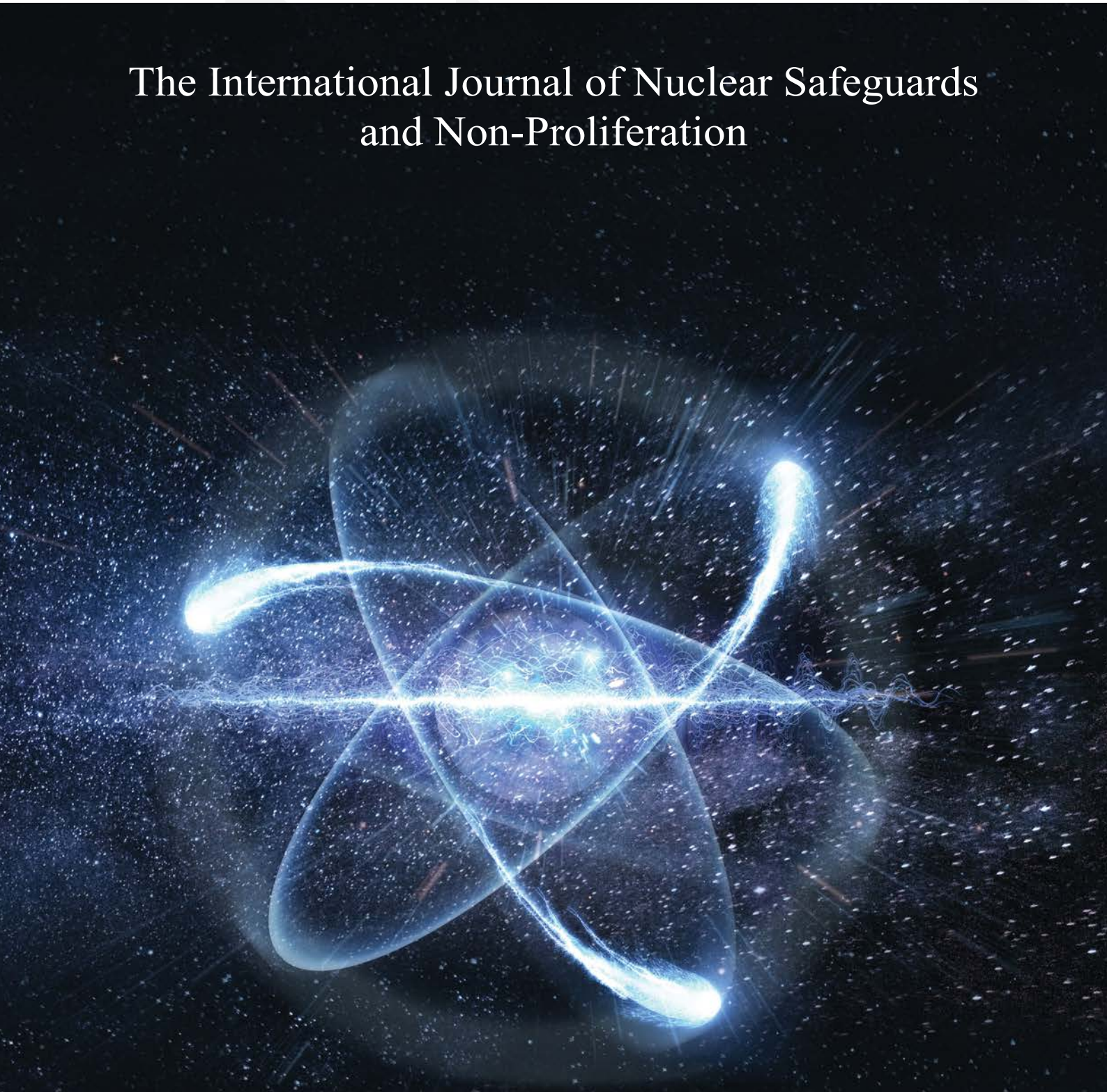
ISSN 1977-5296

ESARDA

European Safeguards Research & Development Association

The International Journal of Nuclear Safeguards and Non-Proliferation



December 2020 Bulletin No.61

ISSN 1977-5296
doi:10.2760/217080
ISBN 978-92-76-18979-4

Number 61
December 2020

European Commission, Joint Research Centre,
Directorate G - Nuclear Safety and Security
Nuclear Security Unit G.II.7
T.P. 800, I-21027 Ispra (VA), Italy
Tel. +39 0332-786182
EC-ESARDA-BULLETIN@ec.europa.eu

ESARDA is an association formed to advance and harmonize research and development for safeguards. More information can be found at the following address:

https://esarda.jrc.ec.europa.eu/

Papers submitted for publication are reviewed by independent authors including members of the Editorial Board.

Manuscripts have to be sent to the Editor (EC-ESARDA-BULLETIN@ec.europa.eu) following the paper guidelines available in the ESARDA Bulletin section of the ESARDA website (https://esarda.jrc.ec.europa.eu/) where the bulletins can also be viewed and downloaded.

Accepted manuscripts are published free of charge.

N.B. Articles and other material in the ESARDA Bulletin do not necessarily present the views or policies of neither ESARDA nor the European Commission.

ESARDA Bulletin is published jointly by ESARDA and the Joint Research Centre of the European Commission and distributed free of charge to over 700 registered members, libraries and institutions worldwide.

The publication is authorized by ESARDA.

Cover designed by Christopher Craig Havenga, (consultant at EC, JRC, G.II.7, Italy)

Printed by
IMPRIMERIE CENTRALE, Luxembourg

ESARDA
European Safeguards Research and Development Association

Bulletin

Contents Issue n° 61

Editorial

Elena Stringa

Dear Readers,

I hope that this issue of the ESARDA Bulletin finds you well.

It is with great pleasure that I open this editorial with very good news: the review of the ESARDA Bulletin for Scopus is complete and the Scopus Content Selection and Advisory Board has advised that the title will be accepted for inclusion in Scopus. The comments of the reviewers were very positive, confirming that articles published in our journal are "scientifically sound and relevant to an international academic or professional audience" in the field of Nuclear Safeguards and Non-Proliferation. Many thanks to the authors and the reviewers who, during the last years, contributed with their work to achieve this goal.

Now that our objective of having our journal included Scopus is very close, I strongly encourage you to continue to submit to ESARDA your valuable work in order to increase citation index and impact factor of our publications.

Academic contributions to be considered for publication in the ESARDA Bulletin should be submitted by email to EC-ESARDA-BULLETIN@ec.europa.eu together with the signed paper submission form that you can find in the ESARDA Bulletin web page, in the publications section of the website. Technical contributions, i.e. articles relevant for the ESARDA community with a content more technical than academic, can be considered for publication in the Connector newsletter by sending them to EC-ESARDA-CONNECTOR@ec.europa.eu.

In the Connector newsletter n.3, edited while I am writing this editorial, you can find updated news from the various ESARDA working groups, as well as the outcome of the ESARDA working groups meetings held in 17-19 November, for the first time in a complete remote format.

As always, I conclude the editorial by thanking Chris Havenga, author of the Bulletin cover, and Andrea De Luca, ESARDA assistant editor, who contributed in a significant way to the realisation of the journal.

Take care and have a pleasant reading.

Dr. Elena Stringa, PhD
Editor of the ESARDA Bulletin - The International Journal
of Nuclear Safeguards and Non-Proliferation
https://esarda.jrc.ec.europa.eu/
EC-ESARDA-BULLETIN@ec.europa.eu
Elena.Stringa@.ec.europa.eu

Heterogeneity effects on nondestructive assay measurements of enrichment in UF_6 cylinders

Allison T. Greaney[1], Susan K. Smith[1], Ramkumar Venkataraman[1], Jason M. Richards[1], Carlos D. Rael[2], Martyn T. Swinhoe[2], Duc T. Vo[2], Ron D. Jeffcoat[3], and Glenn A. Fugate[1],

[1]Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA
[2]Los Alamos National Laboratory, Los Alamos, NM 87545, USA
[3]Savannah River National Laboratory, Aiken, SC 29808, USA
E-mail: fugatega@ornl.gov

Abstract

Measurements were performed using multiple mechanically cooled high-purity germanium detectors at six positions around standard industrial 30B cylinders (2.2 metric ton) of UF_6 to assess if matrix inhomogeneity is detectable and its impacts on the measured apparent uranium enrichment. Uranium enrichment was calculated with FRAM™. Six of the nine cylinders appeared to be homogeneous as indicated by having similarly accurate apparent measured uranium enrichment, at all positions. However, three cylinders appeared to have localized inhomogeneities. This was manifested as very low apparent enrichments, often <10% of the declared value, on one side of the cylinder. Examination of the spectra suggested both elevated ^{234}Th-^{234m}Pa daughter isotopes and reduced ^{235}U were measured at these locations. It is hypothesized that these heterogeneous cylinders may have experienced asymmetric solar heating, which caused volatile UF_6 to sublime preferentially away from the warmed side. Care should be taken during uranium enrichment verification when applying methods that include gamma-rays associated with daughter nuclides to the determination of uranium enrichment for cylinders that are stored in sunlight.

Keywords: enrichment; UF_6; HPGe; FRAM; spectrometry

1. Introduction

Accurate quantification of uranium enrichment in uranium hexafluoride (UF_6) cylinders is necessary for material accounting and nuclear safeguards. Non-destructive assay (NDA) methods using gamma-ray spectroscopy are often used to verify enrichment declarations. NDA enrichment determinations either employ only the 185.7 keV emission associated with the ^{235}U (e.g., the enrichment meter) or ratios of gamma-ray emissions associated with the radioactive daughter nuclides of one or more uranium isotopes. The Fixed Energy Response Function Analysis with Multiple Efficiency (FRAM™) code can perform isotopic analysis on uranium using gamma- and X-ray peaks associated with ^{234}U, ^{235}U, ^{238}U, and ^{234m}Pa.[1] While more complex, use of FRAM™ eliminates the need for measuring the wall thickness of the cylinder or calibrations using known enrichment uranium materials. For low-enriched uranium (LEU) analysis, the software can be run in two modes: "planar" which uses 60 – 250 keV peaks and "coaxial" mode which uses 120 – 1010 keV peaks.[2,3,4] The latter mode is useful for applications to thick-walled walled vessels such as standard industrial 30B cylinders (2.2 metric ton) of UF_6. For an accurate uranium enrichment to be calculated by FRAM™ or any other method that utilizes gamma-ray emissions from uranium daughter products, the UF_6 must be homogenously distributed with respect to uranium and daughter products, and the ^{234}Th and ^{234m}Pa isotopes must be in secular equilibrium with their parent ^{238}U. Given that the half-lives of ^{234}Th and ^{234m}Pa are 24.1 days and 1.17 minutes, respectively, the sample must be at least 3.5 months old (~95% of secular equilibrium) to allow the system to minimize the effects of disequilibrium on the measurement.

The detector geometry, in relation to the distribution of the UF_6 inside the cylinder, is also important for accurate analysis. Depending on the manner in which the cylinder was filled, the UF_6 may be found in different wall thickness distributions. Initially, gas transfers of UF_6 tend to fill with a central void as material sublimes into the cylinder. Liquid filled cylinders initially have a void in the top horizontal half of the cylinder due to volume reduction as the liquid freezes into the denser solid. It should be noted that these are initial geometries which likely progress toward an intermediate state through sublimation and mechanical fracturing over some variable time frame where all surfaces have a significant thickness of UF_6 and the large void is predominantly found in the top half of the cylinder. The distribution of UF_6 around a full cylinder likely meets the ~1 cm infinite thickness requirement of the 185.7 keV gamma-ray for the enrichment meter method at almost any point. Theoretically, enrichment meter and FRAM™ do not require an infinite thickness for higher energy peaks (e.g. ^{234m}Pa) and so do not need an infinite thickness of UF_6. Berndt et al. [5] demonstrated that the detector location and filling profile influence the detector response, and therefore the accuracy of the enrichment, as measured by the enrichment meter measurement. As such, the detector positions were chosen to "guarantee" the requirement of infinite thickness by focusing on the bottom third of the cylinder. Other studies have recently investigated localized spatial impacts related to heel and daughter distribution on spectra and enrichment measurements.

This study presents position-dependent gamma-ray measurements of 30B cylinders made with high purity germanium (HPGe) detectors and analyzed using FRAM™ v5.2 distributed by ORTEC® to determine the uranium enrichment. This approach has been shown to produces quick, accurate results comparable to those of the traditional enrichment meter method [2,6] and has the added advantages that it does not require determination of the wall thickness of the vessel, calibration on a known enrichment uranium material, or an infinite thickness type geometry. This work focused on determining the impacts on enrichment as determined by FRAM™ due to potentially variability within the physical UF_6 distribution and daughter nuclide distribution around the cylinder by performing horizontal profiling.

2. Methods

Gamma-ray spectra were collected around nine 30B UF_6 cylinders that ranged in enrichment between 0.71 and 4.95 weight% ^{235}U. Examination of various enrichment calculation methods [6] and neutron emission rates [7] measurements of these data have been reported elsewhere. The spectra were collected in six positions around the cylinder using commercial, mechanically cooled high purity germanium (HPGe) detectors (**Fig. 1**). These positions were selected to attempt to produce a uniform geometry that assumed a homogenous distribution of uranium and daughter products inside the cylinder without collimation. Both coaxial (three detectors; all ORTEC® portable HPGe detectors) and planar (two detectors; both Canberra/Mirion Falcon 5000 HPGe detectors) crystal types were used. A 10 to 30 minute spectrum was collected using each detector at all six positions around the cylinder (e.g., 5 detectors × 6 positions = 30 gamma-ray spectra per cylinder). The position numbering system was kept constant across all cylinders: positions 1, 2, and 3 were on one side of the cylinder while position 4, 5, and 6 were on the other side (**Fig. 1**). Three of the detectors were angled so that they were in contact with the face of the cylinder near the base where the suspected UF_6 profile was thickest. The two coaxial detectors had to be positioned in a similar geometry but 12 inches from the wall of the cylinder to decrease the overall deadtime of the measurement. Deadtimes varied between 20% and 90%, with high dead times are addressed in section 4.2. The data were initially collected to examine the full gamma-ray spectrum of each cylinder and so the methodology was not optimized for enrichment measurement purposes, *e.g.,* no collimation was used which may have improved the measurement.

The gamma-ray spectra were analyzed with the commercial FRAM™ software in "coaxial" mode, using the ULEU_Cx_120-1010 parameter set. Using this parameter set, the software uses ratios of high and low energy peaks to quantify the isotopic fractions of ^{234}U, ^{235}U, ^{236}U, and ^{238}U. The ^{235}U enrichment is calculated using a ratio between the 185.7 keV peak from ^{235}U and the 1001 keV peak from ^{234m}Pa (^{234}Th), following Eqn. 1.[1] where N^i = number of atoms of isotope i, $C(E)$ = peak area of gamma-ray j of energy E, $T_{1/2,\,i}$ = half-life of isotope i, BR_j = branching ratio of gamma-ray j, and $RE(E_j)$= relative efficiency value of gamma-ray j determined by a nonlinear least squares curve fit to the relative efficiency values at various energies. Thus, this method relies on the attainment of secular equilibrium between ^{238}U and ^{234}Th, unless the date of last chemical separation (in the case of UF_6, gas transfer) is known, in which case the code can make the appropriate correction. The software calculates a relative efficiency curve based on the measured gamma-rays and detector efficiency, as run in "physical efficiency" mode using eleven peaks from uranium isotopes and daughter products: 143, 163, 185, and 205 keV from ^{235}U and 258, 742, 766, 880, 883, 945 and 1001 keV from ^{234m}Pa, and requires no calibration.[4] The software can also be used in "planar" mode, which relies on low energy x-ray and gamma-ray peaks to quantify the isotopic fractions, however 30B cylinder walls are too

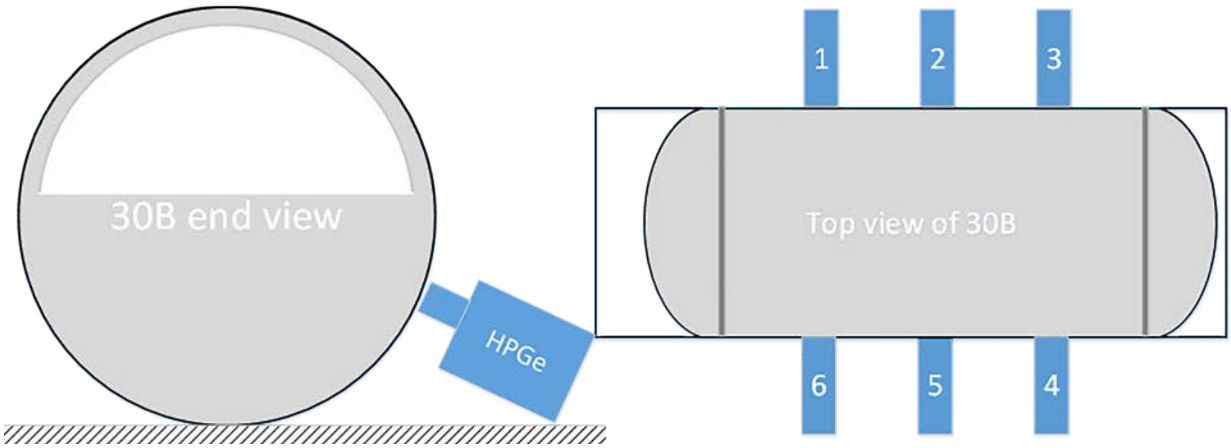


Figure 1: Geometry of the six position analyses around a 30B cylinder. The grey area is to suggest the expected distribution of UF_6 inside the cylinder.

thick for this method. The ratio of N^i and N^k is the enrichment as atom % but is converted by the software to weight % which is more commonly employed in the nuclear industry.

$$\frac{N^i}{N^k} = \frac{C(E_j^i)}{C(E_l^k)} * \frac{T_{1/2}^i}{T_{1/2}^k} * \frac{BR_l^k}{BR_j^i} * \frac{RE(E_l)}{RE(E_j)} \quad (1)$$

3. Results

To assess the accuracy of the measured apparent enrichment, the percent difference between the measured and declared weight% value was calculated following Eqn. 2.

$$\%\ difference = 100 * \frac{\left(^{235}U_{measured} - {}^{235}U_{declared}\right)}{^{235}U_{declared}} \quad (2)$$

Six cylinders showed accurate, relatively constant uranium enrichments at all six positions around the cylinder, represented by calculated apparent enrichments within 20% of the declared value at all positions (**Fig. 2 left**). Three cylinders were found to have different apparent enrichments on either side of the cylinder with side 1-2-3 showing one consistent apparent enrichment within 10% of the declared value, and side 4-5-6 showing a different apparent enrichment (**Fig. 2 right**).

The spectra were examined in Peak Easy v4.98.1 [8] to determine the count rates of the 185.7 keV peak from ^{235}U and the1001 keV peak from ^{234m}Pa. Count rates, relative to the live time, measured by five detectors in six positions around apparently heterogeneous cylinders are shown in **Fig. 3.** There was a correlation between slightly lower 185.7 keV count rate relative to the other side (~5% difference in count rate) and decreased apparent enrichments calculated as calculated by FRAM™ in these heterogeneous cylinders. Variations in 185.7 keV count rate by ~4% as a function of measurement position at the cylinder have been noted by Dufour et al. (2019) [9] so these slight variations in 185.7 keV count rate not significant. The 1001 keV peak count rates show large variation between the two sides of the heterogeneous cylinders, with the apparently low-enriched side showing up to 6x higher ^{234}Th-^{234m}Pa count rate (**Fig. 3**).

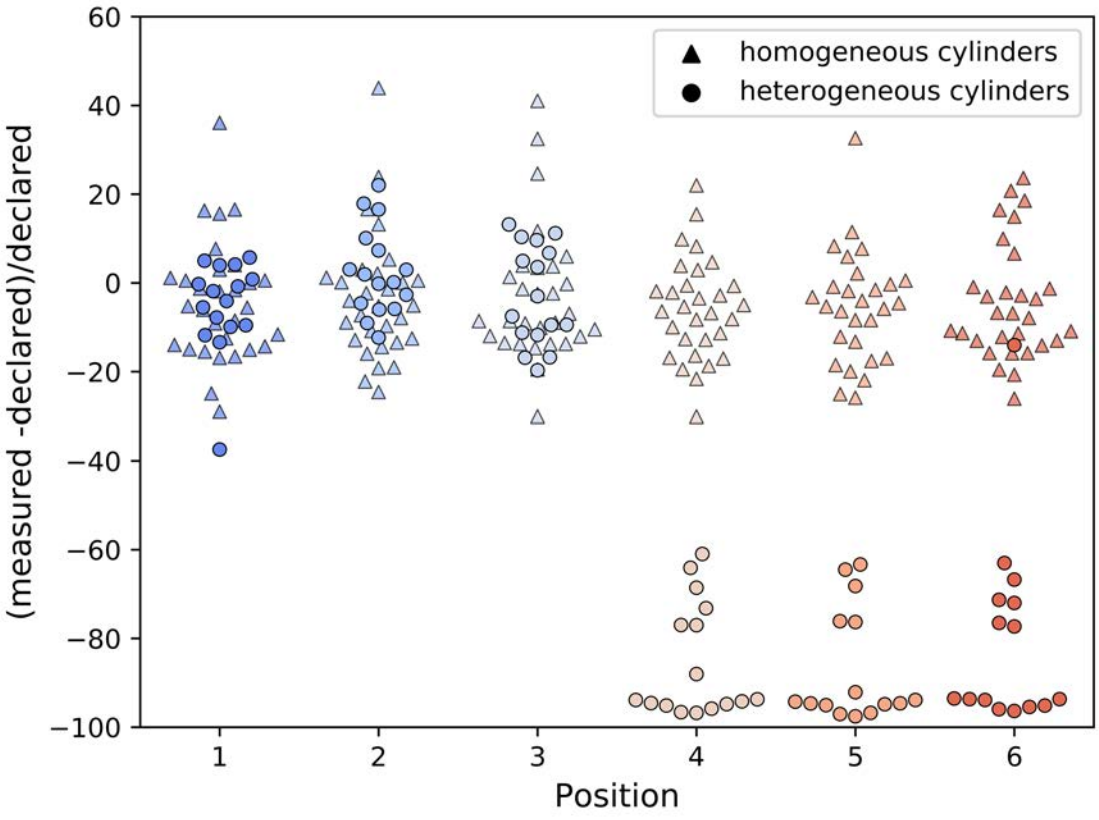


Figure 2: Accuracy of position analyses around the six "homogeneous" cylinders (triangles) and three "heterogeneous" cylinders (circles) plotted as the percent difference of the measured enrichment relative to the declared cylinder tag enrichment.

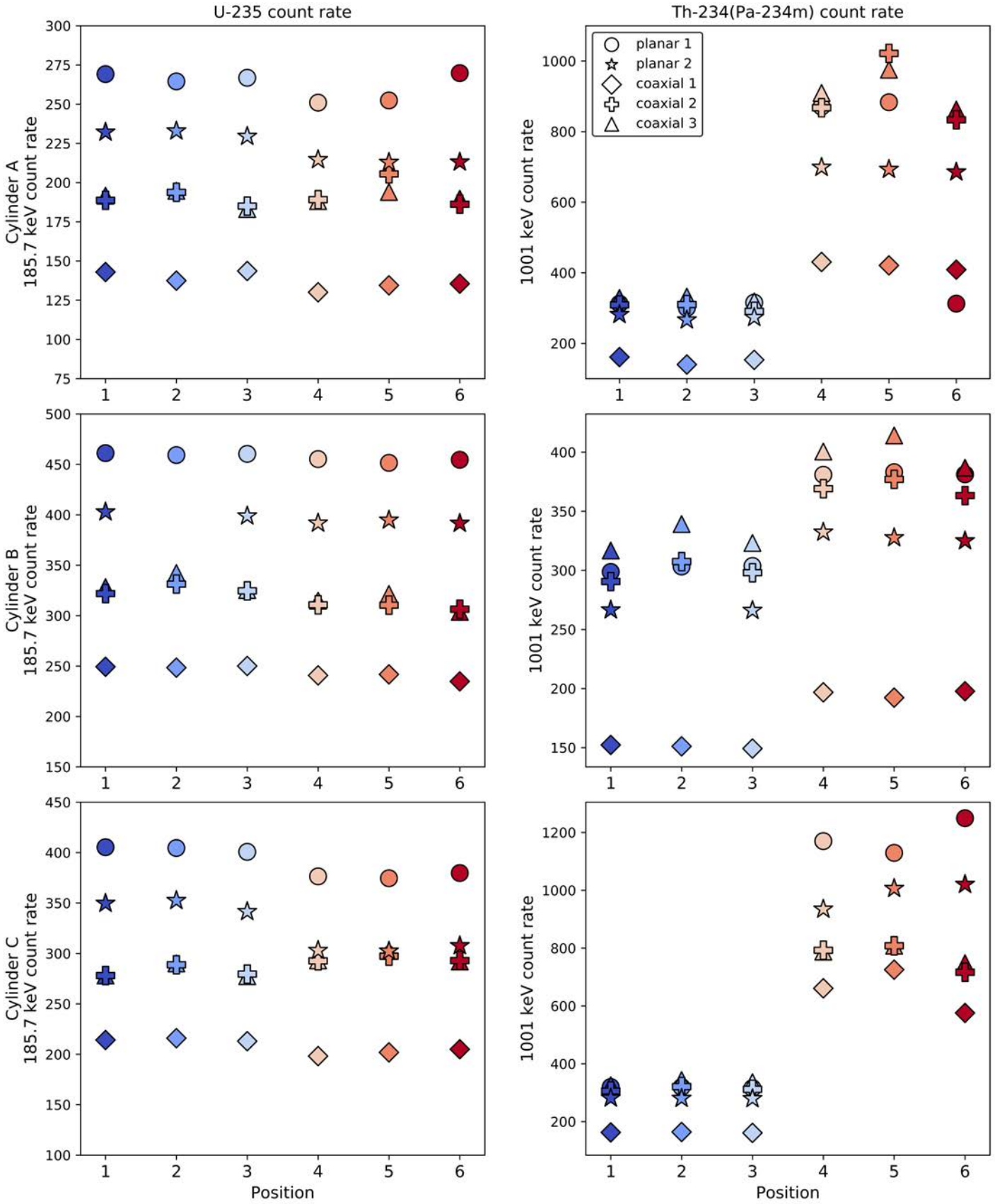


Figure 3: The deadtime corrected 185.7 keV peak count rate (left column) and 1001 peak count rate (right column) plotted by position, in counts per second, for three heterogeneous cylinders. Count rates for Cylinders A, B and C are groups in the top, middle and bottom plots, respectively. Markers denote five different detectors and colors correspond to position (1, 2, 3 = blues; 4, 5, 6 = reds). Measurements at positions 1, 2, and 3 produce accurate ^{235}U apparent enrichments while the measurements at positions 4, 5, and 6 shows apparent enrichment values that are consistently low. The Y-axis range varies because the three cylinders contain different U enrichments. Measurement uncertainty propagated from the total counts is within the data points.

4. Discussion

4.1 Detector performance

Because each detector was used at every position around the cylinder, these data can be used to evaluate detector performance without an added bias of position-dependence. When the homogeneous cylinders are considered, the planar detectors measure, on average, slightly more accurate uranium enrichments (**Fig. 4**). Using FRAM™ to calculate an enrichment, planar detectors average ± 9% accuracy while coaxial detectors average within ±13% accuracy, but the difference between planar and coaxial detector accuracy is not statistically significant (**Fig. 4**). Therefore, we conclude there are no significant differences in the accuracy of the enrichment calculated from spectra measured by planar vs. coaxial detectors and that "coaxial" mode on FRAM™ can be applied to either detector type. Given this, we can generally evaluate the effects of detector placement independently from detector performance.

4.2 Heterogeneous apparent enrichments on opposite side of the cylinder

Six of the cylinders analyzed at six positions showed no significant differences in apparent enrichment related to detector position. Data collected from all positions at these six cylinders showed a similar range of accuracy, as determined by the percent difference between the measured and declared enrichment (**Fig. 2 triangles**). However, three of the cylinders analyzed showed extremely different apparent enrichments related to detector position on each side of the cylinder (**Fig. 2 circles**). Disequilibrium effects between parent and daughter isotopes can be ruled out as the cause of low apparent enrichments, because all eight of these low-enriched cylinders were analyzed between 6 months and 3.5 years after their fill dates. It is therefore assumed that secular equilibrium has been attained on a whole-cylinder scale. While variations can be observed from heel deposits [10], they are not likely to be as systematic as observed in these studies. Thus, we propose that heterogeneous UF_6 and daughter product distribution within the cylinders caused the low apparent enrichment observed on one side of the cylinder. It is assumed that detector placement relative to the filling profile illustrated in **Fig. 1** would have captured a homogeneous UF_6-daughter mixture. However, given that UF_6 can sublime when heated, asymmetric solar heating of the cylinder could result in a heterogeneous UF_6 distribution. A cylinder oriented east –west in a storage yard is likely warmed to a greater degree on its southern side due to solar heating. This process could result in partial UF_6 sublimation and removal from the warmed side, leaving behind nonvolatile daughter nuclides like Th and Pa on the warmed side. The cylinder wall temperatures were not measured during this analytical campaign, so it is important to note that proposed sublimation from solar heating

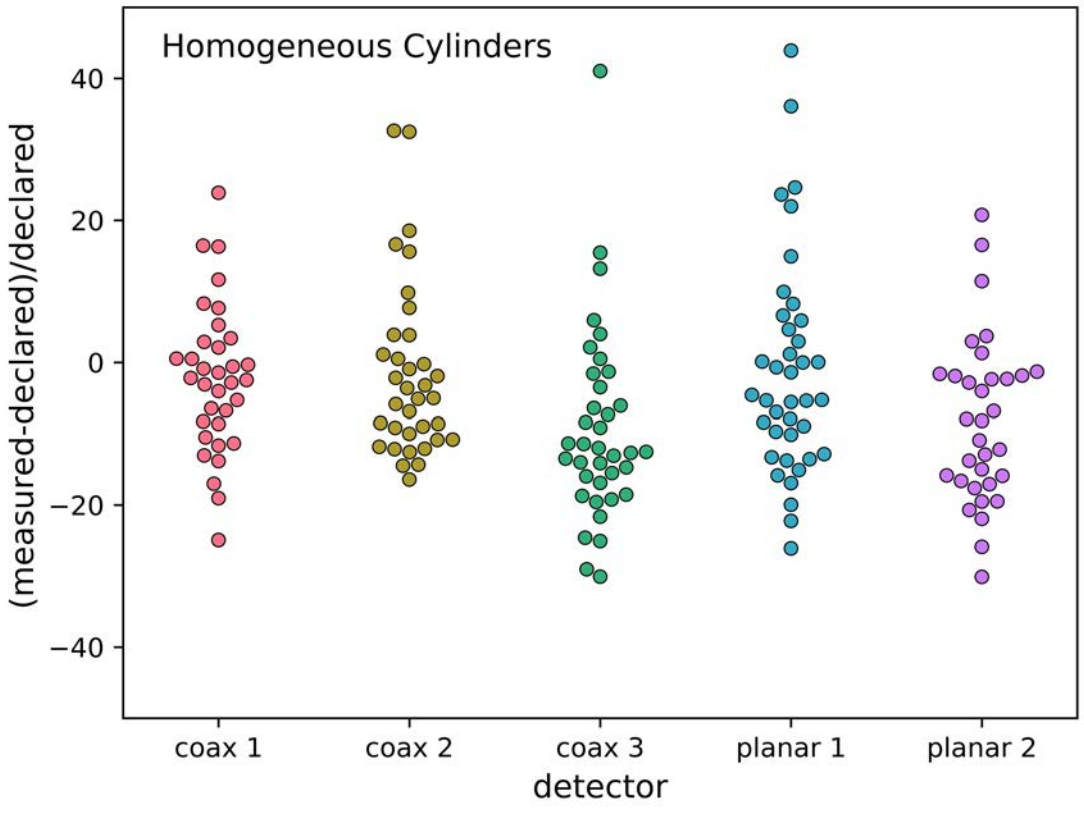


Figure 4: Accuracy of each detector for the six "homogeneous" cylinders plotted as the percent difference of the measured enrichment relative to the declared cylinder tag enrichment.

is only speculative, and further experiments should be completed to test this phenomenon.

The removal of parent U isotopes from the daughter Th products would result in localized disequilibrium between U and Th/Pa on the heated side of the cylinder. The added UF_6 to the cooler side would have minimal impact on the gamma-ray signal observed, as the 185.7 keV line would be self-shielded and the daughter nuclides would not provide significant contributions for months. However, the warmed side would likely have observable gamma-ray effects because UF_6 shielding was removed, allowing for a stronger daughter-product signal. This was observed in several cylinders analyzed here; **Fig. 3** compares the 185.7 and 1001 keV peak count rates at six positions around three cylinders. These cylinders were measured with five detectors, and each are plotted to show that the measured count rates were systematic across multiple detectors and not due to improperly tuned or faulty detectors. For each of these cylinders, positions 1, 2, and 3 show accurate apparent enrichments while positions 4, 5, and 6 show extremely low apparent enrichments. There is a clear 1001 keV peak count rate offset between each side of the cylinder; the low apparent enrichment side has a significantly higher 1001 count rate (**Fig. 3**). This is also true for the 766 keV peak from ^{234m}Pa, implying that the daughter product activity was measurably higher on the suspected warm side of the cylinder than on the cool side.

Additionally, there is a strong relationship between dead time and position. The majority of the spectra collected on the apparently low-enriched side (positions 4, 5, and 6) of the cylinders have exceedingly high dead times (>60% of the total count time), whereas all spectra collected on the "normally" enriched side (positions 1, 2, and 3) have dead times < 60% of the total count time (**Fig. 5**). These high dead times occurred in all detectors used. High dead times are presumably due to the relatively high daughter product activity. This would suggest that cylinders with significant dose differences between their sides may produce erroneous data on the side with higher dose readings, because dose generally correlates to the abundance of high activity ^{234m}Pa. Therefore, dose rate measurements and/or dead time monitoring could be used in the field to quickly identify potentially heterogeneous UF_6 distribution within cylinders.

4.3 Measuring spatial disequilibria within UF_6 cylinders

Because FRAM™ relies on a ratio between the 1001 keV peak and the 185.7 keV peak, it is sensitive to spatial homogeneity between parent U and daughter Th and Pa isotopes. This means FRAM™ calculates a lower apparent enrichment on the warm side as there is increased signal related to the high energy gamma-ray peak signal from ^{238}U. The traditional enrichment meter method is less susceptible to the effects of spatial disequilibria as it only uses the intensity of the ^{235}U peak; however, it is not totally immune. Data

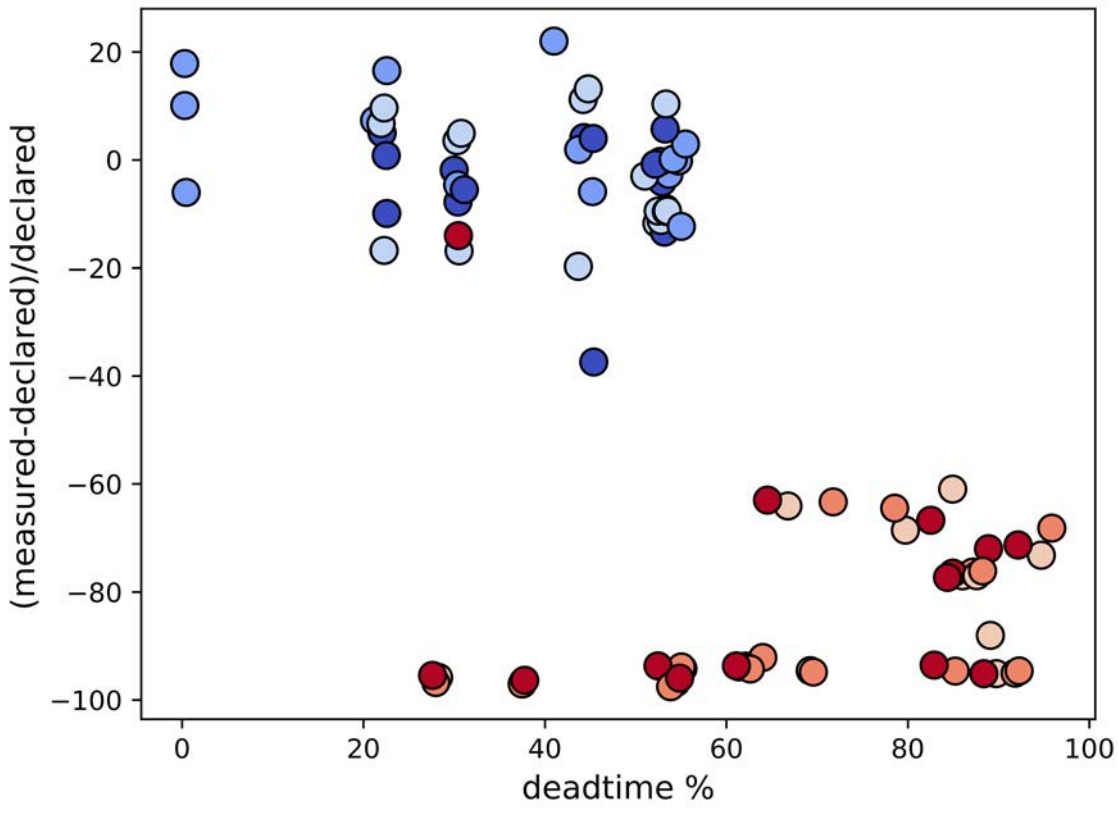


Figure 5: Dead time (% of total time) plotted against the accuracy of the ^{235}U enrichment measurement for the three heterogeneous cylinders. Accuracy is calculated as the percent difference between the measured and declared enrichments. Colors correspond to position (1, 2, 3 = blues; 4, 5, 6 = reds) as in previous figures.

in **Fig. 6** show a comparison between the spectra analyzed with FRAM™ and the 185.7 keV enrichment meter method for the three cylinders presented in **Fig. 3.** The enrichment meter was calibrated on two homogeneous cylinders from this facility with an enrichment of 4.95% and 0.71% ^{235}U. It was assumed that all wall thicknesses were the same, so no corrections were made for wall thickness. The results from the enrichment meter data treatment on the warmed side are much more accurate than the FRAM™ results but are still slightly depressed from the true enrichment (**Fig. 6**). As shown in **Fig. 3,** the 185.7 keV count rates are very similar between the two sides. This is likely because, although some UF_6 was suspected to be have been removed from the warm-side wall, a deposit that was close to infinite thickness for the low energy 185.7 keV gamma-ray remained. Thus, the peak is only slightly altered by suspected material removal. Additionally, the higher dead times resulting from increased daughter product activity likely impacted the apparent enrichments calculated by the enrichment meter. Going forward, the possibility of heating-induced spatial disequilibria and its detection by different spectral analysis methods should be considered when using these methods in the field. Because the cause of the heterogeneities reported here is only speculative (i.e., we do not have confirmation that the heterogeneous cylinders were exposed to more sunlight than the homogeneous cylinders), further study of this phenomenon in laboratory and field settings is encouraged.

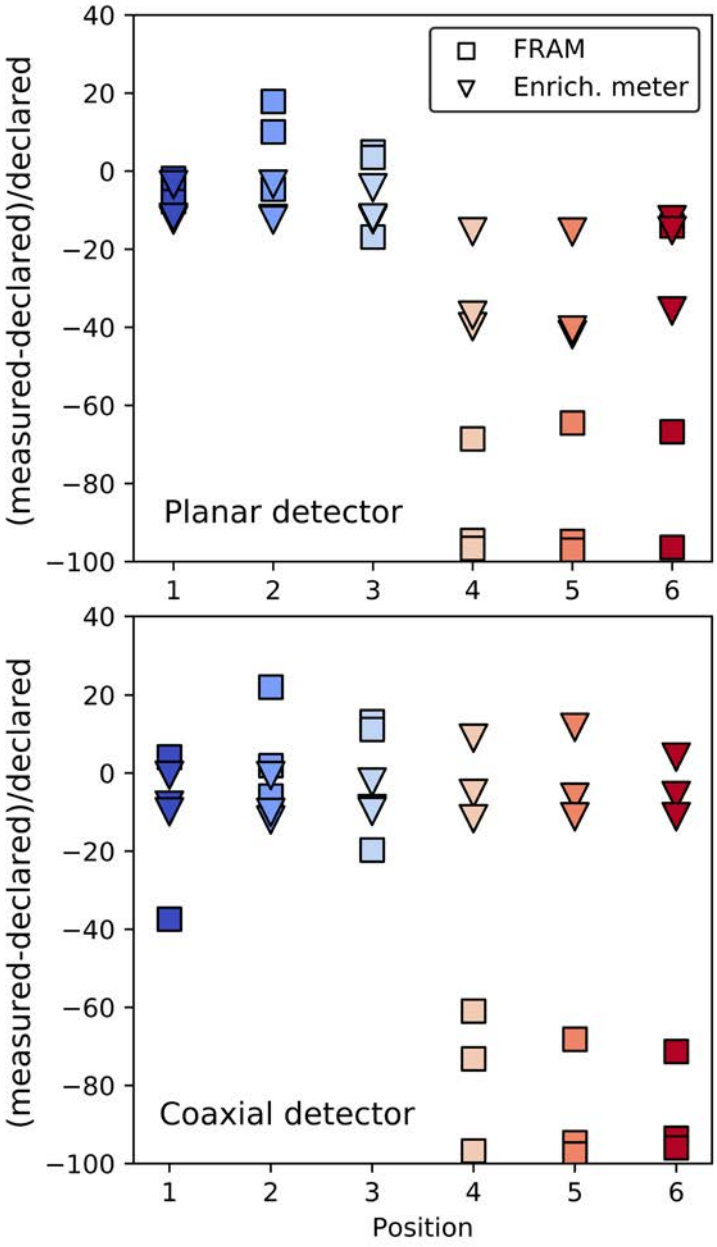


Figure 6: Comparison between FRAM™ (using peaks between 120 – 1010 keV) and the enrichment meter method (using only 185.7 keV) on heterogeneous cylinders. Data measured with two detectors from three heterogeneous cylinders are shown. Because FRAM™ relies on a ratio between the 185.7 and 1001 peaks, it is more sensitive to spatial disequilibria between UF_6 and its daughters in the cylinder. Colors correspond to position (1, 2, 3 = blues; 4, 5, 6 = reds).

5. Conclusions

The application of FRAM™ to determine enrichment of UF_6 in thick-walled cylinders can be performed with similar accuracy using either planar or coaxial detectors. However, caution should be exercised when applying FRAM™ or other enrichment meters that require the use of uranium daughter isotopes for measurements on UF_6 cylinders, especially if there are significant discrepancies between deadtime and/or dose rates at different locations around the cylinders. Cylinder placement within a storage yard should be noted when measurements are made, and if uneven solar heating is suspected, measurements should be taken at multiple points around a cylinder to assess the accuracy of the uranium enrichment measurement. Further studies are needed to assess the effects of solar heating on UF_6 distribution within a cylinder.

6. Acknowledgements

This work was supported by the U.S. National Nuclear Security Administration (NNSA) Office of International Nuclear Safeguards.

7. References

[1] Sampson, T. E.; Kelly, T.A. “PC/FRAM: a code for the nondestructive measurement of the isotopic composition of actinides for safeguards applications” *Appl. Radiat. Isot*. 1997, 48, 1543–1548.

[2] Sampson, T. E.; Verrecchia, G. P. D.; Swinhoe, M. T.; Schwalbach, P.; Gustafsson, J.; Anderson, A. M.; Myatt, J.; Metcalfe, B. “Test and Evaluation of the FRAM Isotopic Analysis Code for EURATOM Applications,” *Annu. Meet. Proc. Inst. Nucl. Mater. Manage*., July 25-29, 1999, Phoenix, Arizona.

[3] Vo, D. T.; Sampson, T.E. “Methods for uranium isotopic analysis with the FRAM isotopic analysis code”, *Annu. Meet. Proc. Inst. Nucl. Mater. Manage*., July 25-29, 1999, Phoenix, Arizona.

[4] Vo, D. T.; Sampson, T.E. FRAM Version 5, User Manual, 2011, Los Alamos National Security, LLC (2011)

[5] Berndt, R.; Franke, E.; Mortreau, P. “^{235}U enrichment or UF_6 mass determination on UF_6 cylinders with non-destructive methods” *Nucl. Instrum. Methods A*, 612 (2010) 309–319.

[6] Greaney, A. T.; Smith, S. K.; Venkataraman, R.; Richards, J. M.; Fugate, G.A. “Comparison of Gamma-Ray Spectral Analysis Methods for Thick-Walled UF_6 Cylinders” *Nucl. Instrum. Methods A*, 2020, 977, 164291

[7] Greaney, A. T.; Smith, S. K.; Venkataraman, R.; Richards, J. M.; Fugate, G.A. “Applications of HPGe-detected high energy gamma rays toward quantifying neutron emission rates and ^{234}U enrichment in UF_6 cylinders” *Nucl. Instrum. Methods A*, 2020, 163912

[8] Rooney, B.; Garner, S.; Felsher, P.; Karpius, P. PeakEasy 4.98, 2018, Los Alamos National Laboratory, Release LA-CC-13-040.

[9] Dufour, J.-L.; Pepin, N.; Deyglun, C.; Weber, A.-L. “Optimisation and uncertainty estimation of the enrichment meter measurement technique for UF6 cylinders”, *ESARDA Bull*., 2019, 59, 2-10.

[10] McFerran, N.; Canion, B.; McDonald, B.; Kulisek, J.; Dreyer, J.; Labov, S.; Enqvist, A. “Gamma-ray spectrum variations for surface measurements of uranium hexafluoride cylinders” *Nucl. Instrum. Methods A*, 2020, 961, 163675.

Fuel rod classification from Passive Gamma Emission Tomography (PGET) of spent nuclear fuel assemblies

Riina Virta[1,2], Rasmus Backholm[3], Tatiana A. Bubba[3], Tapio Helin[4], Mikael Moring[2], Samuli Siltanen[3], Peter Dendooven[1], Tapani Honkamaa[2]

[1]Helsinki Institute of Physics, University of Helsinki, Finland.
[2]Radiation and Nuclear Safety Authority (STUK), Finland.
[3]Department of Mathematics and Statistics of the University of Helsinki, Finland.
[4]School of Engineering Science of LUT University, Lappeenranta, Finland.
Email: riina.virta@helsinki.fi

Abstract:

Safeguarding the disposal of spent nuclear fuel in a geological repository needs an effective, efficient, reliable and robust non-destructive assay (NDA) system to ensure the integrity of the fuel prior to disposal. In the context of the Finnish geological repository, Passive Gamma Emission Tomography (PGET) will be a part of such an NDA system. We report here on the results of PGET measurements at the Finnish nuclear power plants during the years 2017-2020. The PGET prototype device developed by IAEA and partners was used during 2017-2019, whereas an updated device was used in 2020. The PGET device contains two linear arrays of collimated CdZnTe (CZT) gamma ray detectors installed opposite each other inside a torus. Gamma activity profiles are recorded from all angles by rotating the detector arrays around the fuel assembly that has been inserted into the center of the torus. Image reconstruction from the resulting tomographic data is defined as a constrained minimization problem with the function being minimized containing a data fidelity term and regularization terms. The activity and attenuation maps, as well as detector sensitivity corrections, are the variables in the minimization process. The regularization terms ensure that prior information on the (possible) locations of fuel rods and their diameter are taken into account. Fuel rod classification, the main purpose of the PGET method, is based on the difference of the activity of a fuel rod from its immediate neighbors, taking into account its distance from the assembly center. The classification is carried out by a support vector machine. We report on the results for 10 different fuel types with burnups between 5.72 and 55.0 GWd/tU, cooling times between 1.87 and 34.6 years and initial enrichments between 1.9 and 4.4%. For the 77 fuel assemblies measured, the total misclassification rate including misclassifications of missing fuel rods, present rods and water channels, was 0.94% for the Olkiluoto campaigns and 0.66% for the Loviisa campaigns. Further development of the image reconstruction method is discussed. We conclude that the combination of the PGET device and our image reconstruction method provides a reliable base for fuel rod classification. The method is well-suited for nuclear safeguards verification of BWR fuel assemblies in Finland prior to geological disposal. For VVER-440 assemblies, some further work is needed to investigate the ability to detect missing rods near the center of the assembly.

Keywords: geological repository; iterative reconstruction; nuclear fuel; nuclear safeguards; PGET

1. Introduction

Finland will start disposal of spent nuclear fuel in a deep geological repository around the mid-2020's, likely the first country in the world to do so. The construction of the underground facility is ongoing at Olkiluoto, Eurajoki. For safeguards purposes, the disposal needs an effective, efficient, reliable and robust non-destructive assay (NDA) system for spent nuclear fuel verification. The combination of Passive Gamma Emission Tomography (PGET) and Passive Neutrino Albedo Reactivity (PNAR) [1] will be used for this, because a combination of NDA methods will give better confidence of the accuracy of the declaration [2]. At the end of 2017, the International Atomic Energy Agency (IAEA) approved PGET for the verification of spent nuclear fuel. The method can be deployed on all fuel types once the performance has been validated. Other state-of-the-art methods (e.g. Fork detectors [3]) can only detect a gross deviation of material in the fuel assembly, but PGET has been demonstrated to accomplish reliable rod-level detection [4], [5]. This is crucial for ensuring effective nuclear safeguards of the final repository.

First PGET images of fuel assemblies of different types and with varying cooling times and burnups were published in [6] and [7]. These results indicated the need for improved image reconstruction and analysis methods. Recently, we proposed a method in which attenuation and activity images are simultaneously reconstructed from the data by formulating the reconstruction as a constrained minimization problem and solving it with a Levenberg-Marquardt type of method [8]. In the present study, we apply this method to PGET data taken at the two Finnish nuclear power plants (NPPs). We show that high-quality results are produced, enabling to detect a single missing rod in a wide range of different assembly types and parameters. The results are significant in both the context of Finnish safeguards and the global context of nuclear fuel disposal.

2. Materials and methods

2.1 Imaged spent nuclear fuel

PGET measurements were performed at the two Finnish NPPs, Olkiluoto (OL) and Loviisa (LO), during the years 2017-2020. A total of 77 individual assemblies and 4 non-fuel items (e.g. fuel dummies) were measured. The assembly types were VVER-440 at Loviisa and 9 BWR type assemblies at Olkiluoto (SVEA-64, SVEA-96, SVEA-96 OPTIMA, SVEA-100, ATRIUM10, GE12, GE14, 9x9-1AB and 8x8-1). The assemblies were chosen to cover a wide range of operating parameters: burnup from 5.72 to 55.0 GWd/tU, cooling time from 1.87 to 34.6 years and initial enrichment from 1.9 to 4.4 %. We report results from a selection of measurements from the campaigns in Olkiluoto in 2017 and 2019 (OL17, OL19) and Loviisa in 2018 and 2020 (LO18, LO20). The main fuel characteristics are presented in Table 1. The measured assemblies were chosen to best reflect the strengths and future development areas of our software and to represent a wide range of fuel characteristics and specifics (such as 1 to 3 completely removed fuel rods, burnable absorber rods or partial rods), measurement campaigns and assembly types.

2.2 PGET device

The PGET device (see [6] for details) contains 174 highly collimated CdZnTe (CZT) gamma-ray detectors arranged in 2 linear arrays on opposite sides of a torus. These arrays are rotated around the fuel assembly to collect data from all angles. All measurements were done underwater in spent fuel ponds. Measurements were conducted with different numbers of projection angles and measurement times per angle. Some assemblies were measured at different vertical, horizontal and rotational positions. Most of the measurements were done with the PGET central plane at a height of 0.5-1.1 m from the bottom of the assembly, the exact position varying between measurement campaigns. In this work, we show results from measurements with 360 angles and 800 ms projection time per angle (OL17, OL19, LO18) or 924 ms projection time per angle (LO20). Data were collected in four energy windows. The lowest two windows were 400-600 keV and 600-700 keV. The third and fourth windows were 700-1500 keV and above 1500 keV for OL17, OL19 and LO18, and 700-2000 keV and 2000-3000 keV for LO20. The choice of these windows is related to the gamma peaks of the radioactive nuclei present. All image reconstructions shown in this work are from the 600-700 keV window which contains the 661 keV gamma peak from Cs-137, the most abundant gamma ray emitter in spent fuel.

The so-called prototype PGET device was used for the campaigns in 2017, 2018 and 2019. However, some individual detectors were replaced in-between campaigns. The campaign at Loviisa in 2020 used a new PGET device with a more compact design for easier handling and a slightly optimized collimator.

2.3 Data analysis and image reconstruction

2.3.1 Image reconstruction algorithm

The core idea of our image reconstruction strategy is to recover simultaneously the activity and the attenuation map from the data. To enhance performance with real data, we also tweak the detector sensitivity correction during the reconstruction process. This is all attained by formulating the reconstruction task as a constrained minimization problem where the activity and the attenuation maps, as well as the sensitivity correction coefficients, are the variables, and the function being minimized consists of a least squares data fit term and regularization terms [9]. Namely, the minimization problem takes the following form:

$$\min_{\lambda,\mu,c}\begin{Bmatrix} \| H(\mu)\lambda - C(c)\,s \|_2^2 + \alpha_\lambda \| R_\lambda \lambda \|_2^2 \\ +\alpha_\mu \| R_\mu \mu \|_2^2 + \alpha_c \| \log(c) \|_2^2 \\ +\alpha_s \| 1^T \left(s - C(c)\,s \right) \|_2^2 \end{Bmatrix} \tag{1}$$

with the bounds

$$A\begin{bmatrix}\lambda \\ \mu\end{bmatrix} \leq b. \tag{2}$$

#	Type	BU(GWd/tU)	CT(a)	Campaign	Characteristics
1	VVER-440	55.0	6.8	LO20	3 missing (corner), burnable absorber rods
2	VVER-440	42.0	7.9	LO20	1 missing (corner)
3	VVER-440	43.0	2.7	LO18	Burnable absorber rods
4	VVER-440	22.8	27.6	LO18	Activity and attenuation gradient
5	SVEA-64	32.6	20.7	OL19	2 rods in the reactor for 2/4 fuel cycles
6	SVEA-64	32.9	20.7	OL19	Intra-rod activity differences
7	SVEA-96	40.7	8.9	OL17	2 missing (bottom-left and bottom-right quarters)
8	SVEA-96 OPTIMA	39.8	13.7	OL19	Measurements at two heights
9	ATRIUM10	49.7	7.9	OL19	2 missing, measurements at two heights
10	8x8-1	18.6	34.6	OL19	Long CT
11	9x9-1AB	35.0	20.7	OL19	Different assembly type compared to others presented
12	GE12	43.1	11.7	OL19	Measurements at two heights

Table 1: Measured fuel assemblies and their characteristics: assembly type, burnup (BU), cooling time (CT) and measurement campaign (at Loviisa (LO) and Olkiluoto (OL) during the years 2017-2020). The characteristics are from the licence-holder's declaration.

Here λ and μ are the vector forms of the discrete activity and attenuation maps, respectively, and the vector *c* consists of the coefficients used to correct for the detector sensitivity differences in the data sinogram *s*. The least squares term $\| H(\mu)\lambda - C(c)s \|_2^2$ measures how well λ, μ and *c* fit the data *s*. The rest of the terms are regularization terms, which can be understood to incorporate some kind of a priori information in the reconstruction process, namely, they predispose the algorithm towards certain kinds of solutions. The regularization parameters α_λ, α_μ, α_c and α_s balance the contribution of each term.

In more detail, in the data fit term $\| H(\mu)\lambda - C(c)s \|_2^2$, $H(\mu)$ is the system matrix depending on the attenuation map μ: a detailed description of how to implement the system matrix $H(\mu)$ can be found in [8]. The product $H(\mu)\lambda$ is the forward projection, i.e., a sinogram simulated using λ and μ. The system matrix models the effects of attenuation and collimator blurring. The spatial responses are computed based on the given dimensions of the device and assuming an opaque collimator. *C(c)* is a diagonal matrix formed from the detector sensitivity correction coefficients *c* so that in the product vector *C(c)s*, which is the sensitivity-corrected sinogram in vector form, all the elements of *s* corresponding to values from one detector are multiplied by one coefficient in *c*.

The effect of the regularization terms $\alpha_\lambda \| R_\lambda \lambda \|_2^2$ and $\alpha_\mu \| R_\mu \mu \|_2^2$ depends on the choice of the matrices R_λ and R_μ. From the two types of regularization terms introduced in [8], to which we refer for the precise mathematical formulation, we considered in this paper only the geometry aware prior. We consider this prior to be suitable in the context of verification of spent nuclear fuel. This choice for the matrices R_λ and R_μ assumes that the locations and diameters of possible rods, whether they are actually present or not, are known. In practice, this prior asserts that the solution should look approximately like it is made out of rods, each having a uniform activity, with the predefined diameters in the predefined locations. A rod can be missing in these solutions by having zero activity and the attenuation of water.

The term $\alpha_c \| \log(c) \|_2^2$ penalizes large absolute values of $\log(c)$ and so prefers solutions where the coefficients in *c* are close to one, i.e., the corrections made are not large. The last term, $\alpha_s \| 1^T (s - C(c)s) \|_2^2$, where 1 is a vector of ones, requires that the sum of all counts in the corrected sinogram $C(c)s$ is close to the sum of all the counts in the sinogram *s*. The scope of this term is to keep the same "overall scale" of the sinogram after correction. The data sinogram *s* has actually already undergone a preliminary detector sensitivity correction, and the role of the coefficients *c* is only to fine-tune this. The approach we first introduced in [8] did not include the correction coefficients *c*.

The bounds (2) on the activity values and attenuation coefficients used in the minimization process are such that they exclude the possibility of a material with high activity and low attenuation coefficient, which is a physically unlikely case. In practice, to define these bounds, one must give lower and upper bounds for the attenuation coefficients and an upper bound for activity values (the lower bound for activity is always considered to be zero). The bounds can be better understood by visualizing them in the attenuation-activity plane, as depicted in Fig. 1, where the allowed values form a triangle.

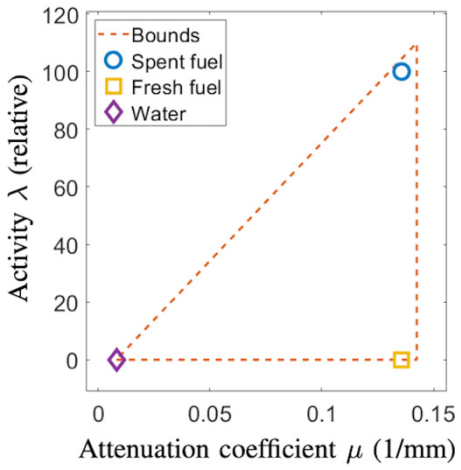


Figure 1: Example of the bounds for the activity values and attenuation coefficients (for 661 keV gamma rays) used in the minimization process. The values inside the triangle are allowed.

To estimate the activity (and attenuation) bounds and to build the matrices R_λ and R_μ for the geometry aware prior, rod locations and diameters are needed. We retrieve this information by identifying the assembly type from an "off-the-shelf" filtered back-projection (FBP) reconstruction, and then using the known grid and rod dimensions for that type of assembly. In case the assembly type has water channels, these are not assumed to be known: rather a full grid of rods is assumed.

In practice, the bounds for attenuation coefficients are estimated by considering the measurement energy window and the materials assumed to be imaged. Once the attenuation coefficients for water and rods are estimated, and locations and diameters for the rods are computed from having identified the assembly type, the upper bound for activity values is estimated by simulating a sinogram using rods with some uniform activity value *a*. The upper bound for activity values is then set so that the ratio of *a* and the maximum value of the simulated sinogram is the same as the ratio of the upper bound and the maximum value of the data sinogram.

Finally, the minimization problem (1) is solved iteratively using a Levenberg-Marquardt type of algorithm [10], that takes the bounds (2) into account [11]. All the reconstructions are computed using 120 evenly spaced measurement angles (every third angle from the total of 360 measurement

angles), which provides a good balance between computational efficiency and quality of the reconstruction in order to perform a reliable rod classification.

2.3.2 Rod classification

To classify rods into missing or present ones, the basic idea is to consider the difference of a rod's activity from the average activity of its immediate neighbors plotted against its distance from the assembly center. Rod activity values are computed as a weighted average of the values of all the pixels that consist at least partly of the rod. The weights are proportional to the fraction of the area of the pixel that is covered by the rod, namely, the border pixels contribute less to the average. The sum of the weights is not normalized to one and as the classification is done on rod-level, the average attenuation and activity values for individual rods can exceed the bounds, as can be seen in Fig. 3.

The classification is carried out by training a support vector machine [12] on the reconstructions of training data sets from varied assemblies, including mock-up fuel and real spent fuel assemblies with different rod placements and intensities. Data from assemblies of mock-up fuel constructed from neutron-activated cobalt rods containing Co-60 were measured at the Atominstitut at the Technical University of Vienna where the PGET system is prepared for spent-fuel measurements [6]. At first, rather than classifying all the rods at once, we begin classifying them as missing one by one starting from the most likely case. After classifying a rod as missing, the differences the classification is based on are recalculated by discarding from the calculations of the neighbor average all the rods classified as missing at that point. This prevents the missing rods from bringing down the neighbor average.

When a rod shows to have at least a couple of missing neighbors, we compute another neighbor average of only the rods that are classified as missing and use that as well: if a rod's activity is close to the average of its missing neighbors' activities, it is classified as missing. This comparison allows classification of larger missing rod areas where rods might not have neighboring rods present. For the second comparison, a second support vector machine was trained using the same reconstructions as training data as was used for the first support vector machine.

The final result is a plot where missing and present rods are separated rather than by a straight line, by a possibly more complex curve.

3. Results

Good quality reconstructions are needed as a basis for detecting anomalies and accurately classifying spent fuel rods. In the following we present results from data from a selection of measurements (see Table 1) and demonstrate the ability to detect missing rods, burnable absorber rods, water channels and intra-rod activity differences.

3.1 Missing and abnormal rod detection

Missing rod detection in hexagonal VVER-440 assemblies is demonstrated for two assemblies in Fig. 2 with activity and attenuation reconstructions and a rod classification figure. Assembly #1 in the top row (BU 55.0 GWd/tU, CT 6.8 a) has three missing fuel rods and five burnable absorber rods (see Section 3.2.) and assembly #2 in the bottom row (BU 42.0 GWd/tU, CT 7.9 a) has one missing rod. In both cases the missing rods are clearly visible in the reconstructions and correctly classified by the algorithm. The central water channel is also classified as missing in both cases (see Section 3.3.).

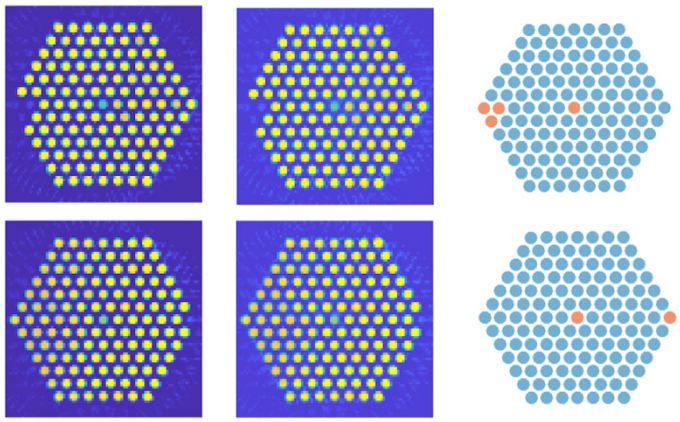


Figure 2: Activity (left column) and attenuation (middle column) reconstructions and classification into missing (orange) and present (blue) rods (right column) for two VVER-440 assemblies. Assembly #1 in the top row has three missing rods, a central water channel and 5 burnable absorber rods near the corners. Assembly #2 in the bottom row has one missing rod and a central water channel.

Fig. 3 shows two classification metric plots for the VVER-440 assembly #1 with three missing rods and five burnable absorber rods (see also Fig. 2, top row). Each circle represents a rod position and the color denotes the ground truth rod type. The rods get grouped by their characteristics and the rod classification is based on the kind of plots on the right. On the left, the three missing rods show low attenuation and low activity, and on the right they have a negative activity difference compared to their neighboring rods. The burnable absorber rods get grouped with the present rods but the water channel position deviates from the rest of the rods and is correctly classified as missing.

Activity and attenuation reconstructions and rod classification for the SVEA-96 assembly #7 (BU 40.7 GWd/tU, CT 8.9 a) with two missing rods is shown in Fig. 4, top row. The four innermost rods are part of the water channel. The bottom row shows the reconstructions and the classification for the SVEA-64 assembly #5 (BU 32.6 GWd/tU, CT 20.7 a) with two fuel rods that have been in the reactor for only two out of the normal four fuel cycles and thus have a different burnup than the other rods in the assembly.

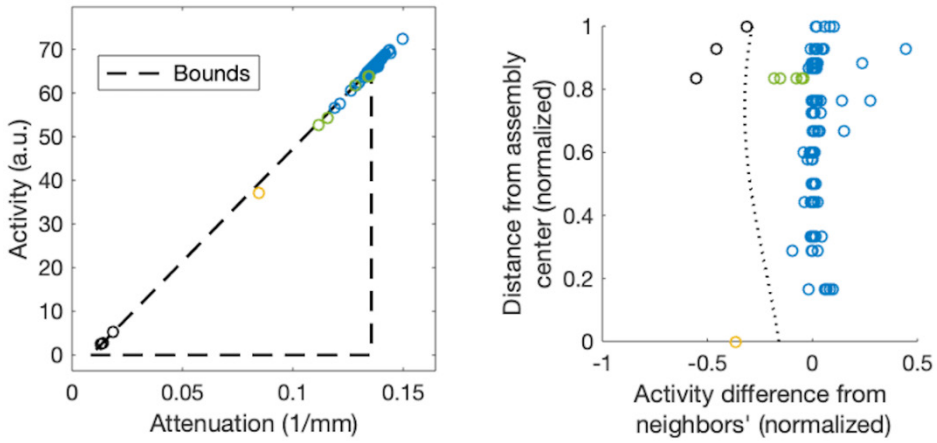


Figure 3: Rod classification plots for the VVER-440 assembly #1 (see also Fig. 2, top row). Linear bounds and average rod values are shown in the attenuation-activity plane on the left and rod activity difference from the neighbors as a function of the distance from the assembly center is shown on the right. Circles represent individual rods and colors denote the ground truth rod type (blue for present, yellow for water channel, black for missing and green for burnable absorber rod). The dotted line on the right represents the classification border.

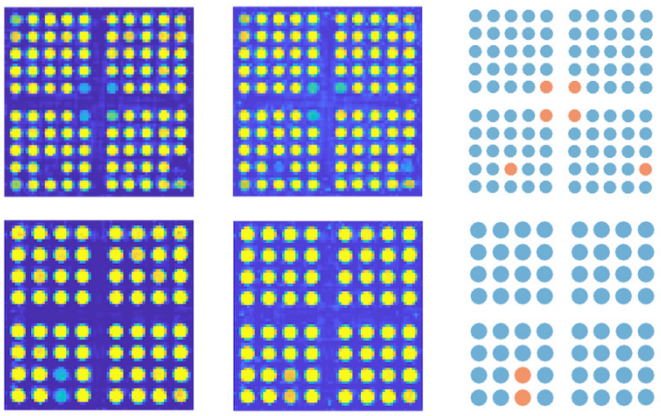


Figure 4: Activity (left column) and attenuation (middle column) reconstructions and classification into missing (orange) and present (blue) rods (right column). The top row shows the SVEA-96 assembly #7 with two missing rods and a water channel (four center-most rods) and the bottom row shows the SVEA-64 assembly #5 with two rods that have been in the reactor for only 2 out of the normal 4 fuel cycles.

An ATRIUM10 assembly has a 3x3 water channel and eight partial rods, which start at the bottom of the assembly and have a length of 2/3 compared to the rest of the rods. Fig. 5 shows attenuation and activity reconstructions and rod classification for the ATRIUM10 assembly #9 (BU 49.7 GWd/tU, CT 7.9 a) at both the normal measurement height as well as at the upper position 1.5 meters higher where the partial rods disappear from view. In the data collected at the higher position, the partial rod positions are classified as missing. This assembly also has two missing rods which are correctly classified at both measurement heights.

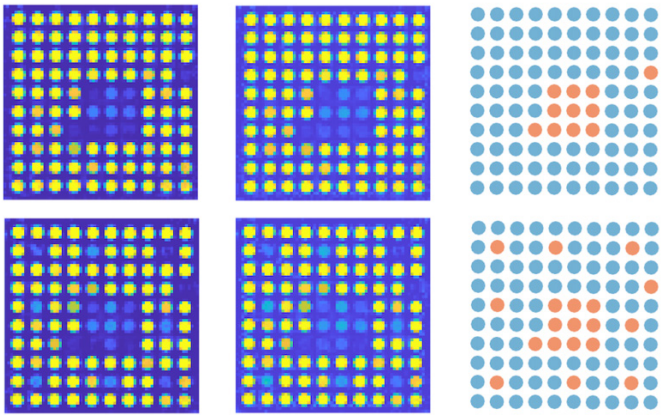


Figure 5: Activity (left column) and attenuation (middle column) reconstructions and classification into missing (orange) and present (blue) rods (right column) for the ATRIUM10 assembly #9 with two missing rods. The top row reconstructions are from data measured at normal height and the bottom row at the upper position, where the partial fuel rods disappear from view.

A total of eight assemblies with one to three missing rods were measured and in all cases the missing rods were accurately classified by the algorithm. Here we have shown results from five of these assemblies.

3.2 Burnable absorber detection

Some fuel assemblies contain fuel rods with added burnable absorber (usually Gd) used to balance the reactivity of the reactor during operation. Once the fuel assembly has reached the end of its operational lifetime, in an optimal case the burnable absorbers have reached the relative burnup of the other rods in the assembly. However, if the assembly has been removed before the absorber material has reached this level, the burnable absorber rods will show up as less active in the reconstructions.

To ensure accurate detection of missing fuel rods and to avoid false alarms, the burnable absorber rods should not be classified as missing by the classification algorithm. Our results show correct classification of burnable absorber rods as present, as demonstrated in Fig. 2 (upper row), where the VVER-440 assembly #1 with five burnable absorber rods is shown. Fig. 6 shows the VVER-440 assembly #3 (BU 43.0 GWd/tU, CT 2.7 a) and similarly, the burnable absorber rods are somewhat visible in the reconstructions but not classified as missing.

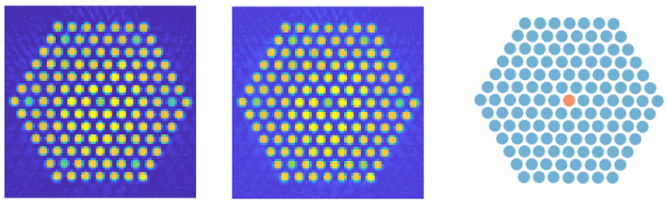


Figure 6: Activity (left) and attenuation (middle) reconstructions and classification into missing (orange) and present (blue) rods (right) for the VVER- 440 assembly #3 with a central water channel and 6 burnable absorber rods near the corners. The water channel is classified as missing but the burnable absorber rods are not.

A total of eight VVER-440 assemblies with burnable absorber rods were measured and all burnable absorber rods were accurately classified as present by the algorithm. Here we have shown results from two of these assemblies.

3.3 Water channel and partial rod detection

The accuracy of the algorithm in detecting missing fuel rods was also tested by its ability to correctly classify water channels and partial rod positions as missing. This is demonstrated with a variety of different assembly types. For VVER-440 type assemblies, the central water channel is accurately classified as missing in Fig. 2 and Fig. 6 (assemblies #1, #2 and #3). For a SVEA-96 assembly, the water channel classification can be seen in Fig. 4 (top row, assembly #7) and for an ATRIUM10 in Fig. 5 (assembly #9).

Fig. 7 shows the water channel classification for the 8x8-1 assembly #10 (BU 18.6 GWd/tU, CT 34.6 a, top row) and the 9x9-1AB assembly #11 (BU 35.0 GWd/tU, CT 20.7 a, bottom row). In both assemblies, the water channel near the center is correctly classified as missing. The former reconstruction also demonstrates the ability to gain accurate results from a long-cooled fuel assembly, which is very relevant in the context of a deep geological repository. Assemblies with even longer cooling times will be expected once the disposal starts.

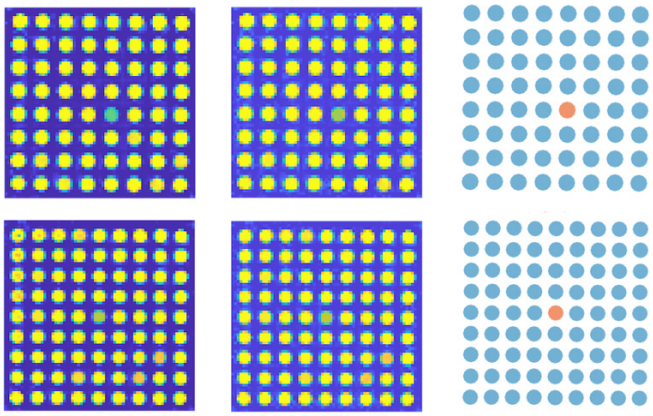


Figure 7: Activity (left column) and attenuation (middle column) reconstructions and a classification into missing (orange) and present (blue) rods (right column) for the 8x8-1 assembly #10 (top row) and the 9x9-1AB assembly #11 (bottom row). Both reconstructions show accurate classification of the water channel near the center.

A GE12 assembly contains two 2x2 water channels and 14 partial rods. Fig. 8 shows the attenuation and activity reconstructions and rod classification for the GE12 assembly #12 (BU 43.1 GWd/tU, CT 11.7 a) both at the normal measurement height and in the upper position, where the partial rods disappear from view. Water channels are visible and correctly classified as missing, as are the partial rods from the higher measurement.

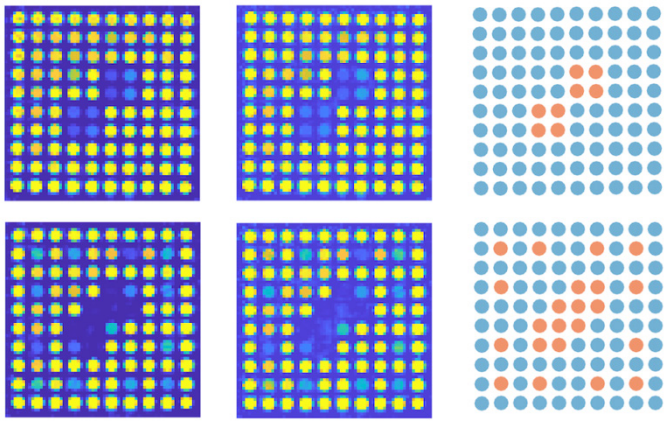


Figure 8: Activity (left column) and attenuation (middle column) reconstructions and rod classification into present (blue) and missing (orange) rods (right column) for the GE12 assembly #12. The top row reconstructions are from data measured at normal height and the bottom row at the upper position, where the 14 partial fuel rods disappear from view.

A SVEA-96 OPTIMA assembly contains 4 rod positions of water channel in the center and 8 partial rods next to them. Fig. 9 shows the SVEA-96 OPTIMA assembly #8 (BU 39.8 GWd/tU, CT 13.7 a) at the normal measurement

height as well as in the upper position where the partial rods disappear from view. The water channel positions in the center are correctly classified as missing as are the 8 partial rod positions in the upper position. Note that in the reconstructions from the upper position measurements there is one extra misclassified rod at the right lower corner and another in the lower-left part, see Section 3.4. for details.

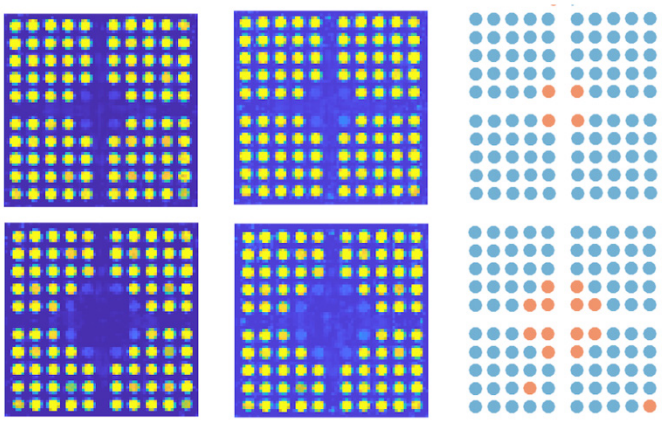


Figure 9: Activity (left column) and attenuation (middle column) reconstructions and rod classification into present (blue) and missing (orange) rods (right column) for the SVEA-96 OPTIMA assembly #8. The top row reconstructions are from data measured at normal height and the bottom row at the upper position, where the partial fuel rods disappear from view. Note the two misclassified rods in the upper position measurements, see Section 3.4.

3.4 Misclassified rods

Occasionally, some rods are falsely classified as missing. The classification algorithm concludes that a certain fuel rod is missing based on limits for activity deviation from the rod's neighbors. The limits are defined by support vector machines trained with labelled training data (see Section 2.3).

In our studies, several different types of erroneous classifications occur. The results strongly suggest that one of the types of misclassifications is related to nearby water channels or partial rods which make it harder to detect whether a rod is present or missing. Another type is related to simplifications in the assembly geometry made before image reconstruction. Sometimes overall poor data quality, for example low gamma counts in certain measurements, cause activity and attenuation variation in the reconstruction and can lead to misclassifications.

The percentages for missing rods (including known water channel positions) and present rods being classified both correctly and faultily are shown in Table 2 for the Olkiluoto and Loviisa measurements separately. The values are calculated by dividing the number of correctly or faultily classified rods by the total number of all rod positions of that type. The total misclassification rate is the fraction of the total number of rod positions that are misclassified.

For all 39 Olkiluoto assemblies (a total of 47 measurements as 8 assemblies were measured at two heights), the amount of misclassified missing rods is 3.92 %. The value for the 38 Loviisa measurements is significantly higher, 65.22 %, which is due to the fact that in most cases the central water channel is not correctly classified. This is further discussed in Section 4. The overall misclassification percentages for all rod types are 0.94 % for the Olkiluoto and 0.66 % for the Loviisa campaigns. In total for all measurements across both measurement locations, the percentage is 0.79 %.

Reference sample	Olkiluoto	Loviisa
Missing rods correctly classified	96.08 %	34.78 %
Missing rods faultily classified	3.92 %	65.22 %
Present rods correctly classified	99.29 %	99.96 %
Present rods faultily classified	0.71 %	0.04 %
Total misclassifications	0.94 %	0.66 %

Table 2: Percentage of rods in a certain category classified correctly or faultily in the two different measurement locations throughout all 85 measurements (47 at Olkiluoto and 38 at Loviisa).

An example of a faulty classification caused by the assembly geometry simplifications can be seen in Fig. 9, where the reconstruction of the SVEA-96 OPTIMA fuel assembly #8 is shown both at normal measurement height and in the upper position. The lower-right corner rod of the assembly is classified as missing in the case where the partial rods are not in view, although the rod is present. The reconstruction in the lower row is more uneven and thus prone to misclassifications. Fig. 10 shows rod classification plots for the assembly #8. In the same way as in Fig. 3, rod positions with similar characteristics get grouped together. The water channels show low activity and attenuation and thus resemble missing rods. In the measurements collected at the upper position, the partial rods behave like water positions and end up with the water channel rods, and at normal measurement height they behave like present rods as they should. The black circles in the rightmost figure of the bottom row are the misclassified rods in the assembly (see Fig. 9). As is clear, the margin to present rods is very small and a minor change in the classification border would result in a correct classification of the rods.

The assembly geometry simplification causing corner rods getting misclassified can be observed in other reconstructions in addition to the previously presented Fig. 9. The problem is limited to a certain assembly type (SVEA-types, especially SVEA-96) and is worst with the data gained from measurements at Olkiluoto 2017. The root cause for this is not yet completely confirmed, but it is related to the corner rods in these types of assemblies being smaller than the other rods and being placed closer to the center of the assembly in order to round the corners. This effect can be observed in the reconstructions.

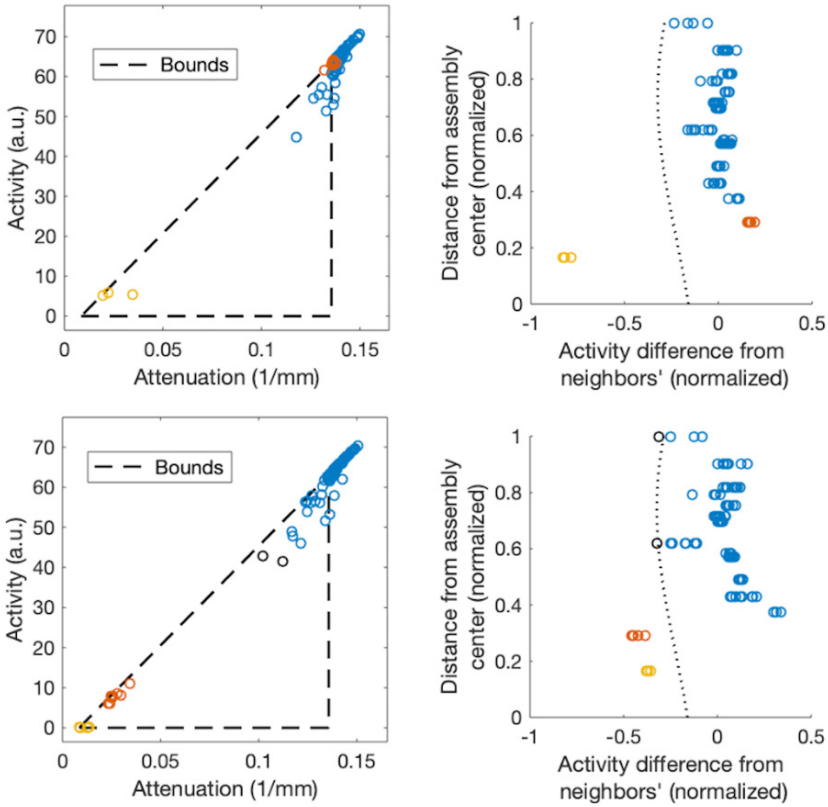


Figure 10: Rod classification plots for the SVEA-96 OPTIMA assembly #8 (see also Fig. 9), normal measurement height in the top row and upper position in the bottom row. Linear bounds and average rod values are shown in the attenuation-activity plane on the left and rod activity difference from its neighbors as a function of the distance from the assembly center is shown on the right. Circles represent individual rods and colors denote the ground truth rod type (blue for present, yellow for water channel, red for partial rod and black for rod that has been misclassified). The dotted line on the right represents the classification border.

3.5 Intra-rod activity intensity deviations

In some cases, especially with thicker rod diameters, the reconstructions show intra-rod activity differences. Fig. 11 shows the SVEA-64 assembly #6 (BU 32.9 GWd/tU, CT 20.7 a), where these differences can be clearly seen as darker spots inside the outermost rods of the activity image.

The intra-rod activity differences are limited to SVEA-64 assemblies and a few 9x9-1AB and 8x8-1 assemblies. These assemblies all have in common a larger pellet diameter (9.5-10.4 mm) than the other assemblies (7.5-8.9 mm). A very small amount of the same phenomenon can be seen at the outer edge of some rods in the activity reconstructions in Fig. 4 for the assembly #5 and in Fig. 7 for the assembly #11.

A similar phenomenon has been studied by Caruso et al. [13]. The intra-rod caesium and europium isotopic distributions were determined by gamma-activity tomography in high-burnup PWR fuel rods. Full-power irradiation at high temperatures causes fission products to diffuse from the hotter central region of the rod to the colder periphery, resulting in a non-uniform distribution of fission products inside the rod. Especially Cs-134 and Cs-137 diffuse easily and can cause significant differences in the fissile material content inside the rod. This phenomenon could explain our results, and the hypothesis is supported by the attenuation reconstruction, which does not show intra-rod differences. If the hypothesis holds, our results show promise in detecting even intra-rod distributions of nuclear material.

It should be noted that our chosen regularization terms (see Section 2.3.) are such that they prefer solutions where the intra-rod activity values and attenuation coefficients are uniform. This means that some of the intra-rod differences in the values are smoothed out in the reconstructions.

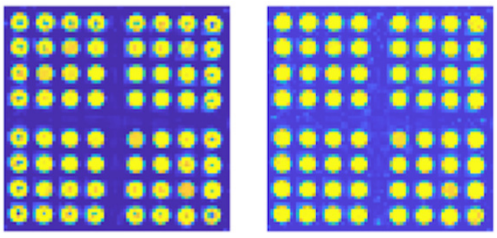


Figure 11: Activity (left) and attenuation (right) reconstructions for a SVEA-64 assembly #6, showing intra-rod activity intensity differences.

3.6 Activity and attenuation gradient

Some assemblies show clear activity and attenuation gradients in the reconstructions. The gradual change of activity values throughout the assembly is normal to certain types of assemblies and is caused by their irradiation

history and their placement in the reactor core during operation. The gradient on the attenuation coefficients, on the other hand, is an artefact of the reconstruction algorithm.

From the perspective of correct rod classification, the assemblies with a gradient demonstrate the ability of the algorithm to perform well even with more difficult assemblies. The gradient could cause some rods to be misclassified, because the algorithm evaluates rods by comparing their activity values to the values of their neighbors. Thus, significantly different values on opposite sides of the evaluated rod might cause the rod to be faultily classified. However, our results show that in most cases the smooth change in the activity throughout the fuel assembly does not cause misclassified rods.

An example of an assembly with a gradient can be seen Fig. 12, where the activity and attenuation reconstruction and rod classification of the VVER-440 assembly #4 (BU 22.8 GWd/tU, CT 27.6 a) is shown.

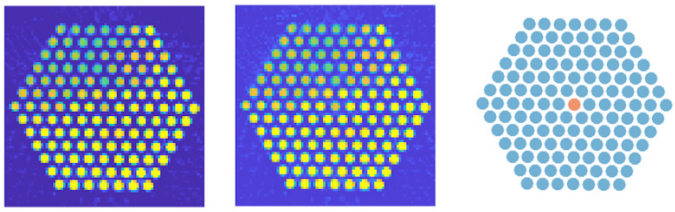


Figure 12: Activity (left) and attenuation (middle) reconstructions and classification into missing (orange) and present (blue) rods (right) for the VVER-440 assembly #4 with a water channel and a visible gradient in activity as well as attenuation. The water channel in the middle is correctly classified as missing.

3.7 Reconstruction robustness against geometry deviations

There is a small amount of prior information assumed about the measured fuel as described in Section 2.3.1. For safeguards purposes, the rod diameters and grid layouts are information that is available from determining the fuel assembly type from an initial FBP image which does not assume any prior information. The robustness of the method against small deviations from the truthful geometry is tested by intentionally changing the assumed rod pitches and diameters and by using a wrong grid layout.

The effect of small variations in the assumed rod pitch is demonstrated for assembly #10 in Fig. 13, where the activity reconstructions and rod classification plots for two different pitches are shown. The assumed rod pitch is intentionally set 1 mm smaller and 1 mm larger compared to the correct value. The reconstruction shows rod shapes distorted at the borders and highly emitting centers that are not in line with each other: a pincushion distortion in case of a smaller pitch and a barrel distortion in case of a larger pitch. With the larger pitch the water channel is misclassified as present.

Variations in assumed rod diameter cause the reconstructed rods to be smaller or larger. Overall this does not affect the quality of the reconstruction but can lead to misclassified rods. This effect needs to be further investigated to understand the underlying reason.

The use of a wrong assembly type as a prior for the reconstruction was also investigated. Results for the same 8x8-1 assembly #10 with both 9x9-1AB and ATRIUM10 type of priors are shown in Fig. 14. The activity reconstructions and rod classification plots for both false priors are shown. Despite the wrong number of rods in the prior, the reconstructions have the correct amount of rods visible, but the distortions in the reconstructions are very clear. The prior does not force the number of rods in the reconstruction to be the same as was assumed, neither do the assumed water channel locations guide the reconstruction result towards outcomes with water in those positions. Misclassifications will occur if the wrong geometry is chosen, but the distortions in the reconstruction allow for easy detection of wrong assumptions.

4. Discussion

The results presented in Section 3 show that the reconstruction method gives a reliable basis for the classification of fuel rods. However, our current method of classifying the rods in only two distinct categories (missing and present) is somewhat insufficient for the safeguard purposes at the disposal in a geological repository. There are more possibilities for nuclear material diversion than just removing a whole rod. The rod can be for example partially removed or replaced with another material. Related to this, also the burnable absorber rods should be somehow identified by the classification because they behave abnormally as well. The burnable absorber rods show that it is actually very difficult to detect rods that are replaced by a material with similar characteristics to spent fuel. Especially with low burnup it is almost impossible to detect these kinds of deviations.

In the case of burnable absorber rods, we know from the licence-holder's declaration that these rods might have a lower burnup and we can respond to the classification result accordingly. However, in other cases such an abnormal result might indicate a replacement of the rod with inactive material of the same density as the fuel or a diversion of a part of the fuel rod. Further investigation could then be done. Replacement scenarios and how to detect them are a topic of future research.

For the above-mentioned reasons there is a need to create additional classification categories to account for cases where the fuel rod might be modified or replaced with some other material. Division could be done into present,

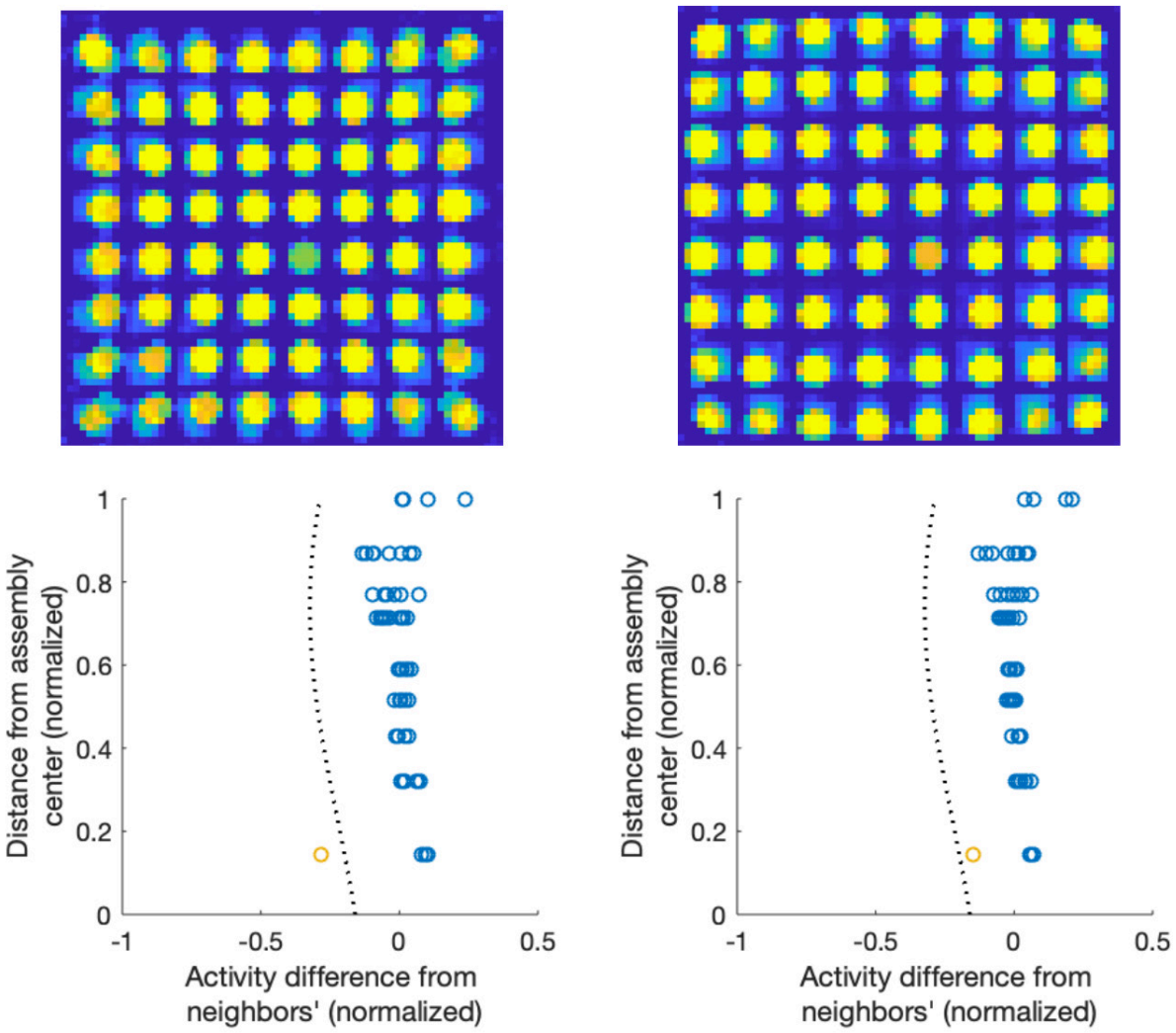


Figure 13: Activity reconstructions (top row) and rod activity differences from their neighbors as a function of the distance from the assembly center (bottom row) for the 8x8-1 type assembly #10 (see Fig. 7 for the reconstruction with correct pitch). Left column: 1 mm smaller rod pitch. Right column: 1 mm larger rod pitch.

abnormal and missing rods and the categories could be given a different priority depending on actions needed. Abnormal assemblies would require further investigations to get the "green light" to proceed with the disposal.

An example of the differing behaviour indicating something abnormal is given by assembly #5 in Fig. 4, where the abnormal fuel rods noticed by the algorithm are present but have a lower burnup than the other rods in the assembly. In the current version of the software, these rods are classified as missing due to their low activity value. Still, the presence of the rods in the assembly is clear from the attenuation image, although the attenuation coefficients also differ slightly from the other rods. In the future version of the classification algorithm, the rods could be classified as "modified" and then investigated in more detail. This would also reduce the possibility of false alarms.

There are some aspects of the software that need revision. The misclassifications caused by geometry assumptions, as presented in Section 3.4, will be addressed to ensure that no false alarms will be given due to simplifications in the geometry assumptions. The issue is demonstrated by the reconstructed activity and attenuation plots in Fig. 9. The corner rods of the assembly can be observed to be somewhat smaller than the other nearby rods. The placement of the corner rod is also a bit different from the other rods in the same row and column such that the overall shape of the assembly is a bit rounded. The reason behind this geometrical arrangement is to even out the reactivity and neutron fluxes in the reactor during operation. The geometry assumptions of this assembly type will be revised and implemented in the software, and this will likely bring down the number of misclassifications for the Olkiluoto campaigns.

The high number of misclassifications, especially in the Loviisa campaigns (see Section 3.4, Table 2), is due to the fact that the self-attenuation of nuclear fuel causes the contribution of the center of the assembly in the gamma data to be very small. This is a problem especially with the larger VVER-440 assemblies, and thus the

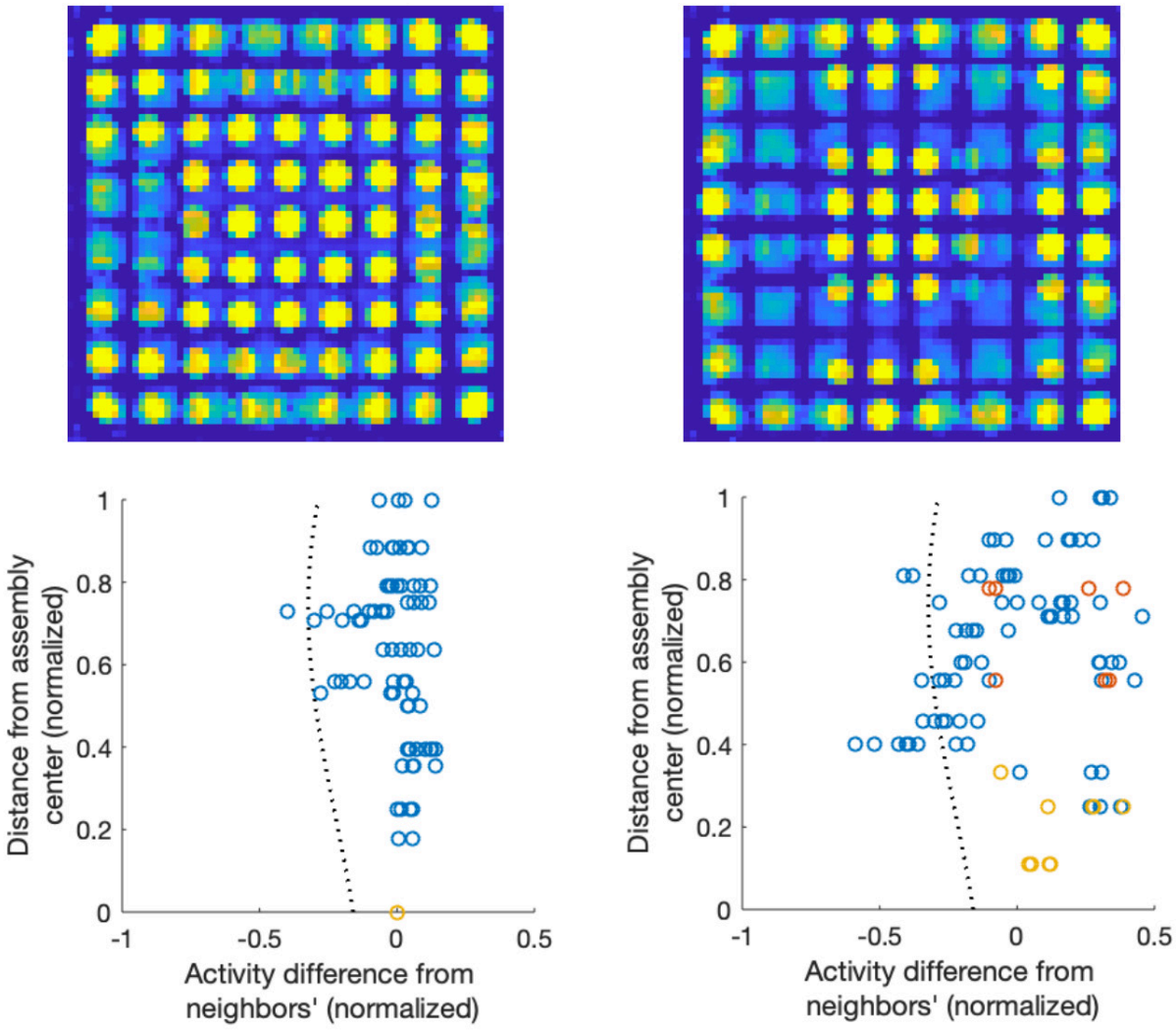


Figure 14: Activity reconstructions (top row) and rod activity differences from their neighbors as a function of the distance from the assembly center for the 8x8-1 #10 assembly with false geometry priors (see Fig. 7 for the reconstruction with correct geometry). Left column: 9x9-1AB geometry prior. Right column: ATRIUM10 geometry prior.

central water channel is often misclassified as present. In fact, excluding the misclassified central channels, the misclassification percentage of missing rods for the Loviisa campaigns is 0. The effect of self-attenuation on the classification of potential missing rods near the center of the assembly needs to be investigated further, especially for long-cooled assemblies. However, at present there are no data available from assemblies with such characteristics. Simulation studies or mockup assembly measurements might be necessary.

Based on the reconstruction robustness results presented in Section 3.7, small variations in the initial assumptions on the assembly geometry cause distortions that will be easily noticed visually. Thus, intentional deviations from the standard assembly geometry can be noticed and using prior information about the assembly type is acceptable from the safeguards point of view.

We see room for improvement in the determination of the predefined activity and attenuation bounds. The presently used triangular bounds, e.g., do not contain high-activity surrogate rods made from less attenuating material such as steel. However, opening up the bounds too much leads to poorer images; an optimum solution will thus need to be found. We are working on making the forward model more realistic by adding gamma ray scattering. Presently, only absorption is taken into account. In the context of incorporating PGET in the geological disposal safeguards activity, the processes of data acquisition, image reconstruction and rod classification will be integrated and automated.

5. Conclusions

The presented results show that the simultaneous reconstruction of activity and attenuation images works as

a reliable basis for fuel rod classification. The developed method is confirmed with data from a wide range of spent fuel assembly types and parameters measured at the Finnish nuclear power plants. Here, a selection of 12 different assemblies was shown, but data from 77 individual fuel assemblies were used in evaluating the total performance of the method.

For the 39 BWR assemblies measured at Olkiluoto (a total of 47 measurements as 8 assemblies were measured at two heights), an overall misclassification rate of 0.94% was achieved and the method shows high accuracy in detecting missing rods. The method is thus well suited for nuclear safeguards verification of BWR fuel assemblies in Finland prior to deep geological disposal.

For the 38 VVER-440 assemblies measured at Loviisa, some further work is still required to investigate the ability to detect missing rods near the center of the assembly. The central water channel is not often correctly classified as missing and might indicate that nearby missing rods would also go undetected, but data is needed to verify this.

We are working on improving the reconstruction method and the classification algorithm. The classification criteria will be revised to include a further category for rods that might be modified but not missing. Other future work includes improving the activity-attenuation bound estimation method and including scattering in the forward model to describe the physical phenomena inside the fuel assembly more realistically.

6. Acknowledgements

TAB acknowledges support by the Academy of Finland postdoctoral grant, decision number 330522.

TAB and SS acknowledge partial support by the Academy of Finland through the Finnish Centre of Excellence in Inverse Modelling and Imaging 2018-2025, decision number 312339.

PD acknowledges partial support by Business Finland under Grant 1845/31/2014.

TH was supported by the Academy of Finland via the decision number 326961.

7. References

[1] S. J. Tobin, P. Peura, C. Bélanger-Champagne, M. Moring, P. Dendooven, and T. Honkamaa; *Utility of Including Passive Neutron Albedo Reactivity in an Integrated NDA System for Encapsulation Safeguards*; ESARDA Bulletin; vol. 56; pp. 12–18; 2018.

[2] ASTOR Group report 2011-2016; *Technologies Potentially Useful for Safeguarding Geological Repositories*; STR-384; Vienna, July 2017.

[3] S. Vaccaro, I. C. Gauld, J. Hu, P. De Baere, J. Peterson, P. Schwalbach, A. Smejkal, A. Tomanin, A. Sjöland, S. Tobin, and D. Wiarda; *Advancing the Fork detector for quantitative spent nuclear fuel verification*; Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment; vol. 888; pp. 202–217; 2018.

[4] T. Honkamaa, F. Levai, A. Turunen, R. Berndt, S. Vaccaro, and P. Schwalbach; *A prototype for passive gamma emission tomography*; IAEA Safeguards Symposium; no. August 2015; 2014.

[5] M.Mayorov, T.White, A.Lebrun, J.Brutscher, J.Keubler, A.Birnbaum, V. Ivanov, T. Honkamaa, P. Peura, and J. Dahlberg; *Gamma Emission Tomography for the Inspection of Spent Nuclear Fuel*; 2017 IEEE Nuclear Science Symposium and Medical Imaging Conference; NSS/MIC 2017 - Conference Proceedings; pp. 2–3; 2018.

[6] T. White, M. Mayorov, A. Lebrun, P. Peura, T. Honkamaa, J. Dahlberg, J. Keubler, V. Ivanov, and A. Turunen; *Application of Passive Gamma Emission Tomography (PGET) for the Verification of Spent Nuclear Fuel*; in Proceedings of 59th Annual Meeting of the Institute of Nuclear Materials Management; Baltimore, MD, USA; 2018.

[7] C. Bélanger-Champagne, P. Peura, P. Eerola, T. Honkamaa, T. White, M. Mayorov, and P. Dendooven; *Effect of Gamma-Ray Energy on Image Quality in Passive Gamma Emission Tomography of Spent Nuclear Fuel*; IEEE Trans. Nucl. Sci.; vol. 66; no. 1; pp. 487–496; 2019.

[8] R. Backholm, T. A. Bubba, C. Bélanger-Champagne, T. Helin, P. Dendooven, and S. Siltanen; *Simultaneous reconstruction of emission and attenuation in passive gamma emission tomography of spent nuclear fuel*; Inverse Problems and Imaging; vol. 14; no. 2; pp. 317–337; 2020.

[9] H. Engl, M. Hanke, and A. Neubauer; *Regularization of inverse problems*; Springer Science & Business Media; 1996.

[10] B. Kaltenbacher, A. Neubauer, and O. Scherzer; *Iterative regularization methods for nonlinear ill-posed problems*; Radon Computational and Applied Mathematics; De Gruyter; 2008.

[11] C. T. Kelley; *Iterative Methods for Optimization*; Philadelphia, PA, USA: SIAM; 1999.

[12] N. Cristianini and J. Shawe-Taylor; *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*; Cambridge University Press; 2000.

[13] S. Caruso, M. F. Murphy, F. Jatuff, and R. Chawla; *Determination of within-rod caesium and europium isotopic distributions in high burnup fuel rods through computerised gamma-ray emission tomography*; Nuclear Engineering and Design; vol. 239; no. 7; pp. 1220–1228; 2009.

A methodology to identify partial defects in spent nuclear fuel using gamma spectroscopy data

Zsolt Elter, Sophie Grape

Uppsala University
Department of Physics and Astronomy, Division of Applied Nuclear Physics
Ångströmlaboratoriet, Lägerhyddsvägen 1
E-mail: zsolt.elter@physics.uu.se, sophie.grape@physics.uu.se

Abstract:

This paper describes a methodology to identify partial defects in modelled spent nuclear fuel using passive gamma spectroscopy data. A fuel library, developed with Serpent2, was used to calculate the material composition of spent nuclear fuel. Two fuel configurations were investigated in this work; one where the fuel assembly configuration was intact and one where 30% of the fuel rods were substituted with stainless steel rods in a random configuration. Emission and detection of gamma radiation from ^{134}Cs, ^{137}Cs and ^{154}Eu was simulated using a model of a passive gamma spectroscopy measurement station mimicking the Clab measurement station in Sweden. A simple HPGe detector model was implemented, and its detector efficiency was assessed using a range of different source energies. Realistic total gamma attenuation coefficients were calculated using the XCOM database. The modelled estimates of detected full-energy peak counts were then used in a Principal Component Analysis in order to investigate whether it was possible to distinguish between intact and partial defect fuel assemblies or not. The results showed that partial defects could be identified using the simultaneous analysis of all three peak intensities, and that the ability to do so increased when only gamma emission energies from ^{154}Eu were considered.

Keywords: safeguards; partial defect; PCA; nuclear fuel; multivariate analysis;

1. Nuclear safeguards and the verification of spent nuclear fuel

1.1 Verification of spent nuclear fuel using the existing safeguards framework

Under the Non-Proliferation Treaty [1], nuclear material needs to be safeguarded to ensure that it is not being diverted and used for non-peaceful applications. For this reason, spent nuclear fuel is regularly verified by nuclear safeguards inspectors. The inspectors are able to perform non-destructive assay (NDA) measurements on the fuel, in order to draw conclusions on the completeness and correctness of declarations. This is especially important before placing the fuel in so-called difficult-to-access storage where re-verification is not possible.

During the past decade, efforts have been ongoing to increase the capability to detect so-called partial defects in spent nuclear fuel, whereby a fraction of the nuclear material has been diverted or substituted. The efforts concern investigations of partial defect detection capability in mainly two different categories of nuclear safeguards instrumentation: i) instrumentation currently used in safeguards inspections, such as the Fork detector [2-4]) or the Digital Cherenkov Viewing Device (DCVD) [5]) and ii) instrumentation under development, or recently developed, for enhanced safeguards assessments such as partial defect detection [6,7]. In addition to these two categories, there are general detection techniques that could be investigated for enhanced safeguards performance and partial defect detection capability, such as passive gamma spectroscopy, which is the topic of this work. Earlier safeguards studies have shown the relevance of this technique [8,9].

The current partial defect detection level for spent nuclear fuel is on 50% of the fuel rods, but with the recent introduction of the passive gamma-emission tomography instrument denoted PGET, it is believed that verification of partial defect level on the single rod level will be possible [7]. However, it has also been shown [10] that if a sampling plan is to be developed for the verification procedure, it would be advantageous to also have additional instruments with partial defect detection capability at levels somewhere between the 50% level and the single pin level. High-resolution gamma spectroscopy is a measurement technique that can be envisaged for this purpose, since some facilities already have the equipment in place and is already used to verify operational parameters such as burnup (e.g. ASEA-ATOM facilities such as all the Swedish nuclear power plants and Clab [11]), whereas others could plan for such equipment when planning for or constructing new facilities (such as the planned encapsulation facility Clink in Sweden).

1.2 This work

It has been reported over the past years that there is a need for efficient and cost-effective safeguards verifications from the International Atomic Energy Agency (IAEA), and the use of machine learning tools and artificial intelligence in nuclear

safeguards is being investigated for different purposes [12-14]. Accordingly, investigating ways to make optimal use of data that has already been, or can be, collected using automatic machine learning tools is of high priority.

Against this background, this work presents a methodology that can be used to investigate whether or not partial defects can be identified in spent nuclear fuel assemblies, using a modelled passive gamma spectroscopy response and differences in gamma attenuation. The software that has been developed for geometric efficiency calculations is already publicly available [15] and the scripts that have been written for this specific work are available on GitHub [16]. The choices made by the authors in this work considering the measurement station, the gamma detector (type, geometry and design) as well as the spent nuclear fuel itself can be defined and changed by the user. For that reason, the results of the analysis here should merely be seen as an example of what information that can be obtained using this methodology, given the choices made by the user.

2. Spent nuclear fuel handling in Sweden

Sweden currently has six nuclear power plants (NPPs) in operation. After discharge from the reactor, the fuel is cooled in a fuel pond at the reactor site for around 1 year before it is shipped to Clab. SKB, the company that owns Clab, handed in an application to construct and operate a final repository for spent nuclear fuel in 2011. So-far, the Swedish government has still not taken a stand on the issue. According to the plans, the fuel to be encapsulated will cover a variety of fuel types (where BWR and PWR fuels are by far the most common types), fuel designs and fuel parameters [17]. The fuel assemblies are expected to have cooling times of up to around 70 years, while burnups are expected to reach up to around 60 MWd/kgHM. Before encapsulation, which will take place in the future Clink facility, the fuel assemblies will reside in water ponds in Clab where they are stored underground, in baskets holding up to 25 fuel assemblies at a time (25 BWR fuel assemblies or 16 PWR fuel assemblies).

Fuel handling equipment exists both at the NPPs and at Clab, but there are essential differences. At the reactor sites, the equipment is constructed to be able to handle manipulation of complete fuel assemblies for loading/unloading into the reactor, but also individual fuel rods in case of fuel damage when a single fuel rod needs to be removed/replaced. Clab has no equipment to handle individual fuel rods at all, and can only handle fuel assemblies in the reception area. At Clab, fuel assemblies are placed in baskets holding multiple fuel assemblies, before they are transported to the underground pools. The fuel handling machine in the underground pool area can only handle baskets. Partial defect verification is primarily of interest before transporting spent nuclear fuel to difficult-to-access storage, which means that such verification will most likely be done in the Clink facility, and neither at the NPPs nor in Clab. What fuel handling equipment or fuel assay instrumentation that will be available in Clink is not yet determined but it seems probable that only fuel elements (and not single fuel rods) will be handled. Based on this information, it appears that the only facilities currently equipped to pull and replace fuel rods are the NPPs, whereas verification of such defects is mainly planned take place in connection to verification before placement in difficult-to-access storage sites, long after such defects may be caused i.e. on long-cooled spent nuclear fuel, unless there are specific reasons to require this type of verification at an earlier stage such as a lost Continuity of Knowledge.

Hence, analysing remaining fission products with relatively long half-lives to possibly identify partial defects make sense, or one should recommend that partial defect verification is performed much earlier.

3. Methodology to calculate and analyze the number of counts in full-energy peaks

This section aims at describing the proposed methodology in different steps.

3.1 Overview of the methodology

A measure of the number of counts in the full-energy peak at a certain energy line in a gamma spectrum can be calculated by

$$f(E) = I_\gamma(E)\cdot\varepsilon_g(E)\cdot\varepsilon_d(E) = I_\gamma(E)\cdot\varepsilon_p(E) \quad (1)$$

where $I_\gamma(E)$ is the emission frequency of a certain energy line, $\varepsilon_g(E)$ is the energy-dependent geometric efficiency which describes the probability of a particle arriving at the detector from the source, and $\varepsilon_d(E)$ is the energy-dependent intrinsic detector efficiency which describes the probability of a particle leaving its full energy inside the detector. The multiplication of the last two efficiency functions can be defined as the full-energy peak efficiency $\varepsilon_p(E)$ of the setup. In gamma spectroscopy measurements of spent nuclear fuel, the emission frequencies $I_\gamma(E)$ depend on the nuclide inventory of the fuel assembly, and thus on its operational history and the amount of radioactive material in the assembly. Thus $I_\gamma(E)$ will be affected by a partial defect, but cannot in itself provide conclusive evidence of a defect. The geometric efficiency $\varepsilon_g(E)$ depends on the experimental setup (i.e. how far away the detector is placed from the source and what kind of absorbing material that is placed between them), the geometry of the fuel assembly, and as it will be shown later, the operational history through the change of the total gamma attenuation coefficient of the fuel. Finally, the detector efficiency $\varepsilon_d(E)$ depends on the geometry of the detector. Therefore, in a given setup (where the location of

the detector and the assembly is fixed), the change in the geometric efficiency should provide evidence of the defect.

In order to calculate the number of counts in different full-energy peaks for a given burnup (BU), cooling time (CT), initial enrichment (IE), the three functions in Eq. (1) have been tackled separately. An overview of the proposed methodology is shown in Figure 1, and the different steps are described in detail in the following subsections. A supplementary jupyter notebook and related python module can be found at [16] to demonstrate how the steps can be performed in practice.

In this work, a partial defect level of 30% has been considered because it constitutes an intermediate level to the detection capabilities of other safeguards instruments available today. 80 fuel rods were thus substituted against stainless steel rods in one random configuration, shown in Step 2 of Figure 1, where the pink fuel rods mark the steel dummy rods. The remaining low-enriched uranium fuel rods have the same material composition as those in the corresponding intact fuel assembly. The selected partial defect level and substitution material can be chosen differently, but the investigation of other configurations is outside the scope of this work and will be targeted in future research.

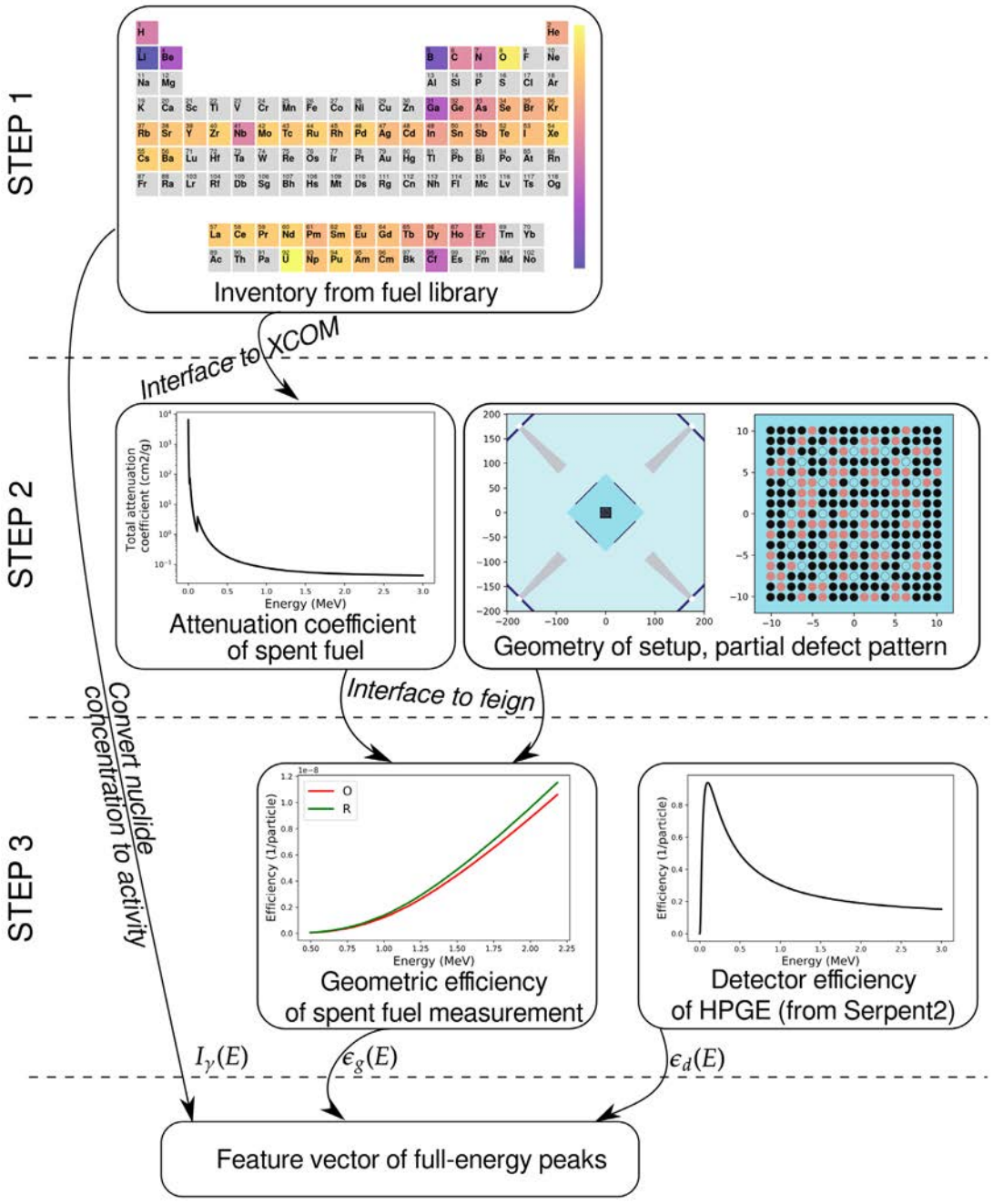


Figure 1. Summary of the proposed methodology

3.2 Step 1 - Determining the spent fuel nuclide inventory

The spent fuel nuclide inventories were sampled from an already existing PWR spent fuel library. The samples in the library were created by performing depletion calculations in a pin cell model with the Serpent2 code [18]. Further details on the fuel library and how it was created can be found in [19]. The fuel library itself is available at [20].

For the current study, 200 spent UO_2 fuel samples with a BU of 20-70 MWd/kgU and CT less than 30 years were randomly selected. Their respective concentrations of various isotopic concentrations were then extracted. As ^{134}Cs, ^{137}Cs and ^{154}Eu contribute significantly to the gamma spectrum at the cooling times considered in this work, only the gamma lines summarized in Table 1 were included. However, the supplemented python code can be extended with additional nuclides and associated energy lines without difficulty. The nuclide concentrations C_n were converted into gamma-line emission activities $A_{n,l}$ according to

$$A_{n,l} = C_n \cdot \frac{ln(2)}{T_{1/2,n}} I_{n,l} \tag{2}$$

where the half-lives of the nuclides $T_{1/2,n}$ and the intensities of the energy lines $I_{n,l}$ are given in Table 1. One has to note here that in order to obtain the actual number of counts in a detector, these per-volume quantities need to be multiplied by the spent nuclear fuel volume contributing to the detector signal. However, as detailed later, in this work only the values per fuel volume were used (no absolute values).

Nuclide	Half-life (y)	γ-lines (MeV)	Intensities (%)
^{134}Cs	2.065	0.563, 0.569, 0.604, 0.795, 0.801, 1.038, 1.167, 1.365	8.338, 15.373, 97.62, 85.46, 8.688, 0.990, 1.790, 3.017
^{137}Cs	30.1	0.662	85.1
^{154}Eu	8.6	0.723, 0.756, 0.873, 0.996, 1.004, 1.246, 1.274, 1.494, 1.596	20.06, 4.52, 12.08, 10.48, 18.01, 0.856, 34.8, 0.698, 1.797

Table 1: The gamma-ray emitting nuclides and their gamma-lines considered in this study.

3.3 Step 2 - Obtaining the accurate total gamma attenuation coefficients

Fresh light water reactor nuclear fuel consists of a mixture of uranium and oxygen, whereas spent nuclear fuel also contains a plethora of lighter fission products and heavier actinides. The difference in material composition as a function of BU impacts the attenuation of gamma rays, and was recently shown to be non-negligible in gamma spectrometry applications [21]. In order to accurately take into account and assess the impact of the change in the attenuating properties of spent nuclear fuel, we have created an interface between the nuclide inventory data and the XCOM software which calculates total attenuation cross-sections of materials [22]. Step 2 of Figure 1 includes an example of the total attenuation coefficient calculated for a spent fuel inventory.

3.4 Step 3 - Determining the full-energy peak efficiency

As pointed out previously, the full energy peak efficiency is determined by considering the geometric efficiency of the measurement setup and the detector efficiency separately.

3.4.1 Geometric efficiency of the modelled setup

In this work, the actual dimensions of the experimental setup are not important per se, but used as an example to show how the methodology can be used for any arbitrary measurement setup. Nevertheless, the dimensions of the passive gamma spectroscopy station at Clab were adopted. At Clab, the spent fuel is placed in a fixture mounted on the pool wall with a ~50 cm distance between the center of the assembly and the pool wall. The fixture is able to rotate the fuel assembly and also move it in axial direction. In the 2m thick concrete pool wall, there is an air-filled hole, partially filled with a steel collimator. The steel collimator is made of two massive steel half cylinders, covered with a steel window. The height of the collimator slit can be adjusted in the range of 1-3 mm to allow for different counting rates in the detector. The length of the steel collimator is 1.2 m, and the distance from the center of the fuel assembly to the end of the collimator is 2.46 m. In the model, between the end of the collimator and the detector, 4 absorber sheets are placed (8 mm of lead, 3 mm of aluminium, 21 mm of steel and 1 mm of copper) in order to filter out low-energy gamma rays. The setup is described in greater detail in [23, 24].

The geometric efficiency of this measurement station was computed with the feign package [15] which implements a 2D ray-tracing method without build-up factors. The package allows the user to define a rectangular fuel assembly, a pool around it, various absorbers and detector points with their associated collimators. A radial view of the geometry of the setup is shown in Step 2 of Figure 1. The geometric efficiency of the setup has been averaged over four detector locations, with each location facing one of the fuel assembly corners, thereby mimicking the rotation of the fuel assembly around the vertical axis in real-life measurements. This averaging plays a role in case the fuel assembly is asymmetric.

In the current work, the geometric efficiency was evaluated for the two cases of an intact 17x17 PWR assembly (noted as “O” for original), and a 30% partial defect 17x17 fuel assembly (later noted as “R” for random). The difference in geometric efficiency curves (as shown in Step 3 in Figure 1) demonstrates how manipulation of the fuel assembly may be identified.

3.4.2 Detector efficiency

A simple HPGe geometry with a crystal diameter and length of 60.5 mm and 61 mm was simulated. The core hole diameter and depth were 12 and 51 mm, respectively. Serpent2 simulations were made with a pen-beam source placed 1 cm above the central axis of the detector crystal for several source energies. A damped exponential function was fitted to the detector response to describe the detector efficiency curve, as shown in Step 3 of Figure 1. For fitting the function

$$ln\left(\epsilon_d\right) = a + b \cdot ln\left(\frac{E}{f}\right) + c \cdot ln^2\left(\frac{E}{f}\right) + d \cdot ln^3\left(\frac{E}{f}\right) + e \cdot ln^4\left(\frac{E}{f}\right) \quad (3)$$

was used, as proposed also in [25]. The fitting and the values of the parameters are available in the supplemented notebooks.

3.5 Creating the feature matrix and PCA

The previous sections summarized how each of the functions of Eq. (1) can be estimated. The following step is to multiply the functions for each considered gamma-ray energy emitted by the given spent fuel. Finally, each fuel sample can then be represented by a multi-dimensional feature vector, with each feature vector corresponding to the full-energy peak counts for all isotopes as described by Eq. (4). The sum of the feature vector was normalized to 1 as shown in Eq. (5), in order to give all features f_i (despite possible differences in magnitude) the same importance in the multivariate analysis (in the current case typically the 137Cs peak would have such an impact).

$$\underline{f} = \begin{pmatrix} f\left(E_{^{137}Cs_{0.662}}\right), f\left(E_{^{134}Cs_{0.563}}\right), \ldots, \\ f\left(E_{^{154}Eu_{0.723}}\right), \ldots, f\left(E_{^{154}Eu_{1.596}}\right) \end{pmatrix} \quad (4)$$

$$\underline{f}_{norm} = \frac{\underline{f}}{\sum_i f_i} \quad (5)$$

Since 200 nuclide inventories were used and the geometric efficiency was calculated for both intact and manipulated fuel, the analysis includes 400 feature vectors, which can be arranged into a feature matrix. In the feature matrix, each row corresponds to the feature vector of a specific spent nuclear fuel sample, and each column corresponds to the full-energy peak counts for a certain gamma-ray energy. The matrix underwent standard scaling meaning that for each column the mean is centered to 0 and the variance is renormalized to 1. The feature matrix is often referred to as a predictor matrix, since it is used to predict targets or responses which can be assigned to each sample. Such a response could be for example the reactor type for classification problems, or the fuel parameters BU and CT for regression problems. In this work, the response was whether or not the fuel assembly was intact. Through a simple encoding, the label “0” was assigned to samples of intact fuel, and the label “1” to all other samples, and all responses were collected in the so-called response vector.

A feature ranking was performed on the data matrix to investigate the importance of the different gamma-ray energies in the classification of the fuels, based on the so-called Pearson’s correlation score between the energies and the responses. The correlation score defined as

$$corr\left(x_i, y\right) = \frac{E\left(x_i \cdot y\right) - E\left(x_i\right) \cdot E\left(y\right)}{\sigma_{x_i} \cdot \sigma_y} \cdot 100(\%) \quad (6)$$

was used to rank the features according to their importance. Here x_i can be a column of the feature matrix, or a vector derived from several columns (eg. the ratio of two peaks).

Since the data is multidimensional (18 full-energy peak counts), it is difficult to visualize. A widely used dimensionality reduction technique called Principal Component Analysis (PCA) [26,27] was applied to investigate patterns within the data set, and analyze whether intact cases could be distinguished as such. PCA uses an orthogonal transformation to convert the original set of possibly correlated features (eg. multiple full-energy peak counts from the same isotope are necessarily strongly correlated) into a set of linearly uncorrelated variables. These variables are called Principal Components (PCs). The number of PCs is less than or equal to the original dimension of the feature vector. In the current study, the Scikit-learn python library [28] was used to perform the PCA.

3.6 Approximations and assumptions made in the analysis

In the analysis, many simplifying assumptions have been made:

- There has been no attempt to estimate an absolute number of counts in any of the selected full energy peaks, only an estimate of the number of counts per source volume and unit time. Due to this, certain effects were neglected such as an estimate of the fraction of the fuel assembly seen by the detector. This is in turn related to the exact dimensions of the collimator slit, which is 8.5 cm wide and variable in height. In an estimation of the absolute number of counts in a detector, a surface source with an angular distribution is a better approximation than the pen-beam source model used in the detector efficiency simulations here.

- No measurement noise is included in the analysis, but it is also not considered to be of major importance to this work since there are no attempts made to estimate an absolute number of counts. However, various (constant) background levels could impact the capability to correctly determine peak areas to use in the analysis, and hence such effects should be looked closer upon in future work, as absolute count rates are also taken into account.
- Since the absolute number of counts is not of interest, there has been no need to look into aspects concerning calibration issues of the detector setup.
- The electronic equipment needed for the gamma spectroscopy measurements has not been considered or modelled.
- There has been no attempt to actually train a classifier algorithm nor to estimate uncertainties in the results from considering different partial defect patterns, substitution materials or uncertainties in the operator-declared values IE, BU or CT. Also, only a standard irradiation cycle has been considered in the depletion calculations.
- Although the change in the attenuation coefficient of the spent fuel due to BU was taken into account, the change in the density and rod radius due to swelling was not considered. Also all rods within the same spent fuel assembly were considered to have identical burnup and inventory.

One could imagine to consider multiple approximations using some constant shape factor. If doing so, the constant would disappear in the normalization of the feature matrix, and thus in a noise-free (i.e. infinitely long) measurement it would have no impact.

4. Results and discussions

4.1 Feature ranking

The correlation between the features and the response vector were evaluated according to Eq. (6) and the correlation scores are given in Table 2. The ^{154}Eu lines have the highest correlation score to the identification of partial defects in the fuel, and the higher the gamma-ray energy is the higher the correlation score is. One can see that the correlation scores are lower than 10% for each feature, which indicates that peak counts from one single gamma-ray energy does not carry much information on whether the fuel is intact or not. The low correlation scores are expected, since replacing fuel rods will impact the probability of low and high energy photons reaching the detector differently. For example, low-energy photons reaching the detector originate most probably from the peripheral rods, since the central rods are shielded. Thus, replacing the peripheral fuel rods will have a larger impact on the low-energy range of the geometric efficiency than on the high-energy range. Accordingly, including multiple peaks in the analysis should correlate more to the response vector.

Nuclide, γ-line (MeV)	Corr. Score (%)	Nuclide, γ-line (MeV)	Corr. Score (%)
^{154}Eu, 1.596	9.55	^{134}Cs, 0.563	1.44
^{154}Eu, 1.494	9.21	^{134}Cs, 0.569	1.39
^{154}Eu, 1.274	7.95	^{134}Cs, 1.038	1.37
^{154}Eu, 1.246	7.73	^{134}Cs, 0.604	1.10
^{154}Eu, 1.004	4.99	^{137}Cs, 0.662	0.85
^{154}Eu, 0.996	4.88	^{154}Eu, 0.756	0.62
^{154}Eu, 0.873	2.92	^{134}Cs, 0.801	0.24
^{134}Cs, 1.365	2.28	^{134}Cs, 0.795	0.21
^{134}Cs, 1.167	1.81	^{154}Eu, 0.723	0.12

Table 2: Correlation scores of normalized peaks to class

One way to include multiple gamma-ray energies in the feature ranking is to calculate peak ratios for all possible combinations of the 18 gamma-ray energies. This was done here and the correlation scores were then re-calculated for the ratios and the response vector. In this case, the feature matrix was not standard scaled, since the scaling would obscure any correlations. The number of correlation scores are more than a hundred, and not easily shown in a table, but it is worth pointing out that the correlation scores dramatically increased. The highest correlation score, 99.34%, was obtained for the ratio of the 1.494 MeV and 1.246 MeV peaks of ^{154}Eu. In general, the ratios involving the different ^{154}Eu and ^{134}Cs gamma-ray energies resulted in very high correlation scores (over 90%), whereas correlation scores in which ^{137}Cs was included were below 7%. This shows that ^{137}Cs, having only a single energy line, is not helpful in classifying partial defects based on the changes in the geometric efficiency and that indeed a multivariate approach is needed when verifying partial defects with passive gamma spectroscopy.

4.2 Principal Component Analysis

Using the feature matrix described earlier, a Principal Component Analysis (PCA) was performed. Figure 2 illustrates the first three Principal Components (PCs), i.e. the three PCs that account for the largest variance in the data when all the 18 gamma lines from Table 1 are included in the matrix. One can observe that the two cases, “O” and “R”, are well-separated in the three-dimensional space spanned by the three first PCs. However, it can be also noticed that only the third PC bears any information on the identification of partial defects. Further investigations show that the variability in the first two PCs is caused by the so-called nuisance parameters BU and CT i.e. parameters which are not of direct interest but that must be taken into account in the analysis. Figure 2, displaying the two parabola-shaped clusters of data points, illustrates that the first two PCs are dominated by the CT-dependence of the fuel samples. The longer the cooling time is, the less pronounced the separation between the two cases is.

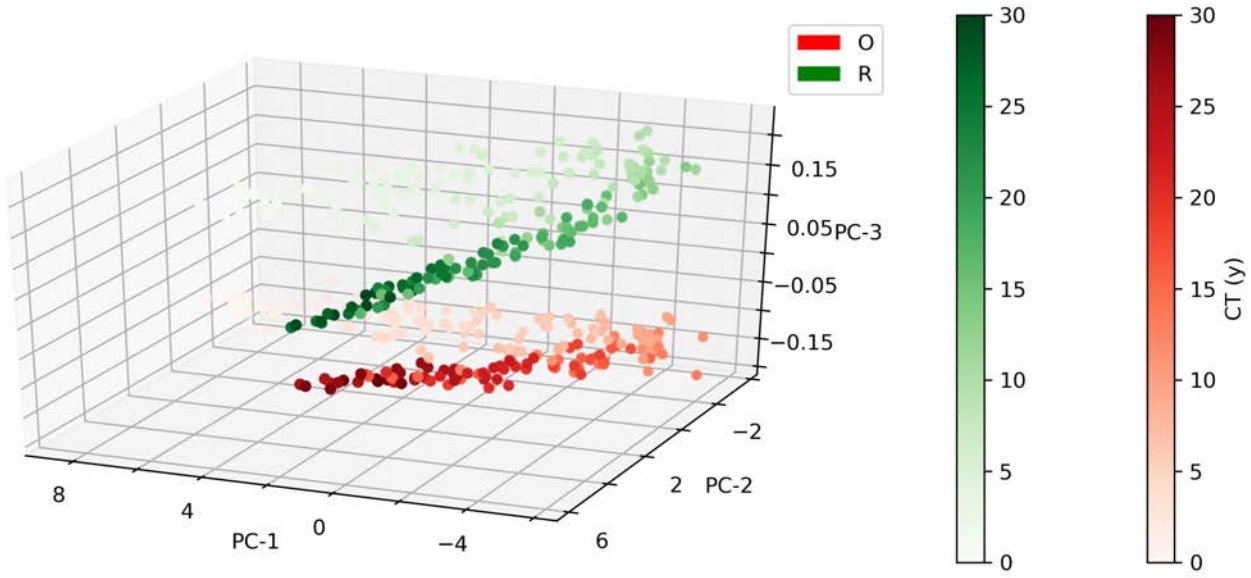


Figure 2: Principal Components of the dataset with all ^{134}Cs,^{137} Cs and ^{154}Eu peaks included.

There are several options available to suppress the importance of BU and CT in the data. Two options are to correct for the CT if that information is available (in practice often that is a value to be verified), or to use ratios of counts in different full-energy peaks. Here, an even more straightforward approach was used: the feature ranking implied that the ^{154}Eu lines carry the most relevant information when classifying fuel samples and thus only information on those gamma-ray energies were kept in the feature matrix. Since the ^{154}Eu concentration in the spent nuclear fuel is the same no matter what ^{154}Eu gamma-ray energy that is analyzed, one can expect that the impact of BU and CT becomes suppressed in the PCs. This can in fact be seen in Figure 3, which also shows the impact of using the actual total gamma attenuation coefficient. The left panel of Figure 3 shows the first two PCs for the case when the total gamma attenuation coefficient from XCOM was taken into account for each fuel sample. One can observe that now already the first PC provides information on the presence of partial defects, and that the spread in mainly the direction of PC-2 is due to the variation of BU (and hence attenuation) among the fuel samples. In the right panel of Figure 3, the first two PCs are illustrated for the case when the total gamma attenuation coefficient corresponding to fresh fuel was used for each sample. It is seen that this simplifying assumption leads to the removal of any BU and CT dependencies in the data, and that all the fuel samples end up at the exact same place on the PC coordinate system. Apparently, the impact of nuisance parameters has been successfully eliminated. It is also clearly seen that it is sufficient to only consider the first PC in order to correctly identify partial defects in the spent nuclear fuel assemblies in this case. However, the difference between the right and left panels show that the impact of the change in the attenuation coefficient is not negligible. Thus, in a practical case when the presence of partial defect (on any level) is to be predicted based on a measured gamma spectrum, the predictor model needs to be trained with data which accounts for that.

5. Conclusions

The current paper described a fast, robust and flexible methodology to estimate and analyze full-energy peak counts in passive gamma spectroscopy measurements of spent nuclear fuel. The methodology evaluates the modelled counts-per-volume and time in a simple HPGe detector, by separately taking into account the activity of different gamma-emitting nuclides in the spent fuel, estimating the geometric efficiency of the measurement setup and simulating the intrinsic detector efficiency of an HPGe detector. The methodology is flexible in the manner that the user may define a different measurement setup or detector design to study.

An application of this methodology was shown here. The purpose was to investigate the possibility of identifying partial defects in spent nuclear fuel. The modelled fuel was intact 17x17 PWR fuel assemblies and manipulated 17x17 PWR fuel assemblies, suffering from a 30% partial defect level. It was shown that the geometric efficiency did in fact depend on the presence of the partial defect, due to the different attenuation of low and high-energy gamma-rays by the fuel assembly itself. For 200 spent nuclear fuel samples, the per-volume and unit time counts in the full-energy gamma peaks of ^{134}Cs, ^{137}Cs and ^{154}Eu were estimated

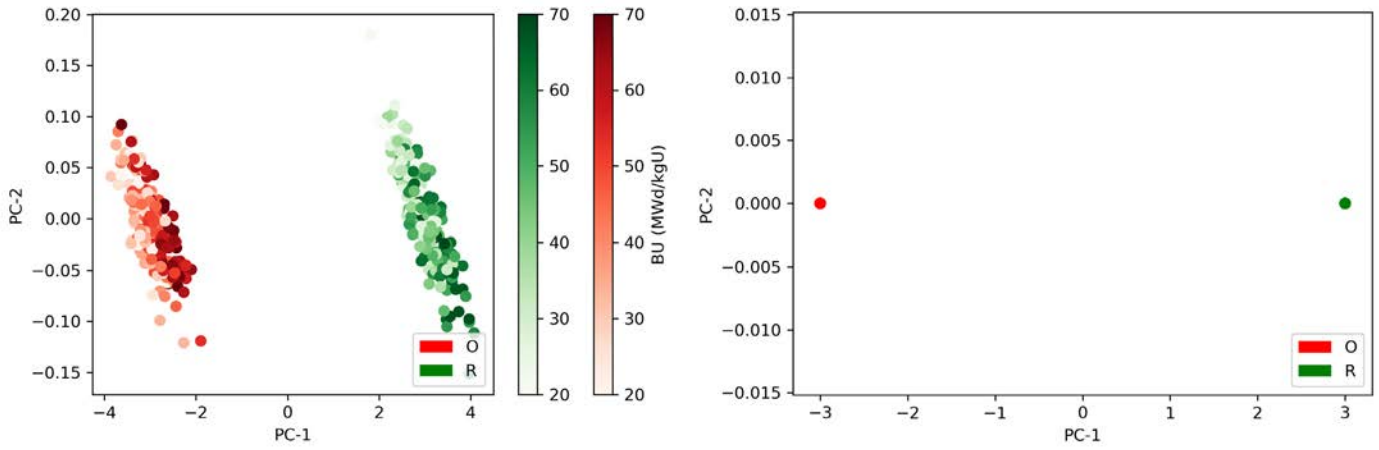


Figure 3: Principal Components of the dataset including only ^{154}Eu peaks. Left: Inventory-dependent attenuation coefficients from the XCOM database are used. Right: The attenuation coefficient of fresh UO_2 fuel is used for all samples.

and analysed with PCA. It was shown that just by using the peak counts-per-volume and time without any calibration, one can distinguish intact and manipulated fuel, however the BU and CT parameters act as nuisance parameters and need to be accounted for. As a solution found here was to only include the ^{154}Eu lines in the analysis. It was also shown in the analysis that considering a more realistic total gamma attenuation coefficient rather than a simplified fixed value, impacts the per volume and time counts in the detector and thus also the classification to some extent.

The possible draw-back identified with the use of only Eu-154 is its relatively short half-life of 8.6 years. In the verification of spent nuclear fuel before encapsulation, it is likely that many fuel assemblies will have a very long CT and thus there will be little Eu-154 left to detect. A solution to this could be to perform the partial defect verification earlier, for instance in connection to receiving the fuel at Clab, or even before shipping it to Clab from the NPP sites. The fact that the receiving facility does not have capability to manipulate spent nuclear fuel on the single rod level, in combination with Containment and Surveillance measures and other tools available to ensure that Continuity of Knowledge is kept, could possibly enable partial defects to be reliably detected.

6. Outlook

Since this paper mainly focused on the methodology and on giving a proof-of-principle, it was not intended to cover a large set of various partial defect scenarios. Nevertheless, we have begun the continuation of this work in which the methodology is used to investigate a large number of different partial defect patterns as well as different partial defect levels and different substitution materials.

Also, as the continuation of this work we would like to investigate how robust the methodology is to measurement noise. For this we will need to estimate the absolute value of the peak counts. In order to achieve that we are currently extending the software feign to handle 3D effects such as the impact of the fan-shaped view-angle of the collimator, and we are also going to improve the detector efficiency simulations to take into account a more realistic surface source.

Finally, it has to be noted that performing the PCA provided only an opportunity to visualize whether intact and manipulated fuel could be identified, although a proper classification method has not been developed. The next step will be to train classification methods, such as for instance artificial neural networks, and assess how well intact fuel can be discriminated from manipulated fuel in the presence of noise and to get an indication of the lowest level of partial defects that can be reliably detected with passive gamma spectroscopy.

7. Acknowledgements

We would like to acknowledge the Swedish Radiation Safety Authority for supporting this work under contracts SSM2017-5979 and SSM2020-996. We would also like to thank Carl Hellesen for the helpful discussions in the early stages of this work. Special thanks to Peter Andersson and Lorenzo Senis for the discussions regarding the simulation of the response function in a passive gamma measurement setup.

Finally, we would like to acknowledge that the coloured periodic table in Figure 1 was created with the python module created by Andrew S. Rosen [29].

8. References

[1] International Atomic Energy Agency (IAEA), *Treaty on the non-proliferation of nuclear weapons*. INFCIRC/140., (1970)

[2] Tiita, A., Saarinen, J., Tarvainen, M., Axell, K., *Investigation on the possibility to use fork detector for partial defect verification of spent lwr fuel assemblies: Final report on Task JNT A 1071 (BEL, FIN, SWE) of the Member States' Support Programme to IAEA Safeguards*. STUK report series, ISSN 0785-9325 ; STUK-YTO-TR 191 (2002)

[3] Meer, K. van der, Coeck, M., *Is the FORK detector a partial defect tester?* Symposium on international safeguards: Addressing verification challenges; Vienna (Austria); IAEA-CN--148/68 (2006)

[4] Rossa R., Borella A., Giani N., *Comparison of machine learning models for the detection of partial defects in spent nuclear fuel*. Annals of Nuclear Energy, Volume 147, (2020)

[5] Branger E., *Enhancing the performance of the Digital Cherenkov Viewing Device: Detecting partial defects in irradiated nuclear fuel assemblies using Cherenkov light*. Doctoral thesis, Uppsala University. ISSN 1651-6214 (2018)

[6] Tobin S.J., et al.. *Quantifying the plutonium mass in spent fuel assemblies with nondestructive assay – an update on the NGSI research effort*. In: Proceedings of the 52nd INMM annual meeting (2011)

[7] Jacobsson Svärd S. et al.; *Tomographic determination of spent fuel assembly pinwise burnup and cooling time for detection of anomalies*, 37th annual ESARDA Symposium

[8] Hellesen C., Grape S., Jansson P., Jacobsson Svärd S., Åberg Lindell M., Andersson P., *Nuclear Spent Fuel Parameter Determination using Multivariate Analysis of Fission Product Gamma Spectra*, Annals of Nuclear Energy, Volume 110, Pages 886-895 (2017)

[9] Elter Zs., Caldeira Balkeståhl L., Grape S., Hellesen C. Partial defect identification in PWR spent fuel using Passive Gamma Spectroscopy, Physor 2018 conference proceedings (2018)

[10] Hellesen C., Grape S., *Non-destructive assay sampling of nuclear fuel before encapsulation*, ESARDA Bulletin No 56, (2018)

[11] *Private communication*, Prof. Ane Håkansson, Uppsala University (2020)

[12] Rockwood L., Mayhew N., Lazarev A., Pfneisl M., *IAEA Safeguards: Staying Ahead of the Game*, Report by the Swedish Radiation Safety Authority, Report number: 2019:14 ISSN: 2000-0456 (2019)

[13] Shoman N., Cipiti B. B., *Unsupervised Machine Learning for Nuclear Safeguards*, Annual Meeting Proceedings of Institute of Nuclear Materials Management, July 22-26, 2018 in Baltimore, Maryland (2018)

[14] Haddal R., Hayden N., *Autonomous systems, artificial intelligence and safeguards*, IAEA Symposium on International Safeguards Building Future Safeguards Capabilities, (2018)

[15] Elter Zs., Cserkaszky A., Grape S., *feign: a Python package to estimate geometric efficiency in passive gamma spectroscopy measurements of nuclear fuel*, The Journal of Open Source Software 4(42):1650 (2019)

[16] Elter Zs. A methodology to identify partial defects in spent nuclear fuel using gamma spectroscopy data (supplementary software) (Version v1.0) Zenodo. http://doi.org/10.5281/zenodo.4250073 (2020)

[17] Svensk Kärnbränslehantering AB, *Spent nuclear fuel for disposal in the KBS-3 repository*, Technical Report TR-10-13 (2010)

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

[19] Elter Zs., Pöder Balkeståhl L., Branger E., Grape S., Pressurized water reactor spent nuclear fuel data library produced with the Serpent2 code, Data-in-Brief, Vol 33, DOI: https://doi.org/10.1016/j.dib.2020.106429 (2020)

[20] Elter, Zs., *Uppsala University Pressurized water reactor spent nuclear fuel data library*, Mendeley Data, V1, doi: 10.17632/8z3smmw63p.1 (2020)

[21] Atak H. et al., *The degradation of gamma-ray mass attenuation of UOX and MOX fuel with nuclear burnup*, Progress in Nuclear Energy, Volume 125, (2020)

[22] Berger M.J. et al., *XCOM: Photon Cross Sections Database*, NBSIR 87-3597, DOI: https://dx.doi.org/10.18434/T48G6X (2010)

[23] Vaccaro S. et al., PWR and BWR spent fuel assembly gamma spectra measurements, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, Volume 833, Pages 208-225 (2016)

[24] Willman, C, *Applications of Gamma Ray Spectroscopy of Spent Nuclear Fuel for Safeguards and Encapsulation.* Doctoral thesis, Uppsala University. ISSN 1651-6214 (2006)

[25] Mattera, A. *Studying neutron-induced fission at IGISOL-4: From neutron source to yield measurements and model comparisons.* Doctoral thesis, Uppsala University. ISSN 1651-6214 (2017)

[26] Hotelling, H. *Analysis of a complex of statistical variables into principal components.* Journal of Educational Psychology, 24, 417–441, and 498–520. (1933)

[27] Hotelling, H. *Relations between two sets of variates* Biometrika. 28 (3/4): 321–377. (1936)

[28] Pedregosa F. et al. *Scikit-learn: Machine Learning in Python*, Journal of Machine Learning Research 12, pp. 2825-2830, (2011)

[29] Rosen, A. S., *ptable_trends*, python module DOI: 10.5281/zenodo.1451880 (2018)

Three-dimensional Positional Analysis of Weapons Grade Plutonium Using Gridded Arrays of Dosimeters

Ryan P O'Mara[1] and Robert B Hayes[2]

[1] Pacific Northwest National Lab
902 Battelle Boulevard, Richland WA 99354
[2]North Carolina State University
Nuclear Engineering Department
2500 Stinson Dr. Raleigh, NC 27520
E-mail: ryan.omara@pnnl.gov

Abstract:

One of the primary challenges to preventing the covert proliferation of nuclear weapons is that no technology currently exists that can economically monitor nuclear activities continuously throughout time. Previous studies have shown that luminescence dosimetry can potentially be used to determine the energy of historical radiation environments and even source positions. This work serves to demonstrate that with an adequate number of radiation dose measurements, it may also be possible to image source material using luminescence dosimetry. To that end, a 4.5 kg sphere of weapons grade plutonium was used to expose two gridded arrays of commercial optically stimulated luminescence dosimeters (OSLDs). The measured position-wise dose was then used to reconstruct the 3-dimensional position of that source material. Expanding upon these findings it may be possible in the future to use luminescence dosimetry in ubiquitous, environmental materials to reconstruct the historical radiation fields in nearly any nuclear facility in the world. Developing such a capability could greatly increase the likelihood of detection of covert nuclear weapons development across broad time scales.

Keywords: Nonproliferation, verification technologies, radiation detection, passive detection

1. Introduction

Luminescence dosimetry has long been a mainstay in the realms of personnel and accident dosimetry, however recent advances have demonstrated these new techniques may also have a place in nuclear nonproliferation and treaty verification. Previous research has shown that using a combination of optically stimulated luminescence (OSL) and thermoluminescence (TL) allows for the assay of nuclear material with surprising resolution [1]. Likewise, doses to surface mount resistors (SMRs), like those found in common personal electronics, have been measured down to background levels using OSL [2]. Even more recently, a linear array of commercially available optically stimulated luminescence dosimeters (OSLDs) were used to localize the position of a weapons grade plutonium (WGPu) source [3]

One of the principle challenges with respect to weapons treaty verification is that many conventional measurement techniques can reveal classified or protected information about the weapons. As a result, signatories of such treaties would be expected object to the use of such instrumentation for verification activities. Gross dose rate information, such as that measured by luminescence dosimeters, on the other hand, can be expected to be much more palatable to treaty signatories since it is unlikely to contain protected design information.

The purpose of this work is to build upon prior studies to show the plethora of information that can be acquired using dose deposition in ubiquitous materials. While the present study was performed using commercial OSLDs, the same data could potentially be obtained from any number of materials ranging from structural bricks to crystalline particulate matter in dust.

1.1 Analytical Source Position Analysis

As done previously [3] when using a single linear distribution of OSLDs, the positions of a spherical source can be approximated as a point source. Under this approximation, the linear array of measured doses will follow the functional form of

$$Y = m_1 / \left(m_2^{\,2} + \left(Z - m_3 \right)^2 \right), \tag{1}$$

where m_2 is the radial position estimate and m_3 axial position estimate, in cylindrical coordinates. In Equation (1), the m_1 parameter is then simply related to the magnitude of the dose delivered. Under this coordinate system, the "z-axis" is the line connecting the linear array of dosimeters. Applying Equation (1) for a single linear array of detectors, it was shown that the position of the source could be determined with a 1° angular resolution [3].

Hayes and O'Mara [3] also showed that by coupling forward particle transport solutions with black-box optimizations routines, the source position and radius could also be simultaneously determined. It was noted, however, that using the full transport solutions were computationally expensive. As a result, it would be advantageous to use simplified methods, such as applying Equation (1), to solve for the approximate source position and then use the more

computationally expensive routines only to solve for other source characteristics, such as the radius. This work builds upon the results reported by Hayes and O'Mara [3], by illustrating a method by which a gridded array (or simply orthogonal linear arrays) can be used to solve for the three-dimensional position of a source. The resultant position could then be used in subsequent full transport models to determine other source characteristics.

2. Materials and Methods

All measurements were made in June of 2019 at the Device Assembly Facility (DAF) at the former Nevada Test Site. The source under investigation was a 4.48-kg sphere of WGPu with a diameter of 7.5876 cm. The WGPu sphere was constructed in 1980 with initial isotopic weight percentages of .02, 93.735, 5.95, 0.2685 and 0.028 for the isotopes of ^{238}Pu, ^{239}Pu, ^{240}Pu, ^{241}Pu and ^{242}Pu respectively with 557 ppm of ^{241}Am.

Two arrays composed of nanodot OSLDs, by Landauer, taped to a foam board backing were constructed and exposed to the source for 85.7 hours. The first OSLD array, termed the large array (LA), was built on a foam board that measured 135 cm long by 90 cm wide. The LA consisted of a 10-by-10 array of OSLDS where the OSLDs on the long axis had a center-to-center spacing of 15 cm and those on the short axis had a center-to-center spacing of 10 cm. The second array, the small array (SA), consisted of an 8-by-8 array of OSLDs, with center-to-center spacings of 10 cm and 7 cm on the long and short axes, respectively. The foam board backing for the SA measured 70 cm by 49 cm in total.

Figure 1 shows the outer configuration of the OSLD arrays with the source in place. It can be seen from the image that the LA was suspended parallel to the floor, above the source using two utility carts. The SA was secured perpendicular to the LA on one of the utility carts. It can also be seen that the source was placed offset from the center of the LA, 20 cm closer to the utility cart supporting the SA.

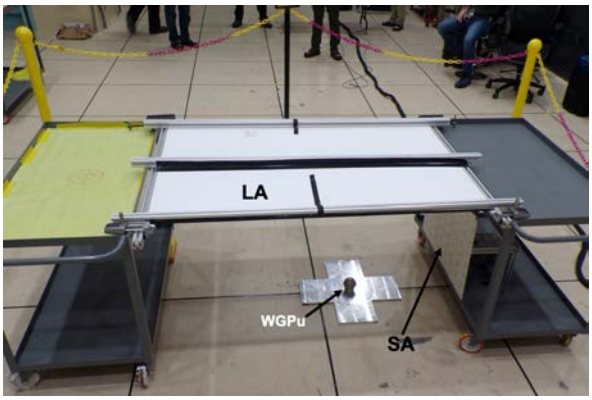


Figure 1: Outer view of the dosimeter arrays. The LA has dosimeters on the opposite side seen here with the SA having dosimeters visible. The clad WGPu was placed on a cross plate of Al on the floor composed of tile on top of concrete.

Figure 2 shows the position of the source from floor level. It can be seen here that the entirety of the SA is positioned above the source. As such, there is no line of OSLDs bisecting the source perpendicular to its central axis where the cladding has a lip. Also shown in Figure 2 is the aluminium stand that the source rests on. The benefit of having both arrays above the source was to reduce the shadowing effect by this stand and to reduce the albedo effect from the floor primarily from the metal cross upon which the WGPu was positioned (Figure 1). There is also a cladding lip around the WGPu which is partially shadowing the lower half of the WGPu to the SA but not the LA as shown in Figure 2.

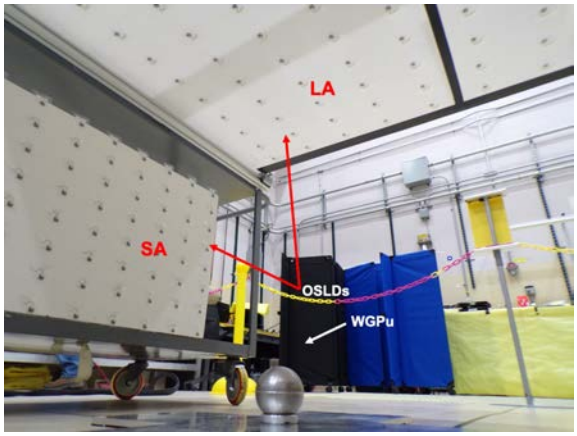


Figure 2: Underside of the exposure geometry, showing the weapons grade plutonium source and both of the dosimeter arrays.

2.1 OSLD Measurement

The dose to each of the OSLDs was measured using a Landauer microStarii® medical dosimetry unit. Prior to the measurement of the OSLD doses, a set of known exposure (NIST traceable doses) OSLDs were utilized to ensure the constancy of the unit calibration and dose estimates. Each dosimeter was read a single time; however, each read consists of a dose estimate from four individual LED pulses. The resultant dose estimate is then the average of the four pulses and the dose uncertainty is its standard deviation.

The nanoDOT™ OSLDs are all calibrated for exposures perpendicular to the top surface of the dosimeter cassette and as a result the dose etimates from the nanodots can be somewhat sensitive to the angle of the dosimeter relative to the irradiation source [4]. The angular dependence of the dosimeters was calculated by a Monte Carlo simulation using MCNP6® [5]. The source's photon spectrum was determined using the Origen module of the SCALE package [6]. The Origen calculation was used to decay the original, measured source material by 39 years and determine the expected present-day isotopic compositions and gamma source rate term. Only those gamma source rate terms with relative contributions greater than 1E-8 were included in the final source specification for the particle transport simulations. The source term was uniformly spread throughout the volume of

the modeled plutonium sphere, in order to account for self-shielding effects.

The photon dose deposition in the Al_2O_3:C chip within each dosimeter set was calculated using a pulse height, energy deposition (*F8) tally. The *F8 tally provides the energy distribution of "pulses" created in a cell modeled as a physical detector [5]. The source spectrum from the Origen calculation was transported mono-directionally at various angles relative to the surface of the simulated dosimeter. Finally, the angular correction factor was calculated as the ratio of the calculated dose deposition for an angle θ to the calculated dose deposition at 90°. The dose measurement for each dosimeter was multiplied by the correction factor calculated for the angle between the dosimeter and the center of the source.

2.2 Source Position Analysis

For each of the two dosimeter arrays, Equation (1) was first fit (using the Levenberg-Marquardt (LM) method) to measured dose profiles both row and column-wise, in order to obtain estimates for each of the two coordinates in the plane of the array (see Figure 3). The final coordinate estimates were calculated as the variance-weighted average of all of the m_2 parameters from the fit to each dosimeter row (or column). Next, the estimated in-plane coordinates were used to calculate the perpendicular distance between the source and the dosimeter plane using the m_3, or radial position parameters from the fits of Equation (1) to each line of dosimeters. Given the estimates of the in-plane (axial) coordinates of the source relative to a given line of dosimeters, the perpendicular distance between the source and the dosimeter plane (e.g. the third spatial coordinate of the source) can be solved by applying the Pythagorean theorem where the length of the hypotenuse of the triangle formed between the source and a given line of detectors is equal to the m_3 parameter given by the fit of Equation (1) to the dose profile. Again, the final estimate of the out-of-plane coordinate was calculated as the variance-weighted average of all of the individual estimates.

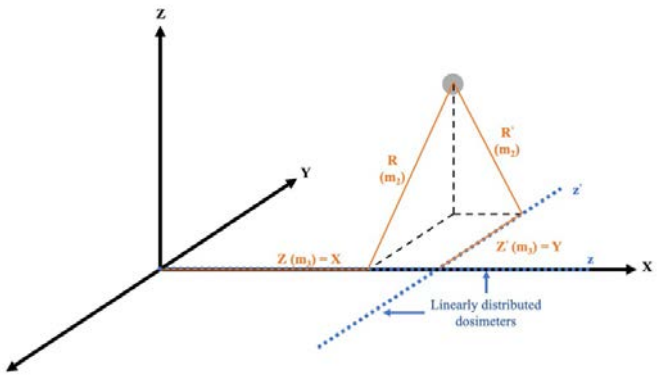


Figure 3: Method for determining the three-dimensional position of a point source using the dose estimates in two orthogonal linear arrays.

With the position estimates determined using the point source approximation, it was hypothesized that an estimate for the source radius would be attainable using forward transport modeling. A simplified model of the exposure geometry was constructing using the MCNP6® code. The modeled geometry included the source, aluminum stand, floor and OSLD arrays. In order to simplify the geometry specification, the utility carts supporting the OSLDs were ignored. Additionally, the OSLD arrays were each modeled as a continuous rectangle of aluminum oxide, AL_2O_3, surrounded by a polyethylene case. Point detector (F5) tallies were placed in the Al_2O_3 region at the locations of the actual dosimeters. While this simplification ignores the spatial extent of the dosimeters, the cost of such a simplification is justified by the greatly reduced computational time required compared to pulse height (*F8) tallies. Further, it is not expected that this approximation will substantially alter the response of the tally compared to the actual energy deposition physics.

2.3 Inverse Transport Methods

The inverse transport problem is a special class of optimization problems where a parameter set in a forward transport model is optimized to match some set of experimental measurements. In many cases, such problems fall into another special class of problems known as black box optimization problems. The term "black box" refers to the non-analytical nature of the forward transport solutions and further, generally implies that gradient information for the parameters of interest is either non-existent or prohibitively expensive to calculate. While many black-box optimization algorithms exist in the literature, this work only focused on relatively simplistic gradient-free, coordinate search methods.

The first, and simplest, solution method (the raster method) consisted of first defining a set of bounds for each of the parameters of interest, the spatial coordinates and radius of the source in this instance, and then computing the chi squared value between simulated and measured doses at equally spaced points within those bounds. Next, the bounds of the previous iteration were moved to bracket the parameter that minimized the chi squared. The subsequent interval was again divided into equally spaced points upon which the chi squared value between simulated and measured doses was calculated. This process of interval refinement was repeated twice and ultimately the parameter value that minimized the chi squared was taken to be the optimal value.

The second method used for the parameter optimization followed the same general approach as the previous method, where each spatial coordinate was optimized independently and sequentially followed by the radius. The main difference was that instead of using a graphical analysis of the chi squared distribution, the second approach used

Brent's method to find the minimum of the chi squared values and the resultant optimum parameter value [7]. The benefit of this approach was that since Brent's method is available in the Scientific Python, *SciPy*, optimization library it was relatively simple to implement and required no interaction with the user [8].

The bounds for each spatial coordinate were selected as a combination of the interval containing the highest dose dosimeters in each direction and any constraints imposed by the physical dimensions of the problem geometry, such as the locations of the floors and the arrays themselves. For example, in the large array, the bounds for the x and y coordinates were taken to be large enough to contain all OSLDs for which a dose greater than 1.6 Gy was measured. This resulted in x-bounds of [-20 cm, 40 cm] and y-bounds of [-25 cm, 45 cm]. The upper boundary of the z-interval was taken to be the largest coordinate of a dosimeter on the small array for which a dose greater than 3 Gy was measured, or between 0 cm and 40 cm, where 0 cm in this case was taken to be the surface aluminum cross plate.

3. Results

Figure 4 shows a graphic rendering of the exposure geometry in addition to the coordinate system used in all of the subsequent analysis and results. The resultant acquired dose profiles for the large and small arrays are shown in Figure 5 and Figure 6 respectively. In these figures, the dose points were plotted on a linear mesh and the resultant dose map surface was generated by linearly interpolating between the measured dose points. Additionally, beneath each surface, in the grid plane, a contour heat map has been plotted.

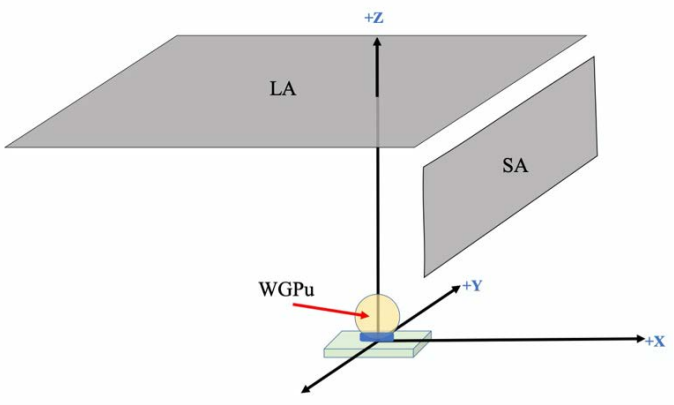


Figure 4: Coordinate system used for the dosimeter arrays with respect to the WGPu source.

The plot in Figure 5 shows an apparent maximum in the dose deposition surface near the (0,0) point in the x-y plane. It was hypothesized that this would be a near optimal measurement scenario since a global maximum implies that the dose measurements have been made on a grid fully containing the source. In contrast, the plot in Figure 6 shows the dose values still increasing at the edge of the z-domain, indicating that the position of the source may fall outside the dose measurement locations in the z-direction. This was indeed the case, as Figure 2 shows, the SA was positioned above the source.

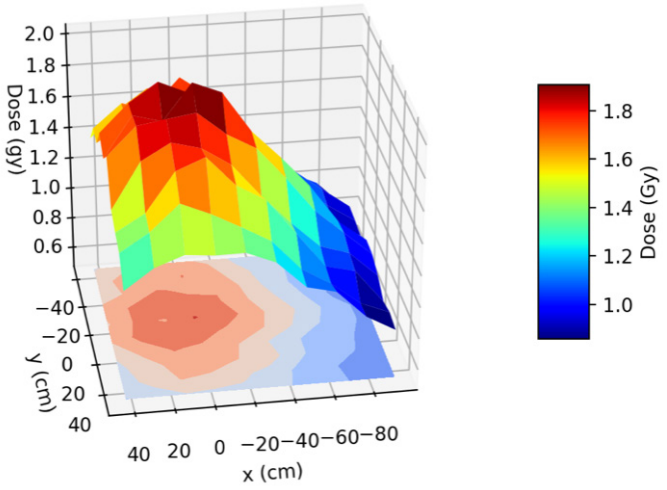


Figure 5: Measured doses to the OSLDs in the large array (LA). The surface map is colored according to the dose, and the doses between adjacent OSLDs were taken as simple linear interpolations. The contour map in the x-y plane shows that a dose gradient exists in both directions, and the OSLD array fully bounded the source distribution.

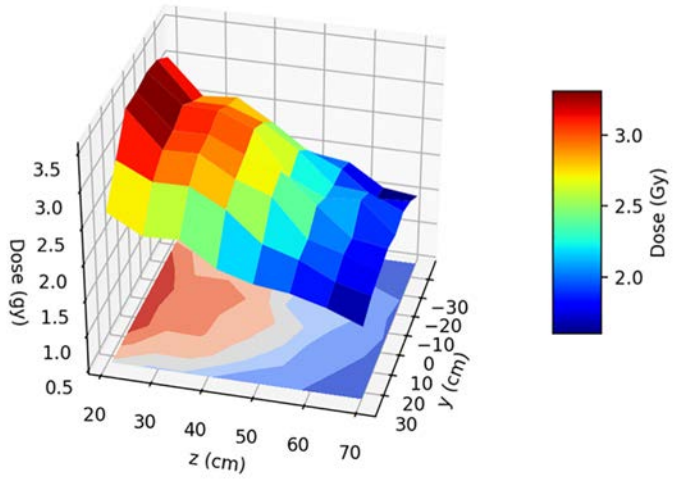


Figure 6: Measured doses to the OSLDs in the small array (SA). The surface map is colored according to the dose, and the doses between adjacent OSLDs were taken as simple linear interpolations. No definite peak is seen in the z-direction, indicating that the dosimeter array does not extend beyond the source location in the z-direction.

Table 1 contains the position estimations calculated from the doses measured in the small and large arrays, respectively, using the analytical method described in Section 2.2. The coordinate system for the dosimeter arrays was defined such that the center of the WGPu sphere was located at position (0,0,4.5) in centimeters.

Array	X (cm)	Y (cm)	Z (cm)
SA	4.4 ± 0.8	-0.2 ± 0.4	10 5
LA	4.6 ± 0.7	-5.6 ± 0.7	2 ± 27

Table 1: Estimated Source Position Coordinates Calculated from the Dosimeters in the Small Array (SA) and Large Array (LA).

It can be seen from the values in Table 1 that the point source approximation method used to estimate the source position tends to poorly resolve all source coordinates. The largest disparity between the assumed true location and the calculated location existed in the y-coordinate estimated from the dosimeters in the large array, followed by the x-coordinate estimates. Referring to the contour plot of the dosimeter measurements in Figure 7, it is apparent that larger measured doses were biased toward the negative y direction. Although steps were taken to align the central axes of each dosimeter array to bisect the center of the source, it is quite reasonable to accept that some uncertainty inevitably existed in the true source locations assumed above. The uncertainty in the x and y positions of the source were taken to be 2 cm each, while the z position uncertainty was taken to be 2 mm. The x and y uncertainty estimates were considerably larger than the z position uncertainty due to the large size of the arrays and the distance between arrays and the source.

It must also be noted that the overall resolution possible in the positional estimates will be inversely correlated to the spacing between the dosimeters, and as result it is expected that observed disagreement between the assumed true and the estimated source positions resulted from a combination of the uncertainty in the true position of the source and resolution limits imposed by the dosimeter spacing chosen. Using the uncertainty estimates for the known source position added in quadrature with position estimate uncertainties, a t-test was performed in order to compare the known source position to the estimates. Table 2 contains the computed *t*-values for each of the coordinates, where the critical *t* at the 95% confidence level was 1.960. It can be seen from the values in Table 2 that only the y-estimate from the SA and the z-estimate from the LA were statistically indistinguishable from the measured values. Therefore, there was disagreement between the known and estimated values that are not accounted for by their associated uncertainties. It was hypothesized that this excess disagreement may have been the result of resolution limits imposed by the spacing of the dosimeters in the arrays.

Array	X	Y	Z
SA	9.13	0.31	3.47
LA	6.87	8.36	0.35

Table 2: t-Statistics Comparing the Known and Estimated Source Positions. Here the null hypothesis was that the known and estimated positions were identical.

Method	Array	X (cm)	Y (cm)	Z (cm)	R(cm)
raster	SA	-1.25	-1.86	10.42	4.49
raster	LA	2.32	-4.86	8.11	4.02
Brent's	SA	5.9	-0.36	18.5	4.22
Brent's	LA	11.0	-9.3	11.1	1.98

Table 3: Estimated Source Position Coordinates Calculated from the Dosimeters in the Small Array (SA) and Large Array (LA).

Table 3 contains the optimum parameters for the source position and radius using each of the inverse transport methodologies described above. It is again apparent that the resolution of the parameter estimates is dependent on the spacing between the dosimeters in the array. One of the challenges with each of the methods tested here was that the total chi-squared values for the arrays were relatively insensitive to changes in the position and source radius for a broad range of values around the true source position. For example, Figure 8 shows the chi-squared value for the simulated and measured dose values for the

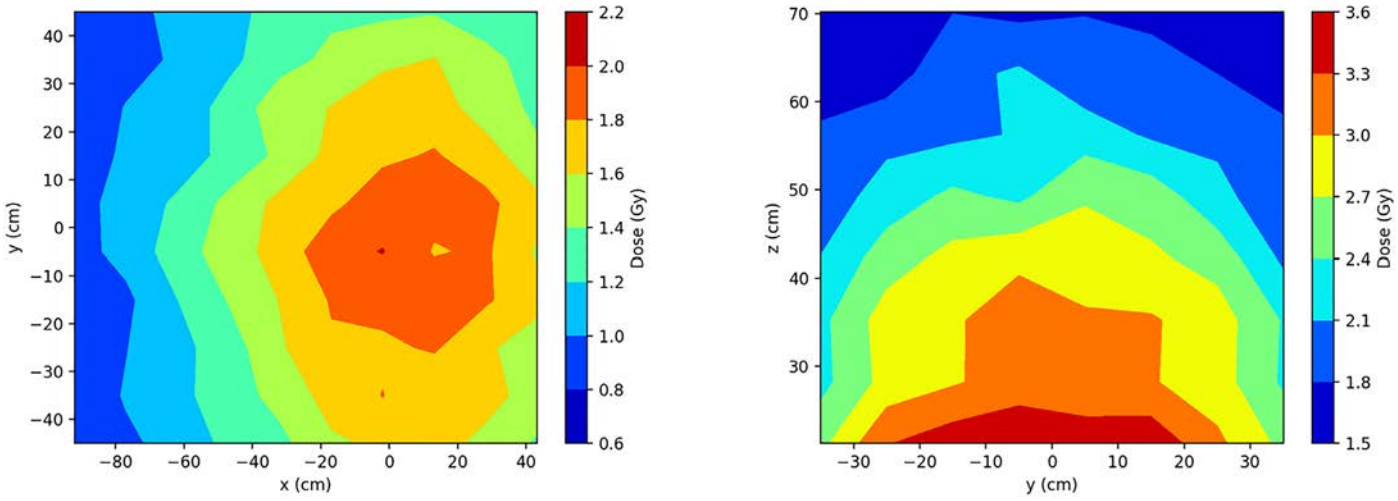


Figure 7: Contour plot of measured doses to the OSLDs in the Large Array (left), and Small Array (right).

dosimeters in the small array. It can be seen that the changes in the computed chi squared were relatively small over a 5-cm range of radius estimates. This problem is compounded by stochastic nosiness inherent to both the measured and simulated element doses resulting in an unavoidable baseline level of noise in the chi squared as a function of position and radius.

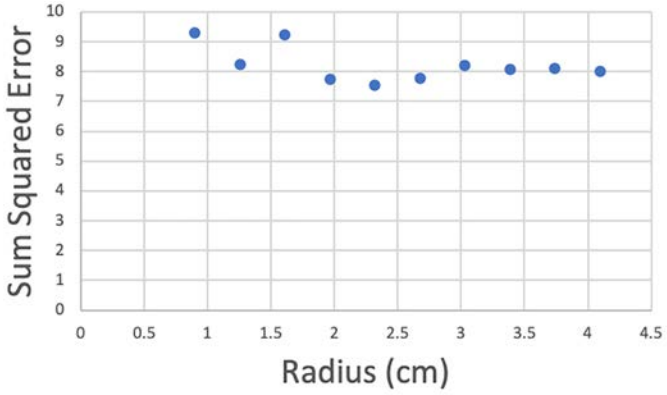


Figure 8: Plot of the chi squared between the simulated and measured dosimeters in the small array as a function of the simulated radius (cm).

Another possible limitation for the raster and Brent's methods is that the final results may be dependent on the order in which the parameters are optimized. The results presented in Table 3 were all computed by optimizing the x-, y-, and z- coordinates followed by the radius. To test this theory, the Brent's method was used to compute estimates for the radius, x-, y-, and z- coordinates (in that order) resulting in estimates that for the radius and x-position (3.73 cm and -5.6 cm, respectively) substantially differed from those presented in Table 3.

4. Conclusions

Previous analyses with a linear array of dosimeters showed promising results for estimating the axial position and radial distance of a source with respect to the line of dosimeters (Hayes and O'Mara, 2020). In this work, two dimensional arrays were used to reconstruct the three-dimensional location of a spherical weapons grade plutonium source. The method consisted of using LM fitting to linearly arrayed dosimeters in orthogonal directions to obtain spatial position estimates. All positional estimates were within 12 cm of the true source position, with the majority of estimates being within 7 cm of the true values. From the position estimates, it was clear that it is important for the dosimeter array to fully contain the profile of the source facing the array. This guarantees that the dose distribution in the array will contain points on each side of the maximum dose.

Using inverse radiation transport methods with coordinate search optimization algorithms, it was again shown that reasonable estimates of the source position (within 10 cm worst case) and source radius (within 2 cm worst case) can be obtained. The major limitation for these methods appeared to be the relatively large range of insensitivity of the chi-squared with respect to the positional and size parameters. Another potential limitation that was discovered during the course of this investigation was that when optimizing multiple parameters, the ordering in which the parameters are optimized can affect the final estimates. The magnitude of this effect, however, can be expected to be inversely proportional to the sensitivity of the results to the parameters being optimized.

In terms of return on investment, the analytical method (Section 1.1) would be the preferable technique for determining the position of a source from a gridded array of dose measurements. Although both of the inverse transport optimization methods required less than 500 model evaluations, and less than 2 hours running time on a single node of a cluster with 32 cores per node, the analytical method returns reasonable results within a matter of seconds on a standard laptop. In terms of making actionable determinations about the presence, or non-presence, of undisclosed source material, the results from the analytical method would be more than sufficient. However, comparable analytical techniques for determining more complex source characteristics such as size, shape, material composition and/or shielding have yet to be developed or tested. As a result, inverse transport optimization methods are, at present, the only and best option for estimating these characteristics. Therefore, future efforts should be devoted to testing the applicability of similar inverse transport optimization methods for estimating more complex, non-positional source characteristics.

Finally, although the dosimetric material used for the dose reconstructions in this study were commercially produced aluminum oxide OSLDs, the ultimate goal is to be able to achieve similar results using minerals derived from ubiquitous materials such as bricks. In general, however, minerals derived from bricks and other earthen based building materials presents additional challenges. Namely, without precise control over the dosimetric material, as there is in commercial OSLDs, there is no guarantee that any samples collected will exhibit sufficient luminescence sensitivity to be useful, especially at low radiation doses. In addition, uncertainties in dose estimates from ubiquitous minerals tend to be higher than those from commercial OSLDs, as a result, the attainable resolution in position and size estimates is decreased for mineral samples. Still, based on the capabilities presented herein, using commercial dosimetry materials, it is likely that the actionable information could be derived using materials derived from environmental materials.

5. Acknowledgements

Acknowledgements to individuals, institutions or funding agencies should be made at the end of the paper before the references.

The authors would like to acknowledge all those who have reviewed these guidelines and contributed to their completion.

6. References

[1] O'Mara, R.P. and Hayes, R.B. *Dose Deposition Profiles in Untreated Brick Material*; *Health Physics*; **114**(4). 2017. 414-420.

[2] Hayes, R.B. and O'Mara, R.P. *Retrospective dosimetry at the natural background level with commercial surface mount resistors; Radiation Measurements*; 112. 2019. 42-48.

[3] Hayes R.B. and O'Mara R.P. *Retrospective Characterization of Special Nuclear Material in Space and Time; Radiation Measurements*; 3. 2020.

[4] Kerns, J.R., Fry, S.F., Sahoo, N., Followill, D.S., Ibbott, G.S. *Angular dependence of the nanoDot OSL dosimeter; Medical Physics;* 38(7). 2011. 3955-3962.

[5] Goorley, T., James, M., Booth, T., Brown, F., Bull, J., Cox, L.J., Durkee, J., Elson, J., Fensin, M., Forster, R.A., Hendricks, J., Hughes, H.G., Johns, R., Kiedrowski, B., Martz, R., Mashnik, S., Mckinney, G., Pelowitz, D., Prael, R., Sweezy, J., Waters, L., Wilcox T., Zukaitis, T. *Initial MCNP6 Release Overview; Nuclear Technology*; 180(**3**). 2010. 298-315.

[6] Bowman S.M. and Gauld I.C. *OrigenARP Primer: How to Perform Isotopic Depletion and Decay Calculations with SCALE/ORIGEN*. ORNL/TM-2010/43, Oak Ridge National Laboratory, Oak Ridge, Tennessee, April 2010.

[7] Brent R.P. *Chapter 4: An Algorithm with Guaranteed Convergence for Finding a Zero of a Function; Algorithms for Minimization without Derivatives*. Prentice-Hall; Englewood Cliffs, NJ. 1973.

[8] Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S.J., Brett, M., Wilson, J., Millman, K.J., Mayorov, N., Nelson, A.R.J., Jones, E., Kern, R., Larson, E., Carey, C.J., Polat, I., Feng, Y., Moore, E.W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E.A., Harris, C.R., Archibald, A.M., Ribeiro, A.H., Pedregosa, F., van Mulbregt, P., *SciPy 1.0 Contributors. "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python; Nature Methods*; 17. 2020. 261-272.

Cognition-Informed Safeguards: Lessons and Recommendations for Safeguards Practitioners from Cognitive Science Research

Zoe Gastelum[1], Arielle Mattes[1], Laura Matzen[2], and Mallory Stites[2]

Sandia National Laboratories, Global Security Division
[1]International Safeguards & Engagements Department
[2]Applied Cognitive Science Department. 1515 Eubank Blvd. SE, Mail Stop 1373, Albuquerque, NM, USA 87123
Email: zgastel@sandia.gov

Abstract:

Despite significant advances in measurement and detection equipment and information analysis methods for international nuclear safeguards, all safeguards verification activities are inherently limited by a common factor: humans. Developments in safeguards equipment and methods are critically important, but so are the care and maintenance of those safeguards practitioners who are expected to use them. The domains of cognitive science and cognitive psychology offer rich information on human performance and cognition. In this article, we summarize key points relevant for international safeguards from extensive literature in the cognitive science sub-domains of attention, cognitive biases, cognitive off-loading and knowledge transfer, the prevalence effect, sleep, stress, task switching and multi-tasking, visual search and visual inspection, wayfinding, and multilingualism. We provide actionable recommendations in a safeguards-relevant context.

Keywords: *cognitive safeguards, international safeguards*

1. Introduction

The verification of international nuclear safeguards is an inherently human activity, combining observations and data collected and analyzed by human inspectors, analysts, technicians, and clerks to meet a series of technical objectives. Dialogue surrounding international safeguards challenges focuses on the rising workload of inspectors and analysts due to new facilities, changing facility types, new data sources, and stagnant budgets; however, even humans in ideal working conditions face the disparaging reality that people make mistakes. By using principals of cognitive science combined with an understanding of safeguards verification tasks and work environments, we can make recommendations to safeguards practitioners that can potentially enhance their performance in this high-impact field.

Some aspects of the international safeguards verification workflow are highly nuanced and unique to the safeguards domain; for example, escorted navigation of complex industrial environments. However, other activities are more common across multiple domains and can benefit from the existing corpus of cognitive science research without significant modification. This article is intended to share lessons and implications from a plethora of peer-reviewed cognitive science literature with international nuclear safeguards practitioners. Most of this literature is not focused specifically on safeguards experiments, but on relevant and closely related activities from which recommendations can be applied to international safeguards activities. Some research has been published by members of our research team focusing explicitly on nuanced safeguards-relevant tasks.

In this article we summarize literature from the cognitive science domain and apply lessons from other fields directly related to safeguards verification tasks to develop cognition-informed safeguards recommendations. The recommendations herein apply to a broad array of safeguards planning and verification activities. The recommendations have also been published as a stand-alone booklet.

2. Attention, Inattentional Blindness and Attentional Misdirection: How Our Minds Focus on What is Relevant

Attention refers to the "means by which we actively process a limited amount of information from the enormous amount of information available through our senses, our stored memories, and our other cognitive processes" [1]. Attention, and what captures our attention, is a significant area of research within the cognitive science domain. Kahneman [2] describes attention as the internal mechanisms that determine or select the significance of stimuli, as well as the degree of that selection. According to Sternberg [1], the three primary types of conscious attention are 1) signal detection, in which one must detect a particular stimulus; 2) selective attention, in which one must choose to attend to some stimuli while ignoring others; and 3) divided attention, in which one allocates their attention to complete more than one task at a time.

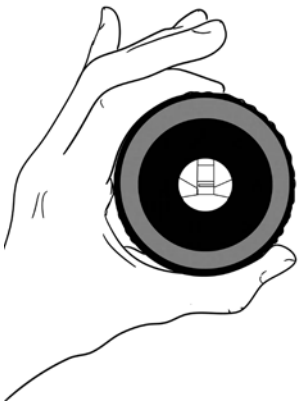


Figure 1: Attention is an internal mechanism that determines the significance and degree of stimuli from the vast amount of information available through our senses. Two mechanisms in which attention fails to capture a relevant event include inattentional blindness and attentional misdirection.

There are two related but unique instances in which our attention fails to capture a relevant event: inattentional blindness and attentional misdirection. *Inattentional blindness* is a phenomenon in which "unexpected objects fail to capture attention" [3]. Inattentional blindness, for example, occurs when someone is paying such close attention to a task that they miss something "in plain sight." One of the most popular illustrations of this was an experiment by Chabris and Simons [4] in which participants were asked to count how many times balls were bounced by individuals wearing white t-shirts in a recorded video. When a woman in a gorilla costume dances through the scene in the middle of the video, about half of the participants miss seeing her the first time. The original video of the experiment is located here: https://www.youtube.com/watch?v=vJG698U2Mvo (accessed June 11, 2020).

Attentional misdirection is defined as the "deliberate diversion of attention away from a visually salient stimulus" [5]. In both inattentional blindness and attentional misdirection, one's attention is directed towards a certain object or activity, but misdirection is done deliberately with the intention of preventing a person from noticing something. Attentional misdirection has been most studied in the context of illusions and magic tricks [6]. There is debate within the cognitive science community about whether attentional misdirection from a third party is distinct from inattentional blindness, but both phenomena offer a situation in which one's attention is focused on a single event or activity causing them to "miss" pertinent information.

Attention is crucial for inspectors working in the field, as they need to be aware of their environment while conducting safeguards activities. Attention is important for maintaining personal safety, correctly completing inspection tasks, and noting any unusual activities. Attention is also critical for analysts and technicians working at the International Atomic Energy Agency (IAEA) Headquarters. Strict attention is necessary for performing detailed analyses, laboratory procedures, maintenance and calibration on sensitive equipment, and other activities.

Inattentional blindness and attentional misdirection can adversely impact safeguards practitioners—especially those working in the field. The close attention necessary for a copious task—like checking many seal identification numbers—is beneficial for good performance but could cause inattentional blindness to an abnormal reconfiguration of equipment in that area. Social cues or gestures by facility operators or regulators can lead to attentional misdirection, causing inspectors to miss potential indications of undeclared nuclear activities.

Recommendations following our review of the attention literature are below.

Allocate Team Roles: An experimental study conducted by part of our research team at Sandia National Laboratories evaluated inattentional blindness in an international safeguards scenario in which participants compared simulated electronic inventory lists to facility documentation [7]. In this digital list comparison activity, participants checked items off one list as they found them in the second list. During the experiment, the background color of the digital screen changed. Participants were asked to announce when they noticed the changes during the scenario, and recount at the end how many times they observed the color changes. The results of the experiment confirmed findings from Chabris & Simons [4] that only about 40-60 percent of the population will notice a change to the environment during an intense concentration activity. We recommend teamwork allowing one person to be intensely focused on a safeguards verification task, while a second person provides support, engages with the operator or escort, and observes the environment.

Notice Opportunities for Distraction and Misdirection: Inspectors working in the field should be aware of opportunities for distraction and misdirection. In studies of participant response to magic tricks, researchers found that repeated exposure (i.e., more than 10 trials), and especially repeated exposure with feedback on performance, increased a person's ability to discern a real event (in this study, a coin toss) from a simulated event in a magic trick [8]. We recommend that safeguards inspectors be exposed to multiple trainings in deceptive environments, such as the Complementary Access trainings, to support in-field recognition of potential misdirection scenarios.

Use Checklists for Important Tasks: Since attention may be diverted to stimuli other than the task at hand, Gawande [9] recommends the development and use of simple checklists to ensure critical tasks are being completed according to the required protocols. While Gawande developed his checklists for the medical community, similar definitions of key steps and procedures to avoid missing an important activity or analysis step can support more effective safeguards

verification results. We recommend that checklists be developed for several safeguards tasks, both in-field and headquarters based. Example of safeguards checklists might include a checklist of sources to search for an open source analysis, or steps to complete in physical inventory verification for a nuclear power reactor.

Use Visual Aids. One study found that when presenting new information, visual aids support better retention of the information [10]. For State Evaluation Groups (SEGs), positive progress has already been made in the domain of incorporating visual information, such as in environmental sampling results visualization [11] and visual representation of nuclear materials flow through a state [12]. We recommend vetting or peer-reviewing visual information to ensure that the initial reaction to the visual data leads to a correct interpretation of visual presentations.

3. Cognitive Biases: Unconscious Predeterminations that Impact Assessments and Decision-Making

Cognitive biases are "cognitions or mental behaviours that prejudice decision quality" due to their "deviations from reality" [13]. These biases—or unconscious shortcuts—have been examined and catalogued in several ways, one of the most interesting being the Cognitive Bias Codex in which the author classifies cognitive biases by four scenarios in which they arise: too much information, not enough meaning, need to act fast, and not knowing what one should remember [14].

Numerous papers describe the cognitive biases that influence decision-making [15, 16, 17]. Examples of a cognitive bias include:

- The availability heuristic, in which the ease of information retrieval (i.e., examples one can easily bring to mind) increases one's perception of the global frequency of an event [18].
- Confirmation bias, which is the tendency to overvalue information that supports an existing belief [14].

Figure 2: Cognitive biases are unconscious shortcuts in decision-making that occur when the human brain has too much information to process, insufficient meaning or context, or does not know what to remember.

Cognitive biases, and how they impact decisions and actions, are especially relevant to safeguards during in-field inspections, interpretations of deviations from state declarations, analysis of open source data, and the integration of multiple sources and types of safeguards-relevant information. Cognitive biases can also impact group scenarios such as SEGs, in which individuals and groups make assessments and judgements about safeguards verification in a state. Gazze, Wilson, Mathews, Reyes, & Schanfein [19] explored many cognitive biases as they specifically relate to SEGs, and recommended three phases of action to counter biases that included specific recommendations for training, peer review, mentoring, and quality assurance.

Our recommendations for managing cognitive biases are listed below.

- Use Structured Analytical Techniques. Some authors suggest the use of structured analytical techniques to counter potential biases [19][21]. There are many different structured analytical techniques, each suitable for different types of analysis or questions. Examples of structured analytical techniques include key assumptions check and analysis of competing hypotheses. Though there has been little experimental research into the impact of using structured analytical techniques, research reviewing a selection of U.S. intelligence documents found the reports that described using these techniques "addressed a broader range of potential outcomes and implications than did other analyses" [22]. We recommend that members of SEGs evaluate which structural analytical techniques are most suitable for their analyses and integrate them into the state evaluation processes.
- Engage a Facilitator to Support Equitable Information Sharing. One study found that, in groups, extroverted individuals may be interpreted as being expert in areas beyond their expertise [23]. Another group of researchers recommended paying special attention to differing viewpoints as part of an overall strategy to combat cognitive bias [24]. We hypothesize that active participation in SEGs may be lower for new staff or staff working at lower levels of fluency in English, potentially compounding cognitive biases. We recommend that an IAEA staff member not associated with the SEGs but familiar with facilitation should mediate SEG discussions of important issues or key analysis. Mediation would ensure that all staff are given an opportunity to participate, and that the group uses techniques to avoid groupthink.

4. Cognitive Off-Loading and Knowledge Transfer: Taking Notes to Remember Information and Share it with Others

Cognitive off-loading is "the act of reducing the mental processing requirements of a task through actions like writing down information or storing information in a cell phone or computer" [25]. Notetaking is one form of

cognitive off-loading for short-term tasks, but it can also be used for longer term memory and knowledge transfer. Some research indicates that beyond just providing a written record for reference, the act of writing notes by hand can improve higher-level comprehension in the short-term and mitigate the forgetting of information over time [26].

Knowledge transfer refers to the handover of insights (know-what) and experiences (know-how) between individuals or teams [27]. The term knowledge transfer can be used to refer to transferring institutional knowledge or to transferring task-specific knowledge. In this work, we consider only task-specific knowledge transfer. Knowledge transfer has been studied in the psychology community especially within shift workers such as medical providers or power plant operators. Bosua and Venkitachalam [27] found that, due to incomplete or inefficient knowledge transfer methods, "incoming workers tend to solve problems with inadequate information, have an incomplete understanding of significant events that occurred in prior shifts, while workers often attempt to solve the same problems across different shifts."

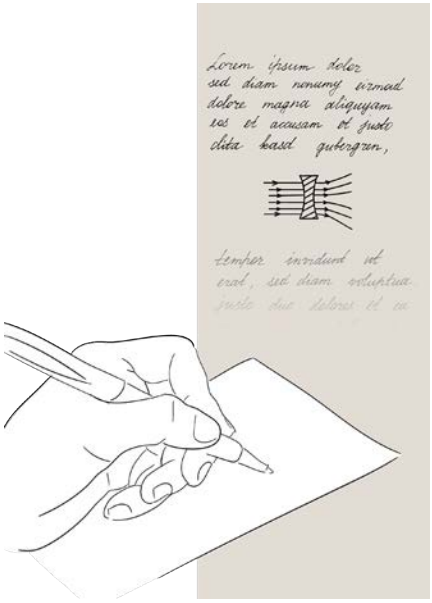


Figure 3: Cognitive off-loading is the act of reducing mental processes through actions like notetaking. Knowledge transfer is the transfer of information—specifically insights and experiences—between individuals or teams.

Notetaking has relevance for many safeguards activities, most notably for on-site inspections. Notetaking can be used to record observations and task status, record personnel information, make illustrations of key pieces of equipment, etc. For headquarters-based activities, notetaking can be used in SEGs to record the analytical findings and interpretations of each SEG member. Notetaking can also be used by individual inspectors or analysts as they complete their activities that will later support SEG activities.

Knowledge transfer is critical as inspection teams may differ between visits to a site. While knowledge transfer is traditionally studied for round-the-clock personnel handoffs, international safeguards inspectors face an additional challenge of the time gap between when they complete their activities and when the next set of inspectors will begin preparing for theirs—which could be weeks or months apart.

Our recommendations for notetaking and knowledge transfer are below.

Recommend Times for Taking Notes. One study on notetaking used a computer-based task in which participants arranged circles on a screen according to visual instructions or patterns provided within the experiment. Some participants were instructed to place visual reminders for themselves at certain points in the task to remind themselves of circle placement, while others could set reminders spontaneously. The researchers found that self-perception of memory capability influences when participants set reminders, with individuals who believe they have high memory capability setting fewer reminders than others. The study found that setting the visual reminders improved performance for all participants regardless of self-perceived memory capacity [28]. For safeguards, we recommend that reminders be set for inspectors or analysts to take notes as checkpoints throughout their activities, such as at natural stopping points during breaks or between tasks.

Define a Structure. In the cognitive science literature, "boundary objects" refer to physical or electronic representations that can be used to transfer knowledge between individuals or organizations. Research in the educational psychology field found that when people are provided with boundary objects such as outlines in advance of an activity, they can use those outlines to support more organized notetaking and demonstrate improved memory [29]. Even if an outline or notes structure cannot be in-hand during an inspection, having the structure in mind while in the field could support better notetaking. This recommendation supports current practices for the IAEA's use of structured forms such as inspection reports.

Take Multimodal Notes. In a study on the impacts of drawing on notetaking [30], researchers found that drawing an item provided support for recalling that item later relative to writing text alone, across multiple settings, instructions, and encoding strategies. The researchers propose that the positive impact on recall of drawing over writing may be due to the integration of multiple types of memory involved in recalling an object from drawing. A similar impact of drawing was shown in an experiment conducted by members of this research team, designed to simulate a safeguards scenario. We found that notetaking improved memory of complex visual arrays relative to memory for items studies without

a notetaking aid. Notes that included both text and drawings/illustrations were most useful—both to the individuals who took the notes and to other people who were given the notes to use to complete a change-detection task without having seen the initial array [31].

Use Cameras When Available, But Not Exclusively. In the same study, we found that the use of digital cameras increased accuracy on some change detection tasks relative to handwritten notes only, especially for subtle changes. Using a camera saved time compared to taking handwritten notes, but using cameras at the time of test took longer than referring to handwritten notes. Further, the use of digital cameras made participants more confident in their findings even when the participant's conclusions from those photographs were wrong, whereas participants with handwritten notes were less confident when they were wrong [31]. Digital cameras are recommended when available, but should always be accompanied with written notes or other sources of information.

Review Inspection Notes. The educational psychology community has repeatedly examined two distinct functions of notetaking—storage and encoding. The *storage function* "suggests that the review of notes stored in written form facilitates retention" by helping students consolidate noted information, stave off the natural forgetting process, or re-learn information that was already forgotten [32]. The *encoding function* "suggests that the process of recording notes facilitates learning even in the absence of review" [32]. In other words, the process of writing the information down increases one's ability to recall it later. For safeguards, both of these functions can support inspection activities. For storage, notes can be shared with others or used to support documentation during a knowledge transfer process. For encoding, inspectors who write down the information and then re-visit the same location will be more likely to remember their notes. Kiewra [32] describes several studies that show the benefit of reviewing notes. In one study, participants who did not attend a lecture but reviewed borrowed notes performed better on an exam than those who attended lecture but didn't review their own notes. We recommend reviewing one's own notes at the conclusion of an inspection activity and any notes that are available in preparation for a new inspection.

Define Hand-off Procedures. In a series of case studies on knowledge transfer in various industries—including manufacturing, information technology, and heavy industry—researchers found that having a defined procedure for preparing for and conducting shift handover activities facilitated effective knowledge transfer [27]. Practices that appeared to have positive impact included periodic mandatory training to review the procedures with staff, and management-defined information to be recorded for handover, including how to document unusual events and the method for doing so (e.g., fill out a form and give it to a single point of contact). This practice required effective infrastructure such as common access via computer systems to all so that staff who needed access could get the required information. For safeguards, the recommendations from this study imply that the IAEA should continue using standardized inspection reporting forms that can be accessed by those with appropriate roles.

Record Inspection Briefings. One study on hand-off strategies from high-consequence failure industries—including ambulance dispatching, railroad, nuclear power, and space shuttle mission control—identified 21 strategies for hand-off of information [33]. One of the most relevant strategies for safeguards recommended was that hand-offs be completed in person and audio recorded for later reference. We recommend that inspection reports presented to management or SEGs be recorded and then used for reference by inspectors preparing for their next visit.

5. Prevalence Effects: Searching for Rare Events or Objects

The prevalence effect refers to the phenomenon in which observers are more likely to miss a rare target (i.e., the thing that is being searched for) than frequent targets [34]. The prevalence effect has important impacts on many visual search activities, such as airport security screenings. The prevalence effect can occur in different manifestations dependent upon the type of task being conducted, and can result in ending the search prematurely, missing a target due to a rare configuration, or missing a target due a premature reflexive response.

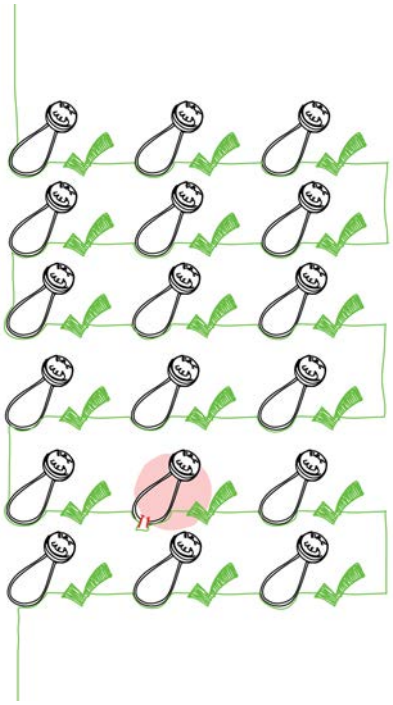


Figure 4: Prevalence effect is the phenomenon in which one is more likely to miss a target that occurs with low prevalence (i.e., low frequency) than targets that occur with higher frequency.

There are many examples within international safeguards in which an inspector or analyst is looking for low prevalence phenomena. A few examples include:

- In verifying containment of equipment or cabinets, or the integrity of seals, very few will have evidence of tamper.
- In reviewing scientific and technical publications, few publications will indicate undeclared research relevant to the nuclear fuel cycle.
- In complementary access visits, few locations will have evidence of undeclared nuclear activities.

Based on the cognitive science literature, our recommendations regarding the prevalence effect for safeguards are below.

Train with Periods of Increased Prevalence. Wolfe and colleagues [35] found that inserting short periods of high prevalence targets can help mitigate the prevalence effect. While this would be impossible to do in natural environments such as inspections or open source information collections, it could be attained by sending inspectors or analysts to trainings in which they experience high prevalence with feedback. There remain several uncertainties regarding the duration of the effectiveness of these trainings, or how training environments can generalize to real-world safeguards activities. We recommend adopting several training methods on a trial basis for safeguards and monitoring both participant feedback and any differences in safeguards outcomes.

Prepare for Field Activities with Brief Periods of High Prevalence. Wolfe, Brunelli, Rubinstein, & Horowitz [36] suggest that "a regimen of a brief high-prevalence block just prior to going to work…might be worth investigating as a countermeasure to the prevalence effect." We further suggest that while realistic high prevalence effects are difficult to simulate for safeguards, a game, mobile application, or virtual or augmented reality in which inspectors encounter high prevalence of a safeguards-interesting theme completed just prior to a safeguards verification activity could counter prevalence effect. For example, inspectors could complete a safeguards-relevant game in which they confirm seal inventory with a high rate of tampered seals, incorrect seal numbers, or missing seals.

6. Sleep Deprivation: How to Get Good, Effective Sleep at Home and Away

Sleep is important to cognitive functions including memory consolidation [37] and attention [38]. Further, *sleep deprivation*—not getting sufficient sleep—can significantly affect functioning, including cognitive and motor performance, as well as mood [39]. According to the Sleep Foundation guidelines, most healthy adults need seven to nine hours of sleep, while older adults need seven to eight hours of sleep per night [40]. However, sleep is often disrupted, especially due to travel across time zones (jet lag) or sleeping in unfamiliar environments such as hotels.

Figure 5: Sleep impacts many cognitive functions, such as memory consolidation and attention. Disrupted sleep can negatively affect cognitive and motor performance, and mood.

Safeguards inspectors and technicians who frequently travel for on-site activities in nuclear facilities are likely to experience disrupted sleep due to time changes and sleeping in unfamiliar places. While our findings will focus on those travel-related sleep disruptions, recommendations for sleep hygiene apply to all people who want an effective and restful sleep.

Our safeguards-specific recommendations for getting good sleep and recovering from disrupted sleep from the cognitive science literature are below.

Provide Time for Sleep Recovery. The sleep research community documents what it calls the *first night effect*, in which the first night of sleep in an unfamiliar environment is disturbed [41]. The sleep disruptions from the first night effect are similar to insomnia. Sleep researchers have observed that the disturbed first night of sleep in a new environment (usually due to participation in a sleep study) can result in poorer memory consolidation [42] and attention [43]. Others found that after a night of sleep deprivation, returning to normal took one to two nights of normal sleep depending on the type of cognitive task studied [44]. This is consistent with findings presented in Alhola and Polo-Kantola [43] that one night of additional rest supported recovery of cognitive function. We recommend grouping closely geolocated inspections together, so that an inspector stays longer in a single time zone. This may help mitigate the detrimental effects of jet lag and first night effects.

Practice Good Sleep Hygiene. In general, sleep researchers recommend good sleep hygiene. These are activities or preparations that promote good sleep and daytime alertness. According to the National Sleep Foundation, good sleep hygiene practices include: limiting daytime naps to 30 minutes; avoiding stimulants such as caffeine and nicotine close to bedtime; using alcohol only in

moderation; exercising regularly (but not too close to bedtime); avoiding food or drinks that may trigger indigestion; ensuring adequate exposure to natural light; establishing a regular, relaxing bedtime routine; and making sure that the sleep environment is pleasant including comfortable mattress and pillows, cool temperatures, low to no light, and white noise machines or other relaxing sounds [45]. Other sleep studies recommended avoiding cognitive arousal in bed, such as worrying, planning, or thinking about important things; not engaging in exciting or emotionally upsetting activities before sleep; and avoiding excessively noisy environments for sleeping [46].

Track Sleep to Identify Poor Sleep or Sleep Deprivation. Alhola and Polo-Kantola [43] suggest keeping a sleep diary to track sleep in order to better help identify when a recovery period is needed. We also recommend tracking sleep with personal devices such as fitness trackers that can monitor the quality and duration of sleep.

7. Stress: Working Under Pressure

Stress is defined in several ways within the cognition community. Staal and colleagues [47] describe three models of psychological stress, which revolve around an individual's cognitive and behavioral responses to stressors within their environments. Descriptions of stress within the cognitive science community range from temporary exam-related anxiety, to strongly stressful single events like natural disasters, to lingering post-traumatic stress disorder from a period of intense stress, to chronic stress from our daily lives (We are unable to make clinical recommendations for how to deal with intense stress, PTSD, anxiety, or other mental health topics. If you believe you are experiencing stress that is greater than typical day-to-day stress described here, seek out the help of a mental health professional or your doctor.). Across the spectrum of stress, cognition can be impacted by stress "with situations that can range from the mild interference that exposure to brief stressors can induce on the ongoing processing of information to the impact of traumatic experiences on the establishment of enduring and devastating memories" [48].

Yerkes and Dodson [49] were some of the first researchers to describe an early increase in performance as stress increases, with performance reaching an optimal point and then degrading as stress continues to increase. This is referred to as the Yerkes-Dodson performance curve. Sandi [48] found that high levels of stress reduce performance, but that mild-to-moderate levels of stress can support attention and memory. This section focuses on periodic, moderately stressful events such as those encountered in some work environments relevant for international safeguards.

Figure 6: Stress is a psychological and physiological response to stressors in an environment. While some mild, short-term stress can enhance attention and memory, ongoing exposure or brief high levels of stress can hurt performance.

Stress can be part of any professional work environment. International nuclear safeguards practitioners, however, may experience additional stress due to a number of unique factors. For inspectors working in the field, stresses can include difficult negotiations with facility operators or national regulators, working in potentially hazardous environments, or working under time constraints to complete inspection tasks. For analysts or technicians working at IAEA Headquarters, stresses can include having limited time to complete many important tasks, unanticipated short bursts of activity from high-priority tasks that require immediate attention, and interpersonal stresses related to working in an international environment with a high staff turnover rate.

We have adapted recommendations for coping with – and building resilience to – stressful situations for safeguards below.

Focus on a Single Task. Ansari, Derakshan & Richards [50] found that higher anxiety individuals (a personality trait measured through a self-assessment) had lower working memory and less ability to shift cognitive tasks. They found that the more anxious an individual is, the more likely they are to get distracted by environmental stimuli and be unable to accurately or quickly complete their required task. We recommend that during work situations of heightened anxiety, safeguards practitioners focus on a single task or activity they have to complete before preparing for or starting the next activity.

Build Resilience Through Training. Staal, Bolton, Yarough & Bourne [47], whose research focuses on military applications, indicate that experience and expertise build resilience to stress that could otherwise impair performance. They recommend that training under pressure may help reduce "choking or panic during subsequent performance under pressure." We recommend that safeguards training include moderately stressful scenario role-playing about difficult negotiations or time-constrained inspection tasks.

Team with Experienced Peers. Another study tested the performance of higher and lower skilled soccer players who, while viewing five-second segments of soccer games from a first-person perspective, were asked to anticipate the next action of the player in possession of the ball [51]. Those leading the study tried to incite high anxiety among the participants by telling them that their results would be compared to those of other players and evaluated by the coach. The study found that the players responded to anxiety differently according to their level of skill, with the lower-skilled players suffering a larger negative impact on task performance than the higher-skilled players. While soccer moves might not be directly relevant to safeguards, we propose that the ability to predict what will happen increases with experience. We suggest pairing more experienced inspectors with less experienced inspectors in potentially anxiety-inducing scenarios in order to mitigate negative performance impacts.

Express Positive Reinforcement to Reduce Stress. Cognitive scientists frequently study evaluation anxiety, which is defined as "a cognitive and emotional experience that commonly arises in social, academic, clinical and vocational settings" in which the stress of being "tested" impacts people's ability to perform the tasks that are required [52]. Coy and colleagues describe one explanation for evaluation anxiety and its negative impact on cognitive performance is through Cognitive Interference Theory, which suggests that "negative off-task self-dialogue…interferes with performance by distracting an individual from the task at hand". In one study, Coy and colleagues proctored exams to groups of students, some of whom received intentionally anxiety-inducing instructions and some of whom received neutral instructions. Those who received the anxiety-inducing instructions reported more negative, off-task self-dialogue after the test, and performed worse on cognitive tests. While safeguards practitioners are not routinely "tested" in this way, and the authors do not examine the impact of positive self-dialogue, we recommend positive self-dialogue before and during potentially stressful situations.

8. Task Switching and Multi-Tasking: Addressing Competing Cognitive Demands on One's Attention

Task switching and multi-tasking exist on a continuum of divided attention in which attention is allocated across multiple tasks at once. *Multi-tasking* refers to situations in which an individual must split their attention between multiple tasks simultaneously. *Task switching* refers to situations in which individuals switch their attention between the sequential performance of two or more tasks that require at least partially different processing sources [53]. Both multi-tasking and task switching are challenging because they draw upon limited processing resources, and often result in a decrement to performance, either through slower or less accurate task performance.

Figure 7: Task switching is when one alternates their attention between tasks. Multi-tasking is when one divides their attention across tasks simultaneously.

Safeguards inspectors working in the field are often required to task switch and multi-task. During a typical onsite activity, an inspector might be checking inventory lists, verifying seals, paying attention to spatial navigation, maintaining situational awareness, talking to their inspection partner or facility operator, and handling equipment. Safeguards analysts working at headquarters are more likely to face task-switching effects between analytical tasks, between tasks and email, and being interrupted.

We have tailored recommendations from the cognitive science literature on task switching and multi-taking for safeguards below.

Support Single-Tasking for Analytical Tasks. Decades of work in cognitive psychology have found both dual-task costs as well as switch costs, in which task performance is lower when people try to simultaneously perform two tasks or switch between two different tasks, respectively [54]. We suggest the promotion of single-tasking within the institutional culture for headquarters-based safeguards activities. For example, we suggest scheduling "focus time" in which there are several hours of uninterrupted work time during the day.

Use Breaks to Switch Tasks. While task switching is generally described as being detrimental to task performance, there are some instances in which switching tasks has been shown to have some benefits to performance. For example, even a brief break during a vigilance task can boost performance [55]. By working only on a single, high-concentration task for 30 minutes and then intentionally switching tasks, task switching can serve as a mental break from cognitively demanding concentration activities. We recommend intense focus on single tasks to the extent possible for durations up to 30 minutes, followed by a brief break or switching to a task that requires different cognitive functions. If an

inspection partner is present, switching roles with that partner can also be used as a cognitive break.

When Multi-Tasking is Required, Choose Non-Overlapping Cognitive Tasks. Research on distracted driving suggests that distractions from cellphone communications—both on a device and hands-free—are harmful to task performance [56]. Similar outcomes have been found in the wayfinding literature, in which secondary spatial and verbal tasks performed by the participants interfered with the encoding of wayfinding information [57]. However, we are constantly task switching and multi-tasking in our daily lives, for example, when listening to a news broadcast while driving or talking with family while cooking. Given that even seemingly unrelated tasks can interfere with each other, we recommend that when multi-tasking and task switching is necessary, safeguards inspectors find tasks that can divide attention between complementary, rather than competing, cognitive capabilities.

9. Visual Search and Inspection

Visual search is "a scan of the environment for particular features—actively looking for something when you are not sure where it will appear" [1]. Search tasks can include activities such as a *feature search*, when a distinctive feature such as color or size defines the search criteria, or a *conjunction search*, during which a specific combination of features is being searched for together. Visual inspection is a similar but specialized activity. *Visual inspection* is "careful and critical examination, especially for flaws…[and] is typically a deliberate, in-depth exacting process that requires more than mere looking or scanning" [55].

Figure 8: Visual search is a scan of the environment for a target or feature that might not be present. Visual inspection is a more deliberate and careful examination, especially for flaws.

Visual search is important for physical inventory and design information verification activities, as well as checking containment of safeguards equipment, examining seals for evidence of tamper, and satellite imagery analysis.

Our safeguards recommendations for visual search and visual inspection, based on the cognitive science literature, are provided below.

Take Time to Avoid Errors. In an overview of human factors literature relevant to visual inspection, researchers cited studies in which time pressure can result in inspectors being more "lenient" and inadvertently letting borderline cases "pass" an inspection [55]. For safeguards, this could possibly translate to not noticing indications of tamper. One study conducted by members of our research team specifically focus on visual search for a safeguards-like inventory verification task. We found that the presentation of list information affects the speed at which participants can complete a list-matching activity, and that presenting the two lists in a very similar order produced the fastest response times. We also found that presenting lists in this fastest configuration resulted in decreased accuracy in the detection of subtle errors (e.g., a transposition of two digits rather than a missing item or more blatant mismatch). Although presenting information to enable more efficient safeguards inspections may be preferred, we recommend that sufficient time should be allocated for all visual inspection tasks to help prevent avoidable mistakes.

Have Sufficient Lighting for the Task. In a literature review of factors that impact visual inspection accuracy, Megaw [68] points out four factors that impact inspector error—visual acuity, lighting condition, time for inspection, and feedback. Time and feedback are covered elsewhere in this section, and visual acuity should already be considered for this type of activity. The overview of research on lighting indicated that "good lighting is essential in reducing visual fatigue," and that adequate lighting can reduce the difference in visibility for low- and high-contrast items. However, the author also notes that too much lighting could induce glare for some activities. IAEA inspectors should ask for appropriate task lighting for their activity, which might mean turning the lights down for some activities. Headquarters-based staff should also ensure that they have sufficient task lighting, for example when analyzing satellite imagery or inspecting seals returned from the field.

Provide Feedback. In a study that looked at the impact of performance feedback on visual search tasks in airframe structural inspections, researchers found that providing feedback on search process and strategy resulted in more improvements to the task performance (measured in time and accuracy) than only providing feedback on the accuracy of the search [69]. Feedback in graphical or visual forms (including visual representations of the participants' search patterns) provided a positive effect on users' performance. While the research focused on an extended visual search task over the entire fuselage of an airplane, we posit that similar feedback on process or strategy could benefit smaller visual search tasks relevant to international

safeguards. We recommend that performance feedback be provided to inspectors and analyst by their peers, and that the feedback include comments on accuracy and strategy or method.

Define a Search Strategy. Using a defined search strategy—such as a snake patterns from side to side, top to bottom—has been shown to improve inspection performance in visual inspection tasks over a random search pattern [70, 71]. We recommend that safeguards inspectors and analysts be trained in specific search patterns in order to be more effective when examining for indications of tamper, searching for items within overhead images, and conducting other safeguards visual searches.

Take Breaks or Switch Tasks Often. In an overview of human factors research on visual inspection, See [55] recommended limiting visual inspection activities to 30 minutes at a time based on anticipated fatigue for this type of task. The authors recommended working in short segments of time, and then either switching tasks or taking a break from inspection for 15 minutes after each 30-minute inspection period. Because safeguards inspectors are time-limited and not able to easily take breaks due to the protective clothing and radiation screening required in many facilities, we recommend switching tasks or roles (e.g., equipment user and recorder) every 30 minutes for similar effect.

Efficiently Display Lists. In a study that examined how lists of surnames with first initials are presented visually on a screen, researchers found that decreasing the screen density—adding white space or a blank row between items—of an electronic list led to faster search times [72]. For these lists, they added more space by suppressing redundant information (in this case repeated surnames) rather than repeating them for each entry along a vertical column. Matzen and colleagues [7] also found that arranging lists in numerical order supported more efficient list searches. We recommend adopting efficient list organization practices for IAEA-controlled information. If an inspector will be searching IAEA-controlled lists (seals inventory lists, for example) or a physical inventory list provided prior to an inspection activity, the list should be appropriate spaced and arranged in alphanumeric order. Repetitive text such as the first portion of ID numbers could be partially truncated to make the lists easier to scan.

10. Wayfinding: Using Sense of Direction and Navigational Aids

Wayfinding refers to how people understand and locate themselves in physical space. It can include navigation (finding the way between desired locations), route knowledge (familiarity with where one has already been), landmark knowledge (locations of points of interest), and site or survey knowledge (an overall conception of the layout of a physical location or area).

Figure 9: Wayfinding is an understanding and ability to locate one's self in physical space, and includes finding the way between desired locations, familiarity with the route one has taken, knowledge of locations of interest, and an overall conception of the layout of an area.

Wayfinding plays an important role in onsite safeguards verification activities. Safeguards inspectors must often navigate through unfamiliar geographic areas to locate nuclear facilities, before being escorted through typically complex industrial spaces. Safeguards inspectors need to know their exact location within a facility, the route they took within the facility to get there, and the location(s) where they observe key equipment or measuring points.

Our wayfinding for safeguards recommendations, based on the cognitive science research, are provided below.

Maintain Active Awareness. In one study on the impact of navigational aids on individual wayfinding activities, researchers found that GPS users travelled more slowly and longer distances in a route navigation task than users of paper maps or experience-based navigation (in which people learned a route from an experimenter) [58]. The study found that GPS users made stops to re-orient themselves more frequently than paper map users, and map users made stops more frequently than experience-based navigation participants. Participants were equally able to reach their end route regardless of their navigation aid. We recommend that when navigating to an unfamiliar nuclear site, inspectors familiar with the site actively engage with newer inspectors to teach them where the site is. For first-time or less familiar visits, we encourage the paper-based maps rather than GPS.

Use Landmarks. One study found that the use of landmarks in navigation instructions help orient people within in an environment and support the development of a cognitive map [59]. The authors recommend providing instructions that include landmarks and using those landmarks to re-orient one's self during navigation for more holistic understanding of an environment. If navigational aids such as GPS are to be used, the authors recommend a system that will reference landmarks in map visualizations. In a similar study, researchers found that local landmarks (e.g., a landmark within

a facility) improve route knowledge, while global landmarks (e.g., external but nearby parking lots, rivers, or mountains) improve the broader knowledge of an area's layout [60]. These types of knowledge contribute to different aspects of wayfinding capabilities. The combination of both local and global landmark features improves performance in both route and survey tasks. The authors recommend including both types of landmarks on maps for best performance. For maps the IAEA controls (such as those provided to IAEA and then sent with inspectors or those created internally), we suggest emphasizing both local landmarks within a facility and global landmarks.

Avoid Distractions. One study investigated the impact of secondary tasks (activities participants are asked to complete during the primary activity) on the encoding of wayfinding information [57]. In the study, participants were led through a virtual path while completing a spatial, verbal, or visual secondary task. Both spatial and verbal secondary tasks interfered with spatial encoding, such that participants in these conditions were more likely to get lost when they repeated the path than those in the visual secondary task or control condition. Based on these findings, we recommend that safeguards inspectors being escorted through an unfamiliar area ask to stop momentarily if they need to pay close attention to a conversation or assess their spatial surroundings.

Use Your Most Efficient Map. In a wayfinding experiment conducted by members of this research team, we designed our task to address the unique needs of safeguards inspection environments, specifically: it was conducted in an indoor, complex industrial environment using guided/passive navigation, and participants were given paper maps to use. Our results showed that individuals with a *low sense of direction* perform worse when using maps on wayfinding tasks—such as developing survey knowledge of a facility—even if the map is studied prior to the navigation task [61]. Our results further showed that individuals with a *good sense of direction* perform better on some wayfinding tasks—like retracing their routes—when using a map. But on other aspects of task performance, such as situational awareness, the map provides no value to those with a good sense of direction and can even be detrimental to performance. Individuals who have a good sense of direction should study a map prior to an inspection if possible, and refer to it in only a limited way during on-site activities to avoid distraction [61]. We further recommend that if wayfinding is necessary during a team-based activity and involves placing important equipment in specific locations or retracing routes, the map should be designated to one member of a team, preferably one with a good sense of direction.

Train on Spatial Navigation Skills. One study found that during wayfinding activities individuals with a poor sense of direction pay less attention to spatial features, landmarks, and orientation than those with a better sense of direction [62]. The authors hypothesize, but have not experimentally confirmed, that additional spatial navigation training could support acquisition of some of these wayfinding skills. We recommend training safeguards inspectors or others who perform in-field activities on basic wayfinding skills.

Assess Your Abilities. Many studies have confirmed that an individual's sense of direction is highly indicative of their performance on wayfinding tasks. Self-assessments such as the Santa Barbara Sense of Direction (SBSOD) scale indicate that people do well in assessing their own capabilities [63]. We recommend that individuals know their own abilities.

11. Working in a Multilingual Environment

Language serves two essential purposes for communication: a) receiving, decoding and comprehending input (from external sources); and b) expressing and producing encoded language output (for external sources) [1]. Language is essential for communications—both verbal and written. One's primary language is referred to as L1, their secondary language L2, and so on. According to Kroll et al. [64], "being bilingual is not only about acquiring and using a second language (L2), but also about the ways that the native or dominant first language (L1) changes in response to the L2." Further, Kroll and colleagues suggest that the use of two languages "may enable bilinguals to develop special expertise that extends beyond language into cognition, shapes the brain networks that support cognitive control, and provides cognitive resources that are protective when individuals are old or cognitively impaired" [64]. Working in a multilingual environment poses unique challenges.

Figure 10: Language is essential for oral and written communication, as it is used to both encode outgoing messages for others and decode messages from external sources. People who are multilingual can communicate in more than one language.

The operating language of the IAEA Department of Safeguards is English. However, many safeguards staff members are not native English speakers, and therefore may have differing degrees of fluency in the English language.

Inspectors often speak different languages in the field, and many staff communicate with stakeholders in other languages. Additionally, analysts search and process information in many different languages, and multilingual ability is highly sought after for this job category.

Our safeguards recommendations for working in multilingual environments based on cognitive science research are below.

Be Aware of Context. One study highlighted the importance of nonverbal communication skills when working with speakers of multiple languages, for improvements to an individual's overall communication abilities, awareness, intelligence, and social interaction [65]. Other researchers found that when following instructions from non-native speakers, listeners rely more on context for interpretation [66]. We recommend that in addition to awareness of the content that individuals are communicating, they also pay close attention to nonverbal communications and the cultural context in which they are being displayed.

Use Multiple Forms of Communication. Research from Lev-Ari [66] suggested that over-relying on context rather than actual language, when interpreting directions from a non-native speaker, caused errors in an item-selection task. We recommend that anyone communicating in a multilingual environment—whether they are a native or secondary speaker of the language—ask for clarification when terms are ambiguous or could be misunderstood, or use additional methods of communication such as summarizing and using visual aids.

Gain Proficiency for Reduced Anxiety and Better Performance. In one study of advanced-level language learners in university-level Spanish courses, the advanced-level students reported lower levels of anxiety related to reading comprehension tasks and follow-up activities than introductory and intermediate-level students [67]. The researchers found that although stress had been found to impair comprehension in lower-level student, it did not have the negative effect on advanced students. This indicates that increased proficiency in operational L2s improves performance due to both comprehension and communication abilities, and the lesser impact from the stress of working in L2. We recommend that safeguards practitioners working in L2, L3, or beyond continue to practice and take coursework, if available, to support better performance.

12. Conclusion

International nuclear safeguards practitioners carry a heavy cognitive load, and their inherent potential for human error could lead to negative outcomes for the global community. Though humans will always make some errors, we can learn how to best support and enable human performance from the vast corpus of cognitive science literature. While the recommendations here may not be relevant for all safeguards verification activities, and may have unintentionally excluded relevant elements of the cognitive science domain, we hope that the recommendations provided here are both relevant and actionable to best support human performance across a range of safeguards tasks. Future research explicitly testing unique safeguards scenarios or environments, and using real safeguards practitioners, will continue to add value to this emerging domain of applied cognitive science for international safeguards.

13. Acknowledgements

The authors thank Jason Bolles and Haley Norris, of Sandia National Laboratories, for their tireless graphic design to support this work. The authors also wish to thank our peer reviewers for their thoughtful feedback. This research was supported by the United States Department of Energy's National Nuclear Security Administration Office of International Nuclear Safeguards (NA-241).

This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

14. References

[1] Sternberg, R. J. (1999). *Cognitive Psychology.* Forth Worth, Texas: Harcourt Brace College Publishers.

[2] Kahneman, D. (1973). *Attention and Effort.* Englewood Cliffs, New Jersey: Prentice-Hall Inc.

[3] Simons, D. J. (2000). Attentional Capture and Inattentional Blindness. *Trends in Cognitive Sciences, 4*(4), 147-155.

[4] Chabris, C., & Simons, D. (2010). *The Invisible Gorilla: How Our Intuitions Deceive Us.* New York: Crown Publishers.

[5] Memmert, D. (2010). The Gap Between Inattentional Blindness and Attentional Misdirection. *Consciousness and Cognition, 19*, 1097-1101.

[6] Kuhn, G., & Findlay, J. M. (2010). Misdirection, Attention and Awareness: Inattentional Blindness Reveals Temporal Relationship Between Eye Movements and Visual Awareness. *The Quarterly Journal of Experimental Psychology, 63*(1), 136-146.

[7] Matzen, L. E., Stites, M. C., Smartt, H. A., & Gastelum, Z. N. (2019). The Impact of Information Presentation on Visual Inspection Performance in the

International Nuclear Safeguards Domain. *Human Interface and the Management of Information. Visual Information and Knowledge Management .* Springer.

[8] Cui, J., Otero-Millan, J., Macknick, S. L., King, M., & Martinez-Conde, S. (2011, September). Social Misdirection Fails to Engance a Magic Illusion. *Frontiers in Human Neuroscience, 5*(103), 1-11.

[9] Gawande, A. (2010). *The Checklist Manifesto: How to Get Things Right.* New York: Metropolitan Books.

[10] Padilla, L. M., Creem-Regehr, S. h., Hegarty, M., & Stefanucci, J. K. (2018). Decision Making with Visualizations: A Cognitive Framework Across Disciplines. *Cognitive Research: Principles and Implications, 3*(29). doi:10.1186/s41235-018-0120-9

[11] Vilece, K., Norman, C., Baute, J., Giaveri, G., Kiryi, M., & Pellechi, M. (2012). Visualization of Environmental Sampling Results at Inspected Facilities. *Proceedings of the Institute of Nuclear Materials Management Annual Meeting.*

[12] Norman, C., Binner, R., Caillou, F., Baute, J., Zhao, K., & Walczak, A. (2015). Dynamic Exploratory Visualization of Nuclear Fuel Cycle Verification Data in Support of the State Evaluation Process. *Proceedings of the Institute of Nuclear Materials Management Annual Meeting.*

[13] Arnott, D. (2005). Cognitive Biases and Decision Support Systems Development: A Design Science Approach. *Information Systems Journal, 16*, 55-78. doi:https://doi.org/10.1111/j.1365-2575.2006.00208.x

[14] Heick, T. (2019, July 3). *The Cognitive Bias Codex: A Visual of 180+ Cognitive Biases*. Retrieved from Teach Thought: https://www.teachthought.com/critical-thinking/the-cognitive-bias-codex-a-visual-of-180-cognitive-biases/

[15] Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic Decision Kaing. *Annual Review of Psychology, 62*, 451-482. doi:10.1146/annurev-psych-120709-145346

[16] James, E. H., & Harland Dean, J. H. (2015). Decision Making Under Pressure. *Proceedings of the 60th Annual Financial Analysts Seminar. Fourth Quarter*, pp. 39-45. Chicago: CFA Institute.

[17] Tversky, A., & Kahneman, D. (1981, January 30). The Framing of Decisions and the Psychology of Choice. *Science, 211*(4481), 453-458. doi:0036-807518110130-0453$01.5010

[18] Schwarz, N., & Strack, F. (1991). Ease of Retrieval as Information - Another Look at the Availability Heuristic. *Journal of Personality adn Social Psychology, 61*(2), 195-202. doi:10.1037//0022-3514.61.2.195

[19] Gazze, C., Wilson, B., Mathews, C., Reyes, G., & Schanfein, M. (2019, November). Improving SEG Assessments by Applying Cognitive Science. *Technical Report*.

[20] Heuer Jr., R. J., & Pherson, R. H. (2011). *Structured Analytic Techniques for Intelligence Analysis.* Washington, D.C. : CQ Press.

[21] U.S. Government. (2009). *A Tradecraft Primer: Structured Analytic Techniques for Improving Intelligence Analysis.* Retrieved from https://www.cia.gov/library/center-for-the-study-of-intelligence/csi-publications/books-and-monographs/Tradecraft%20Primer-apr09.pdf

[22] Artner, S., Girven, R. S., & Bruce, J. B. (2016). *Assessing the Value of Structured Analytic Techniques in the U.S. Intelligence Community.* RAND Corporation. Retrieved from https://pdfs.semanticscholar.org/6ebe/134285707619d5bafd8196bc30a62130bcf6.pdf

[23] Strasser, G., & Abele, S. (2020). Collective Choice, Collaboration, and Communications. *Annual Review of Psychology, 71*, 589-612. doi:https://doi.org/10.1146/annurev-psych-010418-103211

[24] Das, T. K., & Teng, B.-S. (1999). Cognitive Bias and Strategic Decision Processes: An Integrative Perspective. *Journal of Management Studies, 36*(6), 757-778.

[25] Morrison, A. B., & Richmond, L. L. (2020). Offloading Items from Memory: Individual Differences in Cognitive Offloading in a Short-Term Memory Task. *Cognitive Research: Principles and Implications, 5*(1). doi:10.1186/s41235-019-0201-4

[26] Bohay, M., Blakely, D. P., Tamplin, A. K., & Radvansky, G. A. (2011). Note Taking, Review, Memory, and Comprehension. *American Jounral of Psychology, 124*(1), 63-73.

[27] Bosua, R., & Venkitachalam, K. (2015). Fostering Knowledge Transfer and Learning in Shift Work Environments. *Knowledge and Process Management, 22*(1), 22-33.

[28] Boldt, A., & Gilbert, S. J. (2019). Confidence Guides Spontaneous Cognitive Offloading. *Cognitive Research: Principles and Implications, 4*(45). doi:https://doi.org/10.1186/s41235-019-0195-y

[29] Star, S. L., & Griesemer, J. R. (1989). Institutional Ecology, 'Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology. *Social Studies of Science, 19*, 387-420.

[30] Wammes, J. D., Meade, M. E., & Fernandes, M. A. (2016). The Drawing Effect: Evidence for Reliable and Robust Memory Benefits in Free Recall. *The Quarterly Journal of Experimental Psychology*. doi:http://dx.doi.org/10.1080/17470218.2015.1094494

[31] Stites, M. C., Matzen, L. E., Smartt, H. A., & Gastelum, Z. N. (2019). Effects of Note-Taking Method on Knowledge Transfer in Inspection Tasks. *Human Interface and the Management of Information. Visual Information and Knowledge Management .* Springer.

[32] Kiewra, K. A. (1989). A Review of Note-Taking: The Encoding-Storage Paradigm and Beyond. *Educational Psychology Review, 1*(2), 147-172.

[33] Patterson, E. S., Roth, E. M., Woods, D. D., Chow, R., & Gomes, J. O. (2005). Handoff Strategies in Settings with High Consequences for Failure: Lessons for Health Care Operations. *International Journal for Quality in Healthy Care, 16*(2), 125-132.

[34] Rich, A. N., Kunar, M. A., Van Wert, M. J., Hidalgo-Sotelo, B., Horowitz, T. S., & Wolfe, J. M. (2008). Why Do We Miss Rare Targets? Explorting the Boundaries of the Low Prevalence Effect. *Journal of Vision, 8*(15), 1-17.

[35] Wolfe, J. M., Horowitz, T. S., Van Wert, M. J., Kenner, N. M., Place, S. S., & Kibbi, N. (2007, November). Low Target Prevalence is a Stubborn Source of Errors in Visual Search Tasks. *Journal of Experimental Psychology: General, 136*(4), 623-638. doi:10.1037/0096-3445.136.4.623

[36] Wolfe, J. M., Brunelli, D. N., Rubinstein, J., & Horowitz, T. S. (2013). Prevalence Effects in Newly Trained Airport Screeners: Trained Observers Miss Rare Rargets, Too. *Journal of Vision, 13*(3), 1-9. Retrieved from http://www.journalofvision.org/content/13/3/33

[37] Cho, K., Ennaceur, A., Cole, J. C., & Suh, C. K. (2000). Chronic Jet Lag Produces Cognitive Deficits. *The Journal of Neuroscience, 20*. doi:0270-6474/00/200001

[38] Ratcliff, R., & Van Dongen, H. P. (2018, February). The Effects of Sleep Deprivation on Item and Associative Recognition Memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 44*(2), 193-208. doi:doi:10.1037/xlm0000452

[39] Pilcher, J. J., & Huffcutt, A. I. (1996). Effects of Sleep Deprivation on Performance: A Meta-Analysis. *Sleep, 19*(4), 318-326.

[40] Hirshkowitz, M., Whiton, K., Albert, S. M., Alessi, C., Bruni, O., DonCarlos, L., . . . Hillard, P. J. (2015, March). National Sleep Foundation's Sleep Time Duration Recommendations: Methodology and Results. *Sleep Health, 1*(1). doi:10.1016/j.sleh.2014.12.010

[41] Le Bon, O., Staner, L., Hoffmann, G., Dramaix, M., San Sebastian, I., Murphy, J. R., . . . Linkowski, P. (2001). The First-Night Effect May Last More Than One Night. *Journal of Psychiatric Research, 35*, 165-172.

[42] Goerke, M., Cohrs, S., Rodenbeck, A., Grittner, U., Sommer, W., & Kunz, D. (2013). Declarative Memory Consolidation During First Night in a Sleep Lab: The Role of REM Sleep and Cortisol. *Psychoneurodoctrinology, 38*, 1102-1111.

[43] Alhola, P., & Polo-Kantola, P. (2007). Sleep Deprivation: Impact of Cognitive Performance. *Neuropsychiatric Disease and Treatment, 3*(5), 553-567.

[44] Ikegami, K., Ogyu, S., Arakomo, Y., Suzuki, K., Mafune, K., Hiro, H., & Nagata, S. (2009). Recovery of Cognitive Performance and Fatigue after One Night of Sleep Deprivation. *Journal of Occupational Health*.

[45] Suni, E. (2020, August 14). Sleep Hygiene. Retrieved August 26, 2020, from https://www.sleepfoundation.org/articles/sleep-hygiene

[46] Gellis, L. A., & Lichstein, K. L. (2009). Sleep Hyiene Practices of Good and Poor Sleepers in the United States: An Internet-Based Study. *Behavior Therapy, 40*, 1-9.

[47] Staal, M. A., Bolton, A. E., Yaroush, R. A., & Bourne Jr., L. E. (2008). Cognitive Performance and Resilience to Stress. In B. J. Lukey, & V. Tepe, *Biobehavioral Resilience to Stress* (pp. 259-300). Taylor & Frna.

[48] Sandi, C. (2013, May/June). Stress and Cognition. *WIREs Cognitive Science, 4*, 245-261.

[49] Yerkes, R. M., & Dodson, J. D. (1908, November). The Relation of Strength of Stimulus to Rapidity of Habit-Formation. *Thr Journal of Comparative Neurology and Psychology, 18*(5), 459-482.

[50] Ansari, T. L., Derakshan, N., & Richards, A. (2008). Effects of Anxiety on Task Switching: Evidence from the Mixed Antisaccade Task. *Cognitive, Affective & Behavioral Neuroscience, 8*(3), 229-238.

[51] Vater, C., Roca, A., & Williams, M. A. (2015). Effects of Anxiety on Anticipation and Visual Search in Dynamic, Time-Constrained Situations. *Sport, Exercise, and Performance Psychology, 5*(3), 179-192.

[52] Coy, B., O'Brien, W. H., Tabaczynski, T., Northern, J., & Carles, R. (2011). Associations between Evaluation Anxiety, Cognitive Interference and Performance on

Working Memory Tasks. *Applied Cognitive Psychology, 25*, 823-832.

[53] Salvucci, D. D., Taatgen, N. A., & Borst, J. P. (2009). Toward a Unified Theory of the Multitasking Continuum: From Concurrent Performance to Task Switching, Interruption, and Resumption. *Proceedings of the Conference on Human Factors in Computing Systems.* Boston, MA, USA: Association for Computing Machinery.

[54] Strobach, T., Wendt, M., & Janczyk, M. (2018, February). Editorial: Multitasking: Executive Functioning in Dual-Task and Task Switching Situations. *Frontiers in Psychology, 9*(108). doi:10.3389/fpsyg.2018.00108

[55] See, J. E. (2012). *Visual Inspection: A Review of the Literature.* Albuquerque, New Mexico: Sandia National Laboratories.

[56] Strayer, D. L., Watson, J. M., & Drews, F. A. (2011). Cognitive Distraction while Multitasking in the Automobile. In B. Ross, *The Psychology of Learning and Motivation* (Vol. 54, pp. 29-58). Elsivier Inc. Academic Press.

[57] Meilinger, T., Knauff, M., & Bulthoff, H. H. (2008). Working Memory in Wayfinding - A Dual Task Experiment in a Virtual City. Cognitive Science, 32, 755-770.

[58] Ishikawa, T., Fujiwara, H., Imai, O., & Okabe, A. (2008). Wayfinding with a GPS-based Mobile Navigation System: A Comparison with Maps and Direct Experience. Journal of Environmental Psychology, 28, 74-82.

[59] Schwering, A., Krukar, J., Li, R., & Anacta, u. S. (2017). Wayfinding through Orientation. *Spatial Cognition and Computation, 17*(4), 273-303. doi:10.1080/13875868.2017.1322597

[60] Lowen, H., Krukar, J., & Schwering, A. (2019). Spatial Learning with Orientation Maps: The Influence of Different Environmental Features on Spatial Knowledge Acquisition. *International Journal of Geo-Information, 8*(149). doi:doi:10.3390/ijgi8030149

[61] Stites, M. C., Matzen, L. M., & Gastelum, Z. N. (2020, March 20). Where are we going and where have we been? Examining the effects of maps on spatial learning in an indoor guided navigation task. *Cognitive Research: Principles and Implications, 5*(13).

[62] Burte, H., & Montello, D. R. (2017). How sense-of-direction and learning intentionality relate to spatial knowledge acquisition in the environment. *Cognitive Research: Principles and Implications, 2*(18). doi:10.1186/s41235-017-0057-4

[63] Hegarty, M., Richardson, A. E., & Montello, D. R. (2002). Development of a self-report measure of environmental spatial ability. *Intelligence, 30*, 425-447.

[64] Kroll, J. F., Bobb, S. C., & Hoshino, N. (2014, June). Two languages in mind: Bilingualism as a tool to investigate language, cognition, and the brain. *Current Directions in Psychological Science, 23*(3), 159-163. doi:10.1177/0963721414528511.

[65] Hall, J. A., Horgan, T. G., & Murphy, N. A. (2019). Nonverbal Communication. *Annual Review of Psychology, 70*, 271-294. doi:https://doi.org/10.1146/annurev-psych-010418-103145

[66] Lev-Ari, S. (2015, January). Comprehending Non-Native Speakers: Theory and Evidence for Adjustment in Manner of Processing. *Frontiers in Psychology, 5*(1546).

[67] Brantmeier, C. (2005, September). Anxiety about L2 Reading or L2 Reading Tasks? A Study with Advanced Language Learners. *The Reading Matrix, 5*(2), 67-85.

[68] Megaw, E. D. (1979). Factors Affecting VIsual Inspeciton Accuracy. *Applied Ergonomics, 10*(1), 27-32.

[69] Gramopadhye, A. K., Drury, C. G., & Sharit, J. (1997). Feedback Strategies for Visual Search in Airframe Structural Inspection. *International Journal of Industrial Ergonomics*(19), 333-344.

[70] Wang, M.-J. J., Lin, S.-C., & Drury, C. G. (1997). Training for Strategy in Visual Search. *International Journal of Industrial Ergonomics, 20*, 101-108.

[71] Nickles III, G. M., Melloy, B. J., & Granopadhye, A. K. (2003). Comparison of Three Levels of Training Designed to Promote Systematic Search Behavior in Visual Inspection. *International Journal of Industrial Ergonomics*, 32, 331-339.

[72] Bednall, E. S. (1992). The Effect of Screen Format on Visual List Search. *Ergonomics*, 35(4), 369-383. doi:10.1080/00140139208967819

esarda.jrc.ec.europa.eu