and discriminant analysis models, whereas a decrease between -5% and -10% was obtained for “Linear SVM”, “Quadratic SVM”, and “Cubic SVM” models. Several models reached a complete correct classification using the NDA technique with the fuel stored in fresh water, especially when the gamma-ray detector response (P) was used alone or in combinations with other detector responses.

Detector positions considered: external (1 feature per detector response)															
Detector response 1	TH	RES	FA	P	TH	TH	TH	RES	RES	FA	TH	TH	TH	RES	TH
Detector response 2					RES	FA	P	FA	P	P	RES	RES	FA	FA	RES
Detector response 3											FA	P	P	P	FA
Detector response 4															P
Simple tree	67.8	70.1	83.9	84.9	72.0	82.9	85.9	82.2	84.5	86.2	84.5	83.6	87.2	88.8	87.2
Medium tree	68.8	68.4	86.2	94.4	69.4	84.5	94.7	85.9	92.1	99.0	87.2	93.1	97.4	96.4	94.7
Complex tree	63.8	62.5	85.9	**100**	66.8	84.9	97.7	87.2	94.1	99.3	85.9	93.8	98.4	96.1	96.4
Linear discriminant	68.1	69.4	70.4	73.4	71.1	79.9	74.3	83.2	78.9	76.6	79.6	79.6	80.3	81.6	81.3
Quadratic discriminant	67.8	69.7	79.9	85.2	70.4	84.9	85.2	87.8	90.5	85.2	85.2	88.2	85.9	87.5	85.9
Linear SVM	67.4	70.1	76.0	80.6	69.7	83.9	79.9	85.2	87.2	86.5	84.9	84.9	88.5	89.8	90.1
Quadratic SVM	69.7	70.1	81.3	89.5	69.7	86.2	89.1	88.8	94.1	94.7	90.1	92.4	96.4	95.4	96.7
Cubic SVM	39.8	53.9	79.9	96.4	72.0	86.8	94.4	90.5	94.1	96.7	86.8	91.4	97.7	96.4	96.4
Coarse Gaussian SVM	64.5	70.1	78.3	79.3	70.1	83.9	78.0	86.2	83.9	83.9	86.2	84.5	86.5	90.1	90.1
Medium Gaussian SVM	68.4	70.4	80.9	90.8	70.4	85.5	87.5	89.8	95.1	99.3	89.8	93.4	97.4	96.7	96.7
Fine Gaussian SVM	67.4	68.8	88.8	**100**	70.1	85.2	89.8	81.6	94.4	**100**	69.7	71.4	84.9	81.3	68.8
Coarse kNN	64.5	64.5	70.4	64.5	64.5	65.5	64.5	67.4	64.5	67.4	64.5	64.5	67.1	67.4	67.4
Medium kNN	67.4	68.8	87.5	98.7	70.1	83.2	81.3	89.1	92.4	98.4	84.9	86.2	87.8	96.4	89.5
Fine kNN	56.9	54.9	84.9	**100**	62.5	88.5	90.5	88.8	94.7	**100**	86.2	91.4	97.4	98.0	96.1
Cosine kNN	64.5	64.5	75.7	69.4	72.0	80.9	74.0	81.3	82.2	89.8	82.6	81.6	82.9	88.5	85.2
Cubic kNN	66.8	69.7	86.8	98.0	72.0	81.3	81.6	89.5	88.5	97.0	84.9	85.5	85.9	94.1	90.1
Weighted kNN	60.2	59.9	87.2	**100**	68.8	84.2	92.8	88.8	95.4	**100**	86.5	92.1	97.0	97.7	94.1

Table 6: Accuracy (%) of the machine learning models using the detector responses from the external detector positions. Results for the NDA technique with the fuel assembly in air. The cases with complete correct classification are highlighted in bold.

Detector positions considered: central, peripheral, external (3 features per detector response)

Detector response 1	TH	RES	FA	P	TH	TH	TH	RES	RES	FA	TH	TH	TH	RES	TH
Detector response 2					RES	FA	P	FA	P	P	RES	RES	FA	FA	RES
Detector response 3											FA	P	P	P	FA
Detector response 4															P
Simple tree	70.7	70.7	92.4	92.8	72.0	92.4	92.8	92.4	93.1	93.1	92.4	93.1	92.1	92.4	91.4
Medium tree	70.4	71.4	98.0	**100**	72.7	97.4	**100**	97.7	**100**	99.7	96.4	**100**	99.3	99.3	99.0
Complex tree	71.7	68.4	98.0	**100**	71.1	97.4	**100**	98.0	**100**	99.7	97.7	**100**	99.0	99.3	99.3
Linear discriminant	71.1	68.8	90.1	88.8	63.5	87.2	86.2	85.5	85.9	88.5	82.6	84.2	90.5	87.5	88.2
Quadratic discriminant	71.7	66.8	92.4	94.4	66.1	93.1	95.1	93.1	95.1	94.4	91.8	95.1	94.4	95.4	95.4
Linear SVM	89.8	78.6	94.1	94.1	87.5	94.1	93.8	94.1	93.1	96.1	94.7	96.4	94.1	95.4	94.4
Quadratic SVM	92.8	80.6	97.7	98.7	93.1	99.7	99.7	98.7	97.7	98.4	99.0	99.0	98.4	99.3	99.7
Cubic SVM	91.8	82.6	99.3	**100**	92.1	**100**	**100**	98.7	99.7	**100**	98.7	98.4	**100**	99.7	98.7
Coarse Gaussian SVM	74.7	73.4	92.8	94.4	77.6	94.7	94.1	93.1	94.7	96.1	95.1	95.7	97.0	97.7	97.4
Medium Gaussian SVM	83.6	75.0	99.7	**100**	80.9	98.7	99.0	95.4	97.7	**100**	87.8	86.8	98.0	95.7	86.2
Fine Gaussian SVM	79.3	70.1	98.0	**100**	64.5	68.1	77.3	64.5	70.4	98.7	64.5	64.5	67.4	64.5	64.5
Coarse kNN	64.5	64.5	73.4	67.4	64.5	72.0	67.4	70.7	66.8	73.0	69.7	66.8	71.7	71.7	70.1
Medium kNN	76.0	73.4	98.7	**100**	75.3	86.2	87.5	90.1	91.1	99.7	85.2	87.8	90.8	93.4	91.4
Fine kNN	83.6	81.3	99.7	**100**	80.3	98.7	99.7	97.4	97.7	**100**	94.7	95.4	**100**	99.0	97.4
Cosine kNN	81.3	73.0	96.1	99.3	78.6	85.2	86.8	87.2	85.2	97.7	85.5	83.2	85.5	88.2	86.2
Cubic kNN	79.6	73.4	99.3	**100**	75.3	85.5	87.5	88.5	89.8	98.7	84.5	85.9	89.5	92.4	87.5
Weighted kNN	82.2	80.3	99.7	**100**	79.6	98.4	97.7	95.7	96.1	**100**	92.1	91.1	99.0	99.0	96.4

Table 7: Accuracy (%) of the machine learning models using the detector responses from all detector positions. Results for the NDA technique with the fuel assembly in air. The cases with complete correct classification are highlighted in bold.

Detector positions considered: external (1 feature per detector response)															
Detector response 1	TH	RES	FA	P	TH	TH	TH	RES	RES	FA	TH	TH	TH	RES	TH
Detector response 2					RES	FA	P	FA	P	P	RES	RES	FA	FA	RES
Detector response 3											FA	P	P	P	FA
Detector response 4															P
Simple tree	76.0	74.7	82.2	87.2	78.0	82.9	86.2	83.6	87.2	88.5	84.2	85.9	88.2	89.1	88.8
Medium tree	77.6	75.7	85.2	96.1	79.9	88.8	95.7	84.9	97.7	97.4	88.2	95.1	96.4	95.4	94.7
Complex tree	77.6	74.0	86.2	98.7	78.6	88.5	96.4	84.9	96.4	97.0	88.5	95.1	96.1	93.1	96.7
Linear discriminant	71.4	71.1	69.4	74.0	71.7	72.0	79.6	73.4	77.3	79.3	73.0	80.3	80.9	78.6	80.9
Quadratic discriminant	78.3	76.6	78.3	85.5	77.0	81.3	87.8	80.6	85.2	86.8	83.9	88.8	89.8	87.2	91.4
Linear SVM	75.0	73.7	76.6	82.6	74.7	78.0	87.5	76.6	83.9	85.5	77.6	87.5	90.1	84.2	91.4
Quadratic SVM	77.0	75.7	78.6	88.2	77.3	88.8	92.4	85.9	90.1	98.0	86.8	92.8	95.7	95.4	95.7
Cubic SVM	72.0	69.4	78.3	91.8	76.0	89.1	95.1	85.2	93.1	99.0	88.8	94.7	98.4	97.7	97.7
Coarse Gaussian SVM	77.3	74.3	76.6	80.3	77.3	79.3	88.5	77.3	84.2	82.9	79.6	88.8	89.8	88.5	91.8
Medium Gaussian SVM	77.6	76.0	81.6	85.9	78.3	87.2	93.8	83.6	90.1	96.7	87.5	92.4	99.0	97.7	99.3
Fine Gaussian SVM	78.9	77.6	87.2	97.0	76.3	87.5	95.7	82.9	92.8	**100**	78.9	79.6	96.7	93.1	76.6
Coarse kNN	71.4	68.8	70.4	64.5	69.7	70.4	67.4	70.4	67.4	68.4	70.4	67.4	68.1	67.8	68.8
Medium kNN	77.3	73.4	83.6	98.4	78.6	87.2	93.1	82.6	89.1	99.7	85.5	88.5	97.7	91.1	93.1
Fine kNN	74.7	75.0	84.5	97.0	79.3	89.5	97.7	84.2	95.4	**100**	88.8	93.1	**100**	**100**	99.3
Cosine kNN	75.0	64.5	73.7	72.0	72.0	76.6	88.5	76.3	82.6	86.5	78.6	83.6	90.5	87.2	89.8
Cubic kNN	79.9	73.7	83.6	98.7	78.6	89.1	89.5	81.3	87.5	96.7	83.6	89.1	96.4	92.4	93.1
Weighted kNN	74.0	75.3	83.2	97.4	78.6	90.8	97.0	82.9	94.1	**100**	87.8	93.1	99.3	97.7	97.7

Table 8: Accuracy (%) of the machine learning models using the detector responses from the external detector positions. Results for the NDA technique with the fuel assembly in fresh water. The cases with complete correct classification are highlighted in bold.

Detector positions considered: central, peripheral, external (3 features per detector response)																
Detector response 1	TH	RES	FA	P	TH	TH	TH	RES	RES	FA	TH	TH	TH	RES	TH
Detector response 2					RES	FA	P	FA	P	P	RES	RES	FA	FA	RES
Detector response 3											FA	P	P	P	FA
Detector response 4															P
Simple tree	78.6	82.2	93.4	94.1	84.5	93.4	94.1	87.8	88.5	94.1	88.8	88.8	93.8	88.8	89.8
Medium tree	81.6	81.9	98.4	**100**	83.9	98.7	**100**	96.7	97.0	99.7	96.7	96.7	99.7	96.4	95.7
Complex tree	81.6	83.2	99.0	**100**	85.5	99.0	**100**	95.7	97.0	99.7	97.0	95.7	99.3	97.0	96.1
Linear discriminant	75.7	75.0	90.5	86.8	73.0	86.5	86.8	85.5	87.2	91.1	81.3	81.6	89.1	88.5	86.2
Quadratic discriminant	83.6	80.3	93.4	94.7	80.6	93.8	95.4	95.4	96.1	94.7	92.8	95.1	95.7	96.1	95.7
Linear SVM	80.6	81.9	94.1	93.4	82.9	92.8	93.8	95.4	96.7	95.7	95.4	94.4	93.4	95.4	94.4
Quadratic SVM	84.9	84.9	99.0	98.4	87.5	99.7	99.3	99.0	99.7	98.4	99.0	99.7	99.0	99.7	99.7
Cubic SVM	84.2	85.9	**100**	**100**	84.9	**100**	99.3	99.3	**100**	**100**	99.3	99.3	**100**	99.7	**100**
Coarse Gaussian SVM	81.3	83.6	93.4	94.1	83.6	93.8	95.1	95.7	98.4	96.4	94.7	97.4	97.4	99.0	98.0
Medium Gaussian SVM	84.9	82.6	**100**	**100**	84.2	98.7	98.4	97.4	96.7	**100**	89.5	93.1	99.0	97.0	89.1
Fine Gaussian SVM	76.3	69.1	99.7	**100**	65.1	68.1	69.7	65.5	65.8	98.7	64.5	65.1	64.5	65.1	64.5
Coarse KNN	65.1	68.8	73.4	67.4	66.4	72.0	68.1	73.4	71.4	73.4	73.0	69.1	73.0	73.4	73.4
Medium KNN	78.9	81.3	99.3	**100**	82.2	88.2	90.5	94.1	95.7	**100**	90.8	92.8	94.4	99.0	95.4
Fine KNN	81.9	80.9	**100**	**100**	82.9	**100**	**100**	98.0	99.7	**100**	95.1	97.7	**100**	**100**	99.7
Cosine KNN	76.0	77.3	93.8	98.7	81.3	85.5	86.8	90.5	92.1	98.7	88.2	89.8	90.1	95.1	91.1
Cubic KNN	77.3	80.3	97.7	**100**	82.2	85.5	89.5	91.1	94.7	**100**	88.5	90.5	92.4	94.7	93.1
Weighted KNN	82.9	80.6	**100**	**100**	82.9	98.4	98.0	98.0	98.7	**100**	95.1	96.4	99.7	**100**	99.0

Table 9: Accuracy (%) of the machine learning models using the detector responses from all detector positions. Results for the NDA technique with the fuel assembly in fresh water. The cases with complete correct classification are highlighted in bold.

6. Conclusion

The detector responses of a set of detector types were modelled to assess the capability of machine learning models to detect the diversion of fuel pins from a fuel assembly. Several machine learning models were developed and applied for the data analysis for an NDA technique with fuel assembly either in air or in fresh water.

The different models were evaluated in terms of accuracy, which was defined as the percentage of cases with correct classifications compared to the total dataset. The use of the NDA technique for a fuel assembly either in air or in fresh water was also compared using this metric.

The results for the NDA technique showed that the detector response used for the data analysis plays an important role in the accuracy of the model, and in general the gamma-ray detector response obtained the highest accuracy compared to the neutron detector responses. However, the addition of multiple detector responses did not improve significantly the accuracy of the models.

Comparing the use of the NDA technique for a fuel assembly either in air or in fresh water, similar results in terms of accuracy were obtained for the majority of the models. Only in some cases the accuracy calculated with the fuel assembly in fresh water was more than 5% higher compared to the value obtained with the fuel assembly in air.

Very promising results were obtained for the machine learning models using the decision trees, support vector machine, and nearest neighbors techniques. Several models reached a complete correct classification over the dataset used in this study, especially when the gamma-ray detector response was used alone or in combination with other features.

Future work will refine the models developed in this study by investigating different model parameters with the aim of increasing the accuracy of the current models. The uncertainty of the calculated accuracy will also be estimated. In addition, the training database will be further expanded by generating several additional diversion scenarios and including other detector responses to be used as input features.

7. References

[1] International Atomic Energy Agency (IAEA). "*The structure and content of agreements between the Agency and States required in connection with the treaty on the nonproliferation of nuclear weapons*". INFCIRC/153 (corrected), 1972.

[2] International Atomic Energy Agency (IAEA). "*IAEA annual report for 2013*". IAEAGC(58)/3, 2014.

[3] van der Meer K., et al. "*An assessment of the FORK detector as a partial defect tester*". Proceedings of the 27th ESARDA Symposium on Safeguards and Nuclear Material Management, London, 10-12 May, 2005.

[4] Grape S., et al. "*Recent modelling studies for analysing the partial-defect detection capability of the Digital Cherenkov Viewing Device*". Proceedings of the 2013 ESARDA annual meeting, 2013.

[5] White T., et al. "*Application of Passive Gamma Emission Tomography (PGET) to the inspection of spent nuclear fuel*". Proceedings of the 2018 INMM annual meeting, 2018.

[6] Rossa R. "*Advanced non-destructive methods for criticality safety and safeguards of used nuclear fuel*". Ph.D. dissertation at Université libre de Bruxelles, 2016.

[7] Smola A., et al. "*Introduction to Machine Learning*". Cambridge University Press. ISBN 0 521 82583 0, 2008.

[8] Rossa R. "*Comparison of the SINRD and PDET detectors for the detection of fuel pins diversion in PWR fuel assemblies*". Proceedings of the 59th INMM Annual Meeting, 2018.

[9] Kotsiantis S. B. "Supervised Machine Learning: A Review of Classification Techniques". Informatica 31, 2007, Pages 249-268.

[10] Soofi A. A., et al. "*Classification techniques in machine learning: applications and issues*". Journal of basic & applied sciences, 2017, 13, Pages 459-465.

[11] Borella A., et al. "*Signatures from the spent fuel: simulations and interpretation of the data with neural network analysis*". ESARDA Bulletin, No. 55, December 2017, Pages 29-38.

[12] Menlove H. O., et al. "*A resonance self-indication technique for isotopic assay of fissile material*". Nuclear applications, volume 6, 1969, Pages 401-408.

[13] LaFleur A. M. "*Development of self-interrogation neutron resonance densitometry (SINRD) to measure the fissile content in nuclear fuel*". Ph.D. dissertation at Texas A&M University, 2011.

[14] LaFleur A. M., et al. "*Development of self interrogation neutron resonance densitometry to improve detection of partial defects in PWR spent fuel assemblies*". Nuclear Technology, 181, 2013, Pages 354-370.

[15] LaFleur A. M., et al. "*Analysis of experimental measurements of PWR fresh and spent fuel assemblies*

using self-interrogation neutron resonance densitometry". Nuclear Instruments and Methods Section A, 781, 2015, Pages 86-95.

[16] Rossa R., et al. "*Investigation of the self-interrogation neutron resonance densitometry applied to spent fuel using Monte Carlo simulations*". Annals of Nuclear Energy 75, 2015, Pages 176-183.

[17] Sitaraman S., et al. "*Characterization of a safeguards verification methodology to detect pin diversion from pressurized water reactor (PWR) spent fuel assemblies using Monte Carlo techniques*". Proceedings of the 48th INMM annual meeting, 2007.

[18] Ham Y. S., et al. "*Development of a safeguards verification method and instrument to detect pin diversion from pressurized water reactor (PWR) spent fuel assemblies phase I study*". Lawrence Livermore National Laboratory technical report LLNL-TR-409660, 2009.

[19] Ham Y. S., et al. "*A new versatile safeguards tool for verification of PWR spent fuel*". Proceedings of the 53rd INMM annual meeting, 2012.

[20] Ham Y. S., et al. "*Partial defect verification of spent fuel assemblies by PDET: principle and field testing in interim spent fuel storage facility (CLAB) in Sweden*". Proceedings of the 2015 ANIMMA conference, 2015.

[21] Rossa R., et al. "*Development of a reference spent fuel library of 17x17 PWR fuel assemblies*". ESARDA Bulletin n. 50, 2013, Pages 48-60.

[22] Mitchell T. M. "*Machine Learning*". McGraw-Hill editor, 1997.

[23] Bishop C. M. "*Pattern recognition and machine learning*". Springer, 2006, ISBN-13: 978-0387-31073-2.

[24] Oudeyer P-Y. "*The production and recognition of emotions in speech: features and algorithms*". International Journal of Human-Computer Studies, Volume 59, Issues 1–2, 2003, Pages 157-183.

[25] Ngai E. W. T., et al. "*The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature*". Decision Support Systems, Volume 50, Issue 3, 2011, Pages 559-569.

[26] Kourou K., et al. "*Machine learning applications in cancer prognosis and prediction*". Computational and Structural Biotechnology Journal, Volume 13, 2015, Pages 8-17.

[27] Murphy K. P. "*Machine learning: a probabilistic perspective*". Massachusetts Institute of Technology, 2012, ISBN 978-0-262-01802-9.

[28] MathWorks, Inc. "*Statistics and Machine Learning Toolbox™ - User's Guide*", 2018.

[29] Li T., et al. "*Using discriminant analysis for multi-class classification: an experimental investigation*". Knowledge and Information Systems 10 (4), 2006, Pages 453-472.

[30] Kecman V. "*Support Vector Machines – An Introduction*". Studies in Fuzziness and Soft Computing, Springer, 2005.

[31] Iggane M., et al. "*Self-training using a k-Nearest Neighbor as a base classifier reinforced by Support Vector Machines*". International Journal of Computer Applications (0975 – 8887) Volume 56– No. 6, 2012.