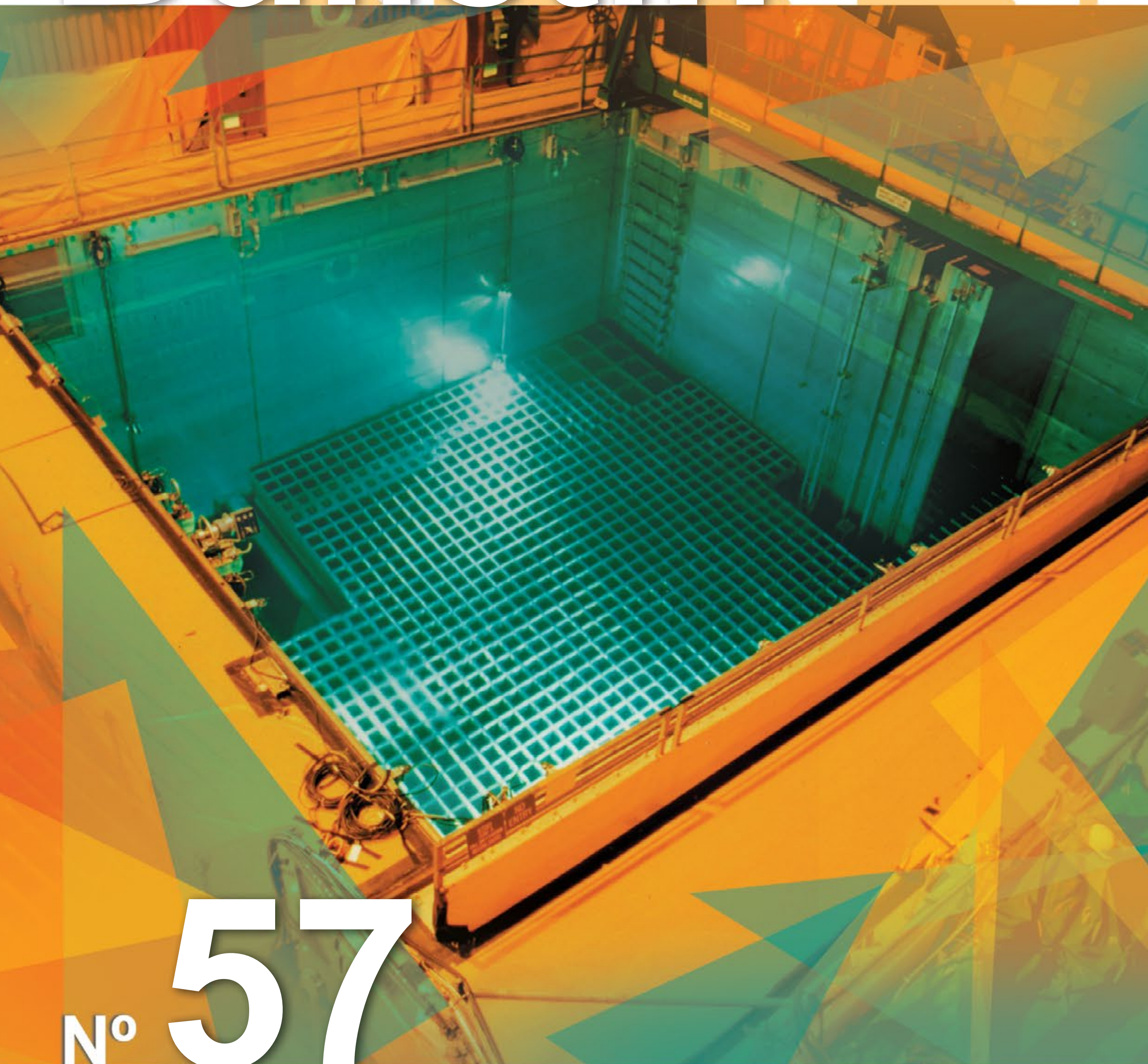
ISSN 1977-5296

ESARDA

European Safeguards Research and Development Association

Bulletin



Nº 57

December 2018

ISSN 1977-5296

Number 57
December 2018

Editor
Elena Stringa
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
Elena Stringa@ec.europa.eu

ESARDA is an association formed to advance and harmonize research and development for safeguards. The Parties to the association are:

Areva, France; ATI, Austria; CEA, France; CNCAN, Romania; EDF, France; ENEA, Italy; European Commission; FZJ, Germany; HAEA, Hungary; MTA EK, Hungary; IRSN, France; MINETUR, Spain; NNL, UK; NRI, Czech Republic; NRPA, Norway; SCK/CEN, Belgium; Sellafield Ltd, UK; SFOE, Switzerland; SSM, Sweden; Springfields Fuels Ltd, UK; ST, Finland; University of Hamburg, Germany; University of Liege, Belgium; University of Uppsala, Sweden; AEA, UK; URENCO, Germany; VATESI, Lithuania; WKK, Germany; PAA, Poland; ORNL, USA

Scientific and technical papers submitted for publication in the peer reviewed section are reviewed by independent authors including members of the Editorial Committee.

Manuscripts are to be sent to the Editor (EC-ESARDA-BULLETIN@ec.europa.eu) following the 'instructions for authors' available on https://esarda.jrc.ec.europa.eu/ where the bulletins can also be viewed and downloaded.

Photos or diagrams should be of high quality.

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 1000 registered members, libraries and institutions Worldwide.

The publication is authorized by ESARDA.

Cover designed by Laura Spirito, (JRC Ispra in Italy)

Printed by
IMPRIMERIE CENTRALE, Luxembourg

ESARDA
European Safeguards Research and Development Association

Bulletin

Table of Content Issue n° 57

Editorial

Peer Reviewed Articles

Editorial

Elena Stringa

Dear readers,

I am pleased to provide you the 57th edition of the ESARDA Bulletin. On behalf of the Editorial Committee, I would like to thank authors and reviewers for the time and energy they have dedicated to the tasks, allowing the publishing in the current Bulletin of high quality articles that are of great interest to ESARDA. All the articles in this issue have been reviewed by at least two independent and expert reviewers, guaranteeing the high standard of the publication.

I encourage ESARDA researchers to publish their work any time so that all the ESARDA community can benefit from the latest relevant research novelties. In order to submit a contribution you are kindly asked to follow the instructions reported in the bulletin section of the ESARDA web site at https://esarda.jrc.ec.europa.eu.

Moreover, I would like to encourage to volunteer as potential reviewer by sending an e-mail to EC-ESARDA-BULLETIN@ec.europa.eu, specifying one's area of expertise among the following:

- Non-destructive analysis
- Destructive analysis
- Arms control and Nuclear Disarmament Verification
- Geological repositories
- Spent Fuel Verification and Spent Fuel Transfer
- Containment & Surveillance
- Process Monitoring
- Statistical Methodologies
- Implementation of Safeguards
- Safeguards Methodologies
- Nuclear Security
- Export control

I also encourage you to cite works published in the ESARDA Bulletin, in order to increase the visibility and interest on ESARDA activities.

I am very pleased to remind you of the following ESARDA events coming in spring 2019:

- April 1-5, 2019: 18th ESARDA Course on Nuclear Safeguards and Non Proliferation to be held at the Joint Research Centre, Ispra (Italy). Information can be found on the ESARDA web site (https://esarda.jrc.ec.europa.eu) under the 'ESARDA course' section.
- May 14-16, 2019: 41st ESARDA Annual Meeting, Symposium on Safeguards and Nuclear Non-Proliferation, marking the 50th anniversary of ESARDA, to be held at the Regina Palace Hotel in Stresa, Italy. Parallel events will also include the ESARDA Working Group meetings and Annette workshops. Information about the Symposium can be found in the ESARDA web site (https://esarda.jrc.ec.europa.eu) under the 'Events' section.

Regarding the ESARDA web site, on behalf of the Editorial Committee, again I would like to address sincere thanks to Andrea De Luca, web master and essential assistant for the ESARDA Bulletin preparation: thank you very much for your engagement, for the pertinent suggestions and fruitful ideas in supporting our work, allowing us to efficiently disseminate ESARDA communications and knowledge.

I would like to take this opportunity also to wish you and your families a fruitful and joyful 2019.

Dr. Elena Stringa
ESARDA Bulletin editor

https://esarda.jrc.ec.europa.eu/
ec-esarda-bulletin@ec.europa.eu
Elena.Stringa@ec.europa.eu

Reporting Security Concerns in the Nuclear/Radiological Industry: New Evidence to the Study of Whistle-blowing

Iryna Iarema and Katherine M. Bachner

Nonproliferation and National Security Department, Brookhaven National Laboratory, Upton, NY, 11973, Office: (631) 344-8271
E-mail: iryna.iarema@hotmail.com, kbachner@bnl.gov

Abstract:

The international community stresses the necessity to report discrepancies in security procedures and build an environment conducive for fostering security culture inside organizations that handle nuclear or radiological materials. In reality, however, there have been a number of instances where reports on nuclear security were not encouraged by organizations or were left without needed corrective actions. Such an attitude, where reports on security matters, instead of serving as an internal 'early warning signal' leading to enhancement of security, have been put aside or entered the public domain after external reporting, is to a great extent caused by a lack of knowledge on how to deal with them and what drives people to report.

This article aims to study challenges and drivers for reporting in the nuclear and radiological sector. First, it discusses the meaning of whistle-blowing and reporting. Secondly, it demonstrates how reporting is encouraged by the international community through IAEA guidance and Nuclear Industry Summit statements. Then by using survey data received from 56 participants, the study examines factors influencing reporting. This analysis is supported by an overview of some real-life examples related to reporting or raising concerns about security procedures in organizations that handle nuclear or radiological materials.

Keywords: whistle-blowing; reporting in organizations; nuclear security; nuclear security culture; drivers of reporting

1. Introduction

With an anticipated expansion of low-carbon energy derived from nuclear technologies or the so-called 'nuclear renaissance' [1], measures should be taken to ensure their safe and secure operation. The role of whistle-blowing is salient in the context of dealing with an insider threat, especially when, as Glynn and Bunn [2] assert, '*nearly all of the documented thefts of highly enriched uranium (HEU) or separated plutonium* [...] *appear to have been perpetrated by insiders*'. This problem is becoming especially acute in the light of amplification of terrorism networks and the risk of their infiltration into organizations that handle nuclear or radiological materials. In such a situation, an employee of a nuclear (or radiological) organization is in the best position to observe a deviance in nuclear security or suspicious behavior and report about it. In addition to that, reliance on reporting could play a pivotal role in deterring and preventing wrongdoing in organizations that handle nuclear or radiological materials.

Being largely instigated by the fact that reporting in the nuclear security field has not yet received a broad discussion in the academic literature, this study will contribute to filling this gap. It hopes to unveil factors that may lead to a greater understanding of the impediments/inducements for the operationalization of whistle-blowing in the nuclear/radiological sector. For example, in relation to nuclear **safety** culture, the IAEA recognised the value of reporting to help continually improve organizational practices and encourages maintaining '*a "blame-free" reporting culture*' to spur '*full reporting of unsafe or unethical practices, incidents and near misses*' [3]. Similarly, increased knowledge about factors influencing reporting of **security** concerns will help to channel management in nuclear organizations in the correct way, bringing practical benefits resulting in better protection of sensitive nuclear materials and facilities.

1.1 Definitions

Although defining whistle-blowing is challenging, there is a need to be explicit about what exactly is meant by the terms we utilize in the current study. Perhaps, the most commonly used definition of whistle-blowing was first provided by Near and Miceli in 1985 [4], according to which whistle-blowing is '*the disclosure by organization members (former or current) of illegal, immoral, or illegitimate practices under the control of their employers, to persons or organizations that may be able to effect action*'. Some [like C. Peters and T. Branch referred in 5] pointed to whistle-blowing as '*the act of disclosing any information that an employee reasonably believes evidences a violation of any law, rule or regulation, mismanagement, corruption, abuse of authority, or threat to public health and safety at the worksite*'. Jubb [6] identified six elements that might be subsumed under the term whistle-blowing: '*act of disclosure, actor, disclosure subject, target, disclosure recipient, and outcome*'.

Interestingly, regulatory provisions in the chemical industry – namely, the U.S. CFATS Act of 2014 [cited in 7] speaks about facilitation of whistle-blower reporting as '***reports** of potential CFATS violations **from employees and contractors** at chemical facilities*', thus expanding the scope of reporting to a broader circle of actors. Building upon this, we would use term *whistle-blowing* with regard to the to the acts of disclosure carried out by an employee or former employee ('*whistle-blower may leave the organization before blowing the whistle*' [4]) or a current or former contractor who reports *internally or externally* about wrongdoing (or lack of actions when they are warranted) in the nuclear or radiological field. Throughout this study, we will also refer to terms such as *reporting* or *informing* and use them interchangeably with *whistle-blowing* to avoid repetition.

2. How reporting of breaches in nuclear security is regarded in international statements

Attention to the issues of reporting security concerns in the nuclear and radiological field is relatively young, nevertheless not without important international commitments, though non-binding. The most prominent in addressing issues of reporting have been: Nuclear Industry Summits and International Atomic Energy Agency (IAEA) guidelines and recommendations.

Started in 2010, Nuclear Industry Summits highlighted the importance of reporting procedures and vigilance for nuclear security matters. In particular, the appeal to '*fostering an open environment for reporting security concerns*' was made in the Joint Statement of the 2012 Seoul Nuclear Industry Summit [8]. In addition to this, the Joint Statement of the 2016 Washington Nuclear Industry Summit [9] also called for '*encouraging employees to report suspicious behavior and/or events through appropriate channels*' [9].

The International Atomic Energy Agency recognizes that the scope of nuclear security extends to '*nuclear and other radioactive material, associated facilities and activities*' [10]. On the level of implementing guides, the IAEA emphasizes the importance of reporting processes for fostering nuclear security culture by such statements as:

> *'Managers need to encourage personnel to report any event that could affect nuclear security. This entails encouraging personnel to provide the security staff with information that could affect security, rather than keeping the information to themselves'* [11].
>
> *'For security, there is the particular need to ensure that staff members understand that adherence to the policy is expected of all personnel. These expectations include protecting information, being aware of potential security concerns and threats, and being vigilant in reporting security incidents'* [11].

The IAEA includes presence of reporting mechanisms to the indicators of a strong nuclear security culture [11]. This international body sees '*protection of individuals who provide information for the purpose of protecting the integrity of nuclear security*' as an antecedent to the establishment of a nuclear power program [12]. Thus a state considering the construction of a nuclear power plant should develop a legislative and regulatory program that contains necessary provisions governing such aspects as whistle-blowing as part of its nuclear security infrastructure [12].

The IAEA model of nuclear security culture has 30 culture characteristics, some of which relate directly to whistle blowing or reporting. For example, to establish and facilitate the process, the model includes characteristics such as a feedback process in management systems, involvement of staff and effective communications in leadership behavior and vigilance in personnel behavior. Culture indicators associated with such characteristics are designed to set standards as well as to provide appropriate tools for periodic implementation of self-assessment, with the focus on whistle blowing and reporting.

Despite the importance of the statements made at Nuclear Industry Summits and recognition of the value of reporting by the IAEA, the process of reporting has not yet been exploited to the full for its capacity to strengthen security inside organizations that handle nuclear or radiological materials. Difficulties arise with practical implementation. Here, Bunn [13] rightfully admits '*Convincing people to report incidents in which they or their colleagues made mistakes or broke the rules is not easy. But experience demonstrates that with the right approach, a culture of reporting can be forged within an organization*'. This begs the question: what is the right approach for encouraging reporting in organizations, and can it be done without deteriorating staff morale? Referring to Miceli et. al. [14], we agree that there are ethical ways to stop wrongdoing via reporting, and information about something which might inflict harm to a large number of people, an organization, the environment, etc. should not be concealed.

Since step-by-step practical guidelines and detailed recommendations are lacking for establishing reporting mechanisms in nuclear and radiological fields, we suggest studying the current state of affairs and the attitudes of professionals toward reporting. In addition to the empirical data gathered by us, this study will also include analysis of the merits of real-life situations on whistle-blowing disclosed in the media or academic literature. A detailed description of our main methodology follows in the next section.

3. Methodology and data

3.1 Description of the methodology

We conducted a survey among people working in the nuclear or radiological industry. The purpose of the survey – to study drivers and challenges for reporting security breaches and actions potentially leading to security breaches within organizations that handle nuclear or radiological materials – was mentioned in the cover letter of an e-mail invitation and its on-line description. The survey consisted of 16 questions and required a total of 10 minutes on average to complete. Questions were formulated in both English and Russian. The survey was available online at the popular on-line cloud-based survey software service for filling out from October, 27, 2016 until November, 7, 2016.

Despite the somewhat sensitive topic, we took several steps to ensure we received enough responses for making this analysis. First of all, the survey was anonymous; the respondents were required only to provide some demographic data. Invitations to fill in the survey were sent to people who work in organizations that deal with nuclear or radiological materials; they were among the professional contacts of the authors or participants at thematic events. The invitation also included a dissemination request. The information about the survey was published on the webpage of the World Institute for Nuclear Security.

As a result, we received 56 completed surveys, which is sufficient, in our view, to make some generalizations on the subject. It is also important to mention that some of the respondents skipped some of the questions, although we suspect that in most cases it happened rather because of an accidental omission than due to purposeful omission. A more detailed account of the profile of our respondents follows.

3.2 Data about survey respondents

3.2.1 Organizational data

Subjects of the study were people working in the nuclear or radiological field (only one person declared that he/she does not work in such a field) who voluntarily participated. They represented different professional roles, which is beneficial for gaining a diverse perspective on whistleblowing in the field. Professional roles were almost equally split (each around 20%) among security specialists, researchers, managers and other categories, a description of which is provided in Figure **1** below.

The majority of our respondents (85.7%) indicated that they have more than 5 years of experience in the nuclear or radiological industry, among whom those who worked in the industry more than 10 years constituted 62.5%. Those who have worked from two to five years in organizations that deal with nuclear or radiological materials comprised 10.7% (see Figure **2**).

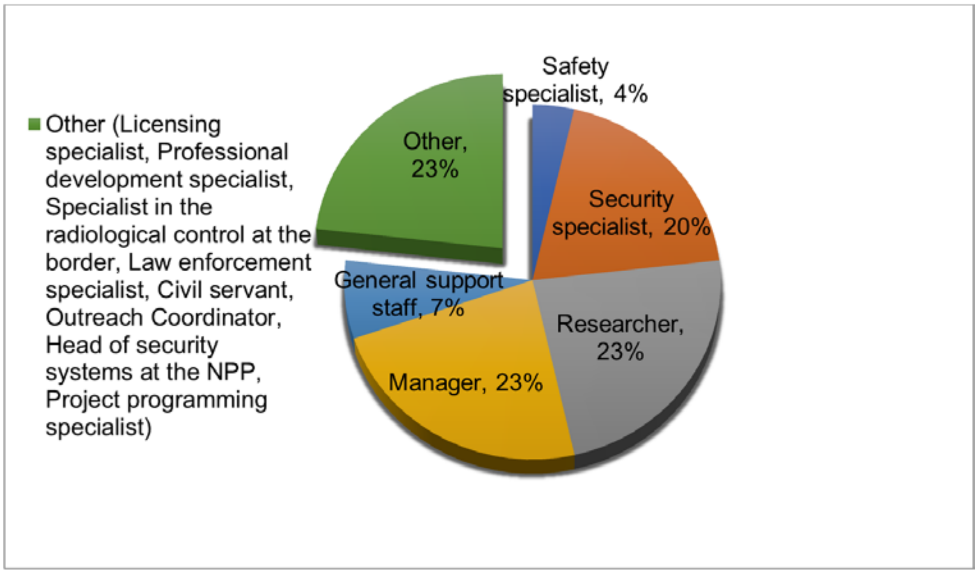


Figure 1: Professional roles of the respondents

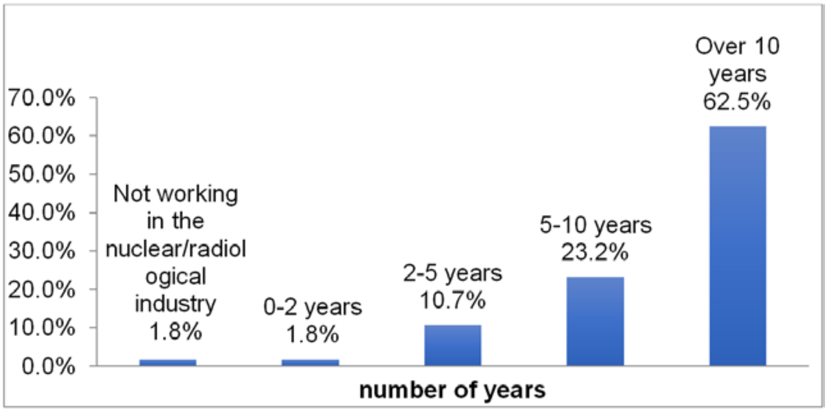


Figure 2: Number of years of experience (of respondents) in nuclear or radiological field

Some studies have indicated that organizational tenure can influence the likelihood of whistle-blowing [15]. Reportedly, newer employees with less experience are less likely to report wrongdoing than more senior fellows, partly because of not being aware of the operational climate in the organization [Dworkin & Baucus, 1998 referred in 15] or appropriate channels for whistle-blowing [Miceli & Near, 1992 referred in 15]. One might see some potential benefits in building upon the argument of organizational tenure to see how to assure a climate favorable for raising valid concerns among less experienced employees and contribute to their empowerment for following organizational procedures of security. For this, one would need to conduct a research study with a larger sample, with follow-up focus groups to develop a reliable picture for the nuclear/radiological industry.

Out of 56 respondents who completed the survey, half work in Ukraine (see Figure **3**). Approximately 29% work in the USA, 5.4% in the U.K. and the rest (approximately 16%) in such countries as Austria, Canada, Germany, Italy, Lebanon, Lithuania, Moldova, and Serbia.

Some studies carried out by Ernst and Young indicated that even in Europe in multinational companies, respondents in the United Kingdom differ from those in France or Austria with regard to their willingness to blow the whistle, where the former would feel more comfortable than the latter [14]. The intention to blow the whistle could also be influenced by regulatory provisions, which in the USA and UK are reportedly more clearly defined than in other countries [14]. In the UK, for instance, the law *'denies protection to whistle-blowers who give information for gain'*, whereas in the USA there is, reportedly, no prohibition against a reward for whistle-blowers [14]. Despite some differences, there are also similarities among countries like the USA and UK, including cultural [14], which also draws our attention to the prospects of exploring cultural phenomena and their influence on reporting mechanisms. This is especially important since in some environments, due to interpersonal tensions, whistle-blowing may be used as a tool to avenge personal grievances or injuries outside the security area.

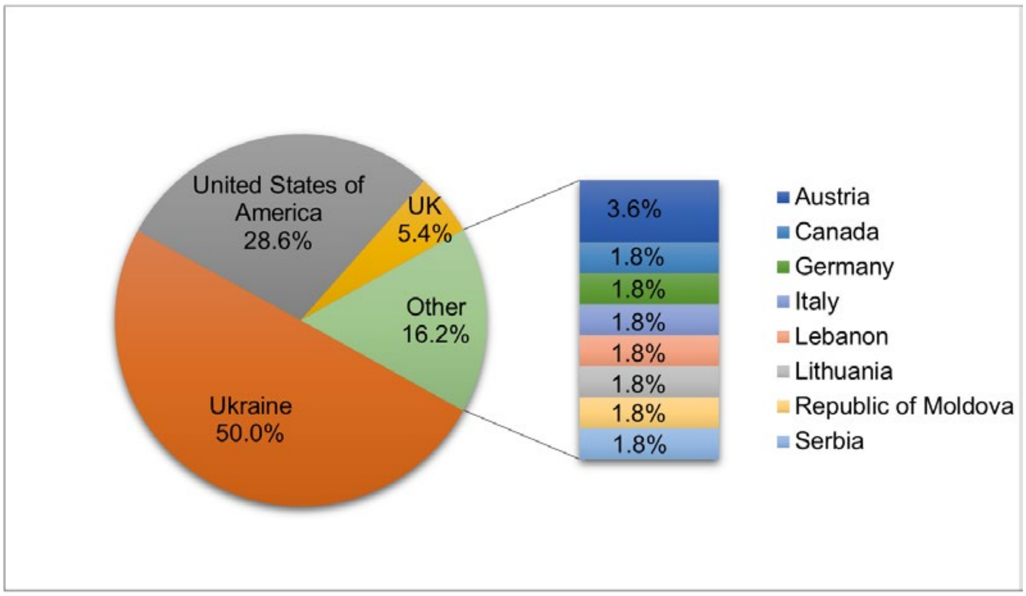


Figure 3: Countries, where respondents work

3.2.2 Demographic data

Out of 56 respondents who completed the survey, 82% were male and 18% female, which reflects the male-dominated character of the industry. There are a number of ways in which data about gender could be considered in a study about reporting non-compliances in the nuclear/radiological field. For example, Miceli [16] in her work summarized the existent research argument about the importance of the gender variable on the propensity to blow a whistle. She pointed to some studies which assert that '*women are more likely than men to blow the whistle*' [16] due, as some suggest, to their '*lower tolerance for illegal and unethical behaviors*' [Yu & Zhang, 2006 referred in 15]. Contrary to expectations that women are more likely to blow the whistle, some have found out that in fact they are less likely, due to lower managerial positions and greater risk of retaliation [15].

The idea that gender, tenure and preference with regard to the recipient of the complaint might be interrelated [referred in 17] should receive more analysis, and so far, based on our data (where the sample size is relatively small), we cannot build a consistent pattern with regard to a gender variable and, therefore, will neither deny nor agree with the statement that gender influences the reporting in nuclear or radiological organizations. However, this is something that might be interesting to explore in the future, especially, keeping in mind efforts towards the expansion of women engagement in the nuclear sphere through the IAEA policies to attract qualified female employees to work in the IAEA [see Resources for Women at 18] and activities carried out by the Women in Nuclear professional network [see 19]. Testing whether involvement of women might influence the reporting behavior in the nuclear industry, and whether gender influences proneness to select internal versus external reporting mechanisms, could bring some value added in the context of providing an opportunity to choose reporting channels that will suit all genders.

Among our respondents, we received a good representation of different age categories (see Figure **4**). The largest group (almost 33%) was those whose members are aged 35-44, followed by a group (23.6%) of people who said they are 55-64 years old. Those participants aged 25-34 and 45-54 formed groups which are the same in size (each 18.2%). Professionals who are older than 65 comprised 7.3% of the respondents.

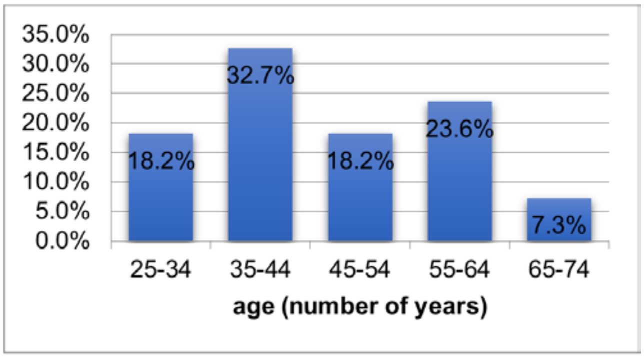


Figure 4: Age of respondents

Similar to the research data that focused on organizational tenure, studies about the age variable concluded that older members of organizations are more likely to report wrongdoing than their younger counterparts, which is explained by their better understanding of , the systems of control and their greater authority within the organizations, leaving them less hesitant to blow the whistle [summarized in 15]. On the other hand, elderly individuals who have extensive work expertise and experience in some situations may have second thoughts about reporting for fear of being forced into retirement as most likely reprisals.

The result showed that most of our respondents received a higher education, with almost 20% being holders of doctoral degrees (see Figure **5**). Around 70% of the participants have master degrees, 7.1% -bachelor degrees, and the remaining 3.6% are evenly split between vocational and other training.

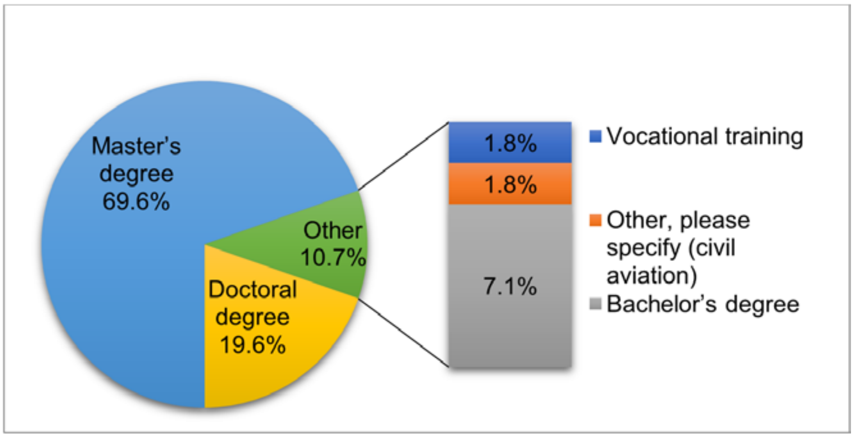


Figure 5: Level of education of the respondents

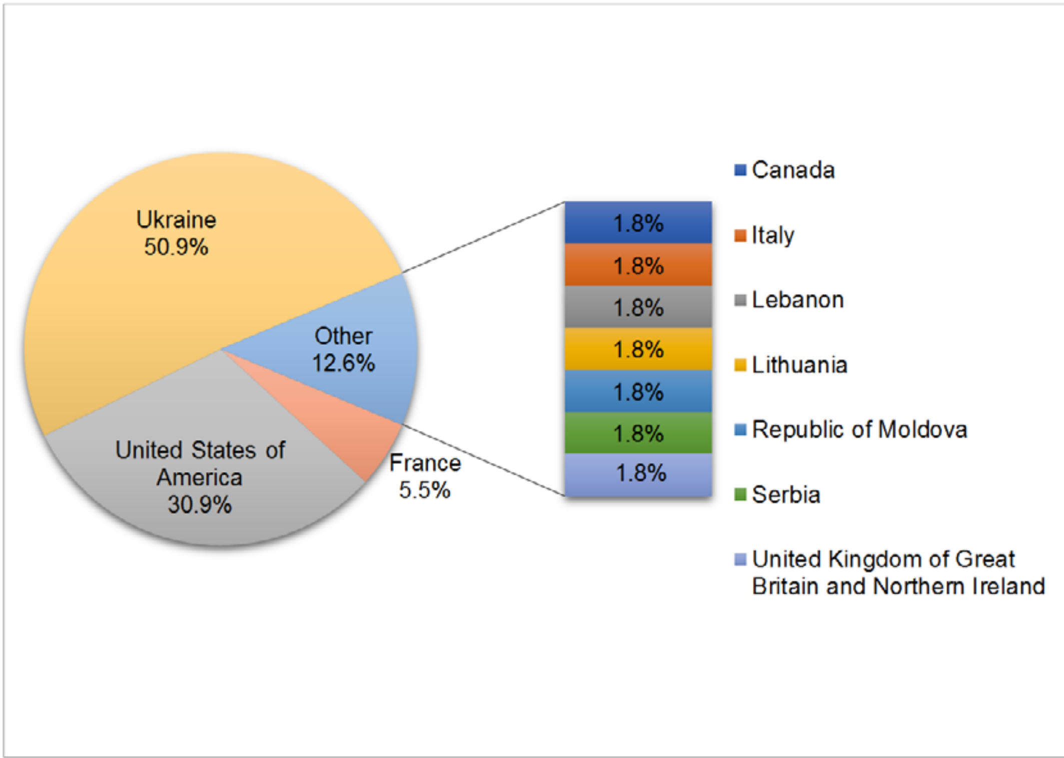


Figure 6: Nationalities of the respondents (N=55, 1 person skipped the question)

Caillier [20] suggested that '*more educated employees are expected to be more likely to blow the whistle than less educated employees, for the reason that the former may have a greater ability to find a job elsewhere if they face reprisals*'. Since in our case almost all of the respondents are relatively highly or highly educated, we will not attempt to draw conclusions based on differences in the respondents' educational level.

We tried to get responses from representatives of different nations, and therefore did not limit distribution of the survey to participants from particular nations; nonetheless, the fact that the survey was available only in the English and Russian languages limited answers to the survey to only those who speak one of these two languages.

The data on citizenship of the respondents is shown in Figure 6. More than half (50.9%) of respondents said they were Ukrainians, 30.9% indicated they were Americans, French persons constituted 5.5% of participants, and the rest (12.6% in total) was evenly split among citizens of Canada, Italy, Lebanon, Lithuania, Moldova, Serbia and the United Kingdom. As can be noticed, the profile of countries is quite diverse, although two major groups based on the country of citizenship can be singled out: Ukrainians and Americans. This, we believe, can be further explored in comparative analysis of some of the results of the survey; in particular, it can be tested whether there is any correlation between an attitude towards whistle-blowing and nationality.

Different researches have acknowledged a mediatory role of national cultural characteristics on whistle-blowing. For example, Miceli et al. [14] warned that research on whistle-blowing done for North American settings '*may not generalize to other cultures, nor even to* [all] *areas in North America*'. A comparison of whistle-blowing on the cultural level was done by Keenan [21], who examined American and Indian managers' propensity to blow the whistle. Ahmad et al. [15] looked at Malaysian whistleblowers in connection with the theory of prosocial behavior. Considering all of this evidence, it seems that attention to culture as a societal environment is a promising field for new discoveries in the complex whistle-blowing problematics.

4. Findings and their interpretation

4.1 What does a violation leading to a security breach entail and what constitutes wrongdoing?

There has been recognition among researchers that '*individuals differ in their perception of what constitutes wrongdoing*', thus some of them might go unnoticed [22]. Lack of security standards in the nuclear field exacerbates the problem with definitions of wrongdoing, since an organization should establish its own security measures based on the design basis threat. Prima facie, wrongdoing in nuclear security can be described as an act initiated from outside or within an organization that bypasses or contravenes security policies, practices, or procedures. However, some may take a deeper view that loopholes or gaps in security constitute a breach in security by revealing a weakness that can be exploited with a malicious intent. Therefore, a starting point of our investigation of reporting procedures in the nuclear or radiological industry was to determine

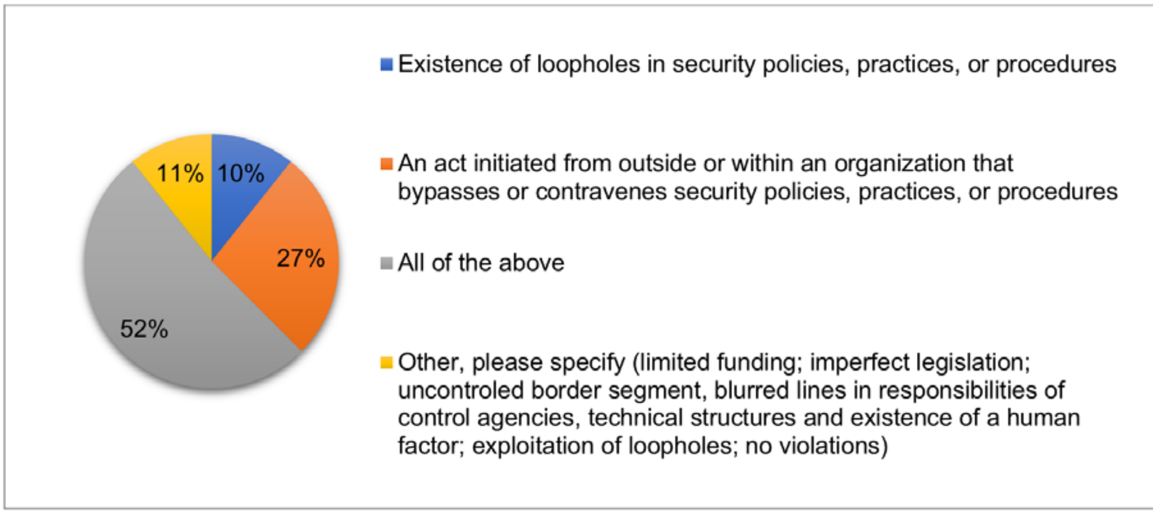


Figure 7: Perception of a violation by the respondents

how our respondents understand a violation that could lead to a security breach. The question itself might provoke certain misinterpretations; therefore, apart from providing possible options to choose from for an answer, we left room for the respondents to suggest their own answer that might best reflect their judgement. The results of the replies are presented in Figure **7**.

The majority of respondents (52%) hold a comprehensive definition of violation, which includes both: an act of an insider or outsider that contravenes security policies, practices and procedures, as well as existence of loopholes in security policies practices and procedures. On the other hand, 27% of the respondents believe that the definition of a violation should be restricted to an act initiated from outside or within an organization that bypasses or contravenes security policies, practices, or procedures. Ten percent argues that loopholes in security policies, practices or procedures are the reason for violations that lead to a security breach.

This inconsistency may be related partly to a matter of linguistics; however, individual assumptions will also affect it. Semantically, one could assume that there is a difference between a wrongdoing and a violation. If the former entails an activity or instance of doing something '*illegal, illegitimate or immoral*', the latter, one could suggest, encompasses not only an act but also a condition that is being violated or leads to a wrongdoing, or in some cases creates favorable conditions for a wrongdoing. For example, non-working cameras at the Y-12 security complex was a clear violation of security procedure. A U.S. DoE report [23] acknowledged that one critical camera with a view on the penetration area was out of service for almost six months, which contributed to '*delays in assessing alarms and identifying the trespassers*' [23]. Still, one may describe it not as an act but rather the lack of an act or a negligent attitude to security policies that allowed anti-nuclear activists, allegedly followers of the Plowshares movement [24], to break into the premises of the Y-12 complex where nuclear weapons-grade uranium was stored [23].

Another example of loopholes could concern the existence of human reliability programs. In some countries, the presence of such programs would be prescribed by law, while others might not have such regulations. In the latter case, it will not considered a violation for a non-vetted person to receive access to sensitive nuclear material. Thus, the answer to the definition of a violation will to a large extent depend on how one regards the issue of security, which is largely conditioned by the environment in which one is living.

In addition to the organizational or individual perception of violation or wrongdoing, the cultural setting can provide its own influence on the understanding of the term. In that regard, Miceli et al. [14] posited that definition of what constitutes wrongfulness may vary from country to country. For example, giving valuable presents or money in some countries is considered as a cost of doing business [14] or as act of 'gratitude', whereas in other cultures this is classified as bribery and is totally unacceptable.

In our case, there was not any clear pattern revealed in terms of a nationality-based preference for definitions of a violation. The replies were distributed more or less equally between different categories of answers by representatives of different countries. However, we would suggest listing all answers that people provided in the 'Other' section, where some interesting observations can be made. One respondent (USA) claimed that '*the existence of*

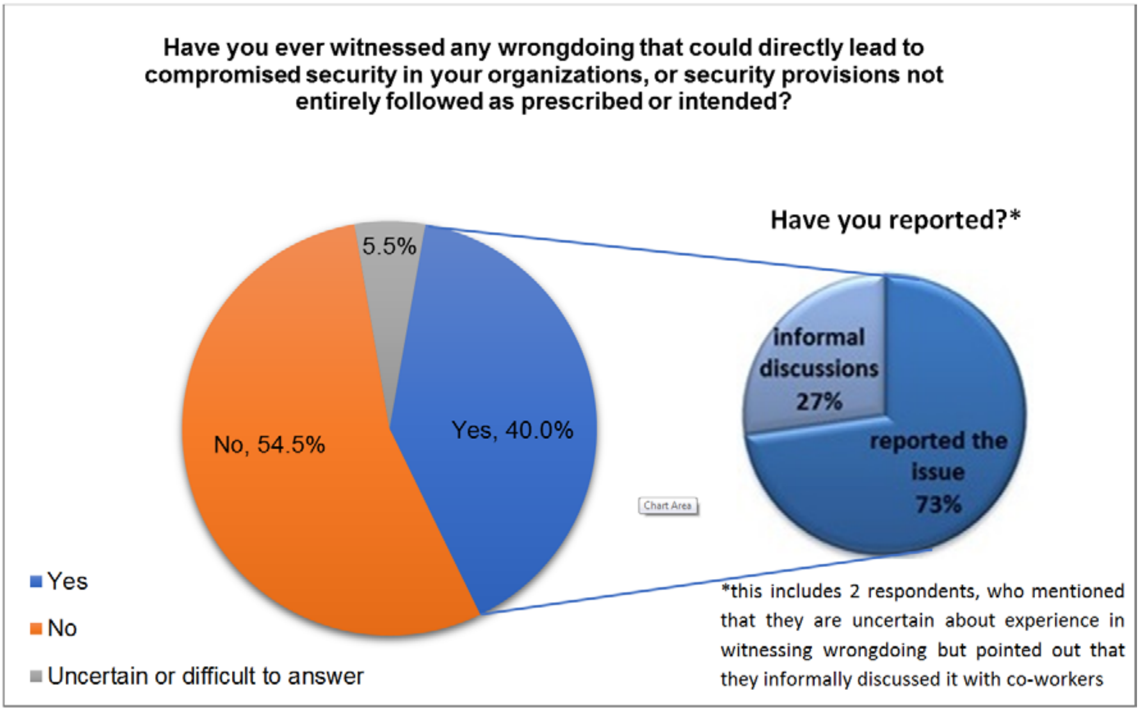


Figure 8: Experience in witnessing any wrongdoing and reporting

loopholes is not the violation, but the intentional exploitation of those loopholes. There will always be gaps to be exploited.' Another person (Ukraine) believed that '*no violations exist*', probably meaning that he has not yet experienced them in his practice. Some other answers pointed to drawbacks and weaknesses in security procedures and policies such as: '*limited funding*' (Ukraine), '*shortcomings in the legislation at the current stage, lack of highly professional staff, lack of funding*' (Ukraine), '*uncontrolled territory of the border*' (Moldova), '*blurred responsibilities among controlling organs, human and technical factors*' (Ukraine).

4.2 Experience with witnessing wrongdoing and reporting on it; the underlying reasons for reporting

It is hard to deny that whistle-blowing is a challenging and risky enterprise. A person who witnesses wrongdoing can raise an issue about an organizational problem, foster a solution to it, or vice versa, disrupt the functioning of legitimate activities [25] if, for example, allegations are not well-grounded. We asked our respondents about their experience in observing wrongdoing or security provisions not followed as prescribed or intended and whether this observation spurred them to report. The results of the survey on these questions are presented in Figure **8**.

The chart depicts that the majority (54.5%) postulated that they have not experienced witnessing any wrongdoing or neglect with regard to security procedures in their organizations. Forty percent of the respondents admitted that they have been faced with a wrongdoing or improperly followed security procedures, while a few (5.5%) expressed difficulty or uncertainty in answering such a question.

All those who answered positively to the question about witnessing a wrongdoing stated that somehow they called attention to the issue, either through informal discussion (27% - this includes two answers of those who were uncertain about whether they saw or not the wrongdoing etc.) or reporting (73%). Differentiating between two types of actions (i.e. reporting and informal discussions) is commonplace in the academic world, where the latter (informal discussions) is not equated to whistle-blowing. Miceli and Near [26], well established researchers in the area of whistle-blowing who insist that discussing informally with co-workers or family members is not reporting, build their argument on the fact that only discussions with those who might influence or affect the situation (i.e. to bring changes) constitute whistle-blowing.

To operationalize further the motives of those who have reported on alleged wrongdoing etc., we asked our respondents to select all factors from among those listed in the survey that have motivated them to report. The scale of responses is depicted in Figure **9**. Amongst all motives, one that was selected a substantial amount of times (43.5%) was the reason that '*Security is everybody's responsibility and I feel obliged to report*'. Thirteen percent of employees contended that '*A negligent attitude towards one's duties is detrimental and I did not want to work with people who do not align themselves with organizational standards*'. The same degree of response (6.5% each) was given to such reasons as '*The person in question might have had more detrimental motives in mind' and 'I was particularly worried about the risk of terrorism against my country or organization*', which indicates an acute

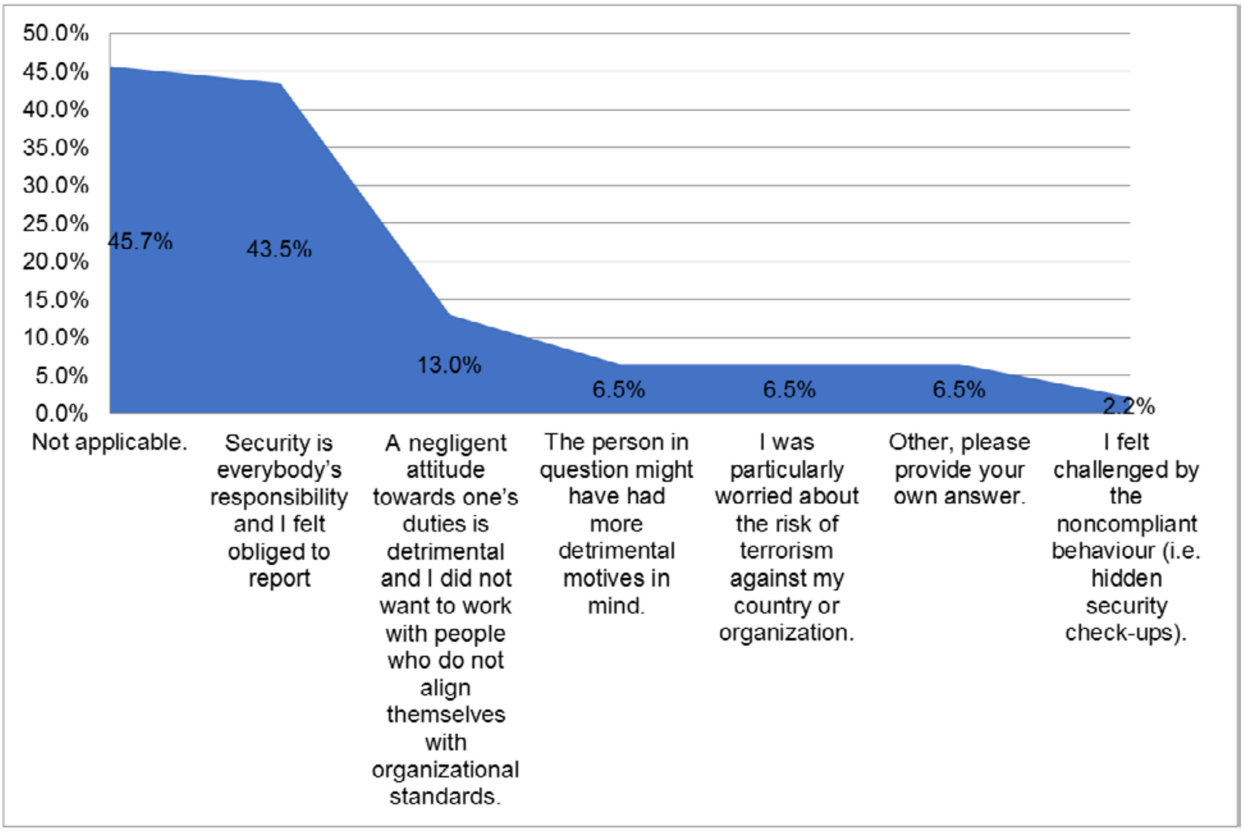


Figure 9: Factors that motivated reporting

perception of threat that potentially could concern nuclear or radiological materials handled by organizations, where respondents who have chosen these types of answer work.

Those who answered this question had a choice to provide their own explanations as well. We provide below the list of responses in the 'Other' section (6.5% selected this option) for the attention of the reader.

These '*Other*' responses include such statements:, '*I was part of a project which provided material control and protection oversight and we were responsible for monitoring, reviewing and reporting the results of our findings to those facilities where we were engaged*' (USA), '*Violations usually dealt with wrong exploitation of equipment and my duties concern the arrangement of uninterrupted working conditions*' (Ukraine), '*There is no such thing as not important issues in security; wishful thinking is the biggest enemy*' (Ukraine). One person stated that '*To report about violations is among my professional duties*' (Ukraine).

In that regard, there have been some discussions in the literature about whether to recognize reports that are considered to be part of a job description as whistle-blowing or not. An interesting observation was made by Dozier and Miceli [25] who posited that even if job descriptions require uncovering of violations, '*enthusiastic pursuit of this goal may not be rewarded in the organization*'; therefore, an individual may sometimes come into dissent with established 'organizational norms' when reporting misconduct. In this context, an exemplifying case is that of Richard Levernier, a nuclear security professional with more than 20 years of work experience, who reportedly pointed out that the possibility that suicide terrorists would not need to exit from nuclear facilities was overlooked by contingency planning scenarios [27]. After reporting as part of his job on weaknesses in security systems at nuclear power plants, he allegedly was reassigned to administrative work [27]. Numerous cases are described in the literature about employees going public when the organization fails or is unwilling to correct wrongdoing; in some cases, where they reveal weaknesses in an organization's system, they are faced with retaliation.

We also would like to indicate some factors that have influenced some people who participated in our survey not to report the issue or an alleged wrongdoing but to discuss informally with co-workers. In particular, they indicated that: '*Minor violations are better dealt with internally...*' (France) – this probably means that the person discussed the issue internally in contrast to external reporting, '*Knowledge that there is no solution for correcting a situation or it will require insurmountable financial and human resources from my organization*' (Ukraine), '*I am not in a position to personally witness security violations. As a researcher, I only learn of such issues after they've happened*' (USA), '*It was not direct wrongdoing, but rather lack of strong security culture/awareness and/or*

Figure 10: Drivers for reporting (determinants of intent to whistle-blow)

a compliance-based culture as opposed to a principle-based culture. Nothing really reportable' (USA), '*The issue was minor and could not affect security within the organization*' (Ukraine (5 persons), USA (1 person)), '*Belief that no action will be taken if I report*' (Ukraine). Some of the respondents from Ukraine just referred to a particular requirement in their national legislation (the piece of legislation mentioned by them concerns coordinating co-operation between governmental bodies in the event of detection of sources of ionizing radiation in an unauthorized possession, rather than speaking about reporting procedures inside an organization).

4.3 Drivers for reporting (determinants of intent to report)

We asked respondents (hypothetically if they were to observe noncompliance with a security requirement) to rank in the order of importance factors that might influence their decision to report. Although some of the rankings were missing from some of the applications, the aggregate picture looks as follows (see Figure **10**).

On average, for most of the people, ***certainty that noncompliance took place*** was the most significant determinant of an intention to report. Rationally, it can be explained by precautionary measures taken before denouncing something that may not correspond to reality, thus generating a 'false alarm'. Hence, most of the respondents are restrained from ungrounded accusations. Study of Miceli and Near [28] illustrated that '*convincing evidence that noncompliance took place*' can affect whistle-blowing behavior. As for future research, it might be interesting to study how certainty about the violation or wrongdoing is formed. In some cases, disconformity or divergence from security procedures might be clearly observed and wrongdoing might be apparent; in others the activity might be questionable; moreover, as Gundlach et al. [29] noted, wrongdoer manipulation tactics (like, 'false apologies' etc.) might reduce certainty and influence a decision to blow a whistle.

The ***significance of the issue (seriousness of the violation)*** has a strong positive relationship with the likelihood of reporting about a violation or wrongdoing. This was confirmed in our study, where this factor holds the second position in order of importance for its potential of triggering reporting behavior. In accordance with that, the results obtained in a study of U.S. federal employees showed that salience of the wrongdoing (which means that it was '*either very serious or very frequent*') has a strong influence on the likelihood of reporting by those who witness it and who hold the evidence [30]. As for the definition of seriousness of a wrongdoing, intriguingly, scientists have found that individuals on average '*perceive physically harmful acts as more serious than financial wrongdoing*' since the effects of the former usually can have a direct link to the risk posed to human health [31]. When the issue seems less serious, Keenan [32] posited that organizational propensity, which includes the amount of encouragement to report, will play a decisive role in whistle-blowing. Of course, it will also depend on how an organization tolerates wrongdoing, if, as Miceli and Near [26] noticed, it '*doesn't discourage wrongdoing* [, it] *would probably also not discourage retaliation*', which, one would assume, will increase the expected costs of reporting.

If observing a violation, there is reason for one to suspect ***bad intentions in a wrongdoer,*** this may foster a decision to report, but here we should remember the crucial role of certainty that non-compliance took place and the degree of seriousness of the violation. Barnett, Bass and Brown [33] studied how our own ethical judgement influences a decision to report on peers. Therefore, the moral standards of an observing employee and other individual variables influence the whistle-blowing process; however, we do not explore them in our study. A lot has been done in that field by other authors [see 33, 34, 35]. Our findings, which showed that certainty about wrongdoing taking place, its seriousness and the ***systemic nature of non-compliance*** of a wrongdoer are the most powerful triggering factors of whistle-blowing behavior in organizations that handle nuclear or radiological materials, are, in general, in line with the results of investigations carried out in other professional networks [28].

Assurances that the issue will be investigated in a fair manner, as well as that actions will be taken after reporting, were quite important (at fourth place) to the employees who took part in our survey. Here we would like to draw upon an example from Sandia National Laboratories to demonstrate a lack of fair investigation after breaches in security were reported by an employee of an organization that deals with nuclear materials. When Shawn Carpenter informed his superiors at the Sandia lab about systematic cyber espionage on major U.S. defense and military government agencies and their contractors, he was faced with an order to keep this secret to himself, since the computers attacked did not belong to the organization in question but to other governmental bodies and literally were assumed to be other persons' business [36]. 'Disobedience' and the subsequent external report of Carpenter resulted in his security clearance being revoked and termination of employment [36]. Evidence of reluctance from the side of organizations to investigate the issues raised by concerned employees can easily dissuade some from internal reporting or instigate them to public disclosure of violations. In contrast to that, perceived higher levels of organizational justice are positively associated with internal whistle-blowing behavior [37]. Hence, one may conclude, if the organization is not trusted, and is believed not to treat an issue in a fair manner, this may provoke silence even about salient violations or, conversely, disclosure of acts to the media or other external parties.

Experience and knowledge about reporting mechanisms influence reporting behavior [22]. This applies when they exist and employees are familiar with them, but what if no such special reporting mechanisms are established to deal with nuclear security in an organization? To receive an answer to this question, we asked our respondents to assign a ranking criterion for ***availability of complaint channels (adequate reporting mechanisms) such as designated hot line, web-page, etc.,*** based on the role it would play in motivating them to report an observed alleged wrongdoing. Our respondents placed it in the top five of the factors (out of ten). In our view, this is quite high in the ranking, which agrees with the statement that '*open-door policies, telephone "hotlines" and formal "whistle-blowing procedures" are* […] *likely to have a strong influence on individuals' decision whether to report perceived wrongdoing*' [33]. In the study by Glynn and Bunn [2] on the casino and pharmaceutical industries, they provided an example that in a number of casinos, anonymous tip-lines were an effective mechanism to enhance a security program; this could be borrowed for the nuclear security field. Establishing international 24/7 toll-free hotlines is becoming commonplace in some multinational corporations for establishing contact with a whistle-blower or those seeking advice regarding the reporting procedures to be followed; the latter may, if desired, remain anonymous [14]. Setting up such internal communication channels, as Barnett [38] contends, '*may increase the likelihood that employees discuss such concerns internally*'.

Next in the ranking of the determinants of reporting behavior, respondents of the survey put ***certainty that actions will be taken after reporting.*** A 'sleeping guard' case, as we call it, can serve as an antagonism of what a person who decides to report expects from an organization. When an employee, Kerry Beal, discovered that his colleagues in the security team at the Peach Bottom nuclear power plant in Pennsylvania took '*regular naps in what they called 'the ready room'*', he reported to supervisors, who allegedly told him '*to be a team player*' [39]. Resorting to the regional office of the Nuclear Regulatory Commission also did not bring the anticipated relief, since the plant owner to whom the issue was transferred '*said it found no evidence of guards* [being] *asleep on the job*' and the matter was considered concluded [39]. Obviously, transferring the issue to the organization, where security was not upheld, carried a risk of the issue being covered up and no corrective actions taken. Management of organizations in the nuclear and radiological field should bear in mind that such an indifferent attitude towards reporting is a manifestation of lax nuclear security culture and not something a vigilant employee who conveys a concern would expect to exist within his/her employer. Miceli et al. [14] warned that an unwelcome attitude from managers frequently deters employees from speaking up about observed wrongdoing; '*they believe nothing can or will be done to correct the problems; and* [...] *these beliefs are often well-founded*'.

Certainty that the organization has the resources to correct misconduct was considered to be a factor that might influence the reporting behavior of the employees in organizations that handle nuclear or radiological materials. As per the survey results, it occupied a somewhat intermediate position in the ranking of importance. Shockingly, as Bunn et al. [40] noted, if an organization is constrained financially, it might discount or even punish employees who try to enunciate security concerns.

Interestingly, by putting ***assurances in anonymity and confidentiality to exclude retaliation*** lower in ranking, people expressed less confidence that the presence of these measures would encourage them to blow the whistle. This might lend support to the theory of public service motivation (which will be discussed later in our paper); according to which individuals in public service (and most of the organizations that handle nuclear or radiological materials in the countries we examined were government-owned) are conducive to altruistic, public service motives also associated with self-sacrifice [20].

In addition to that, Brewer and Selden [34] posited that, in general, '*whistle blowers are probably less concerned about job security*'. Furthermore, a significant body of empirical research proved that overall retaliation will not suppress whistle blowing [34, 38]. Although the fear of retaliation does not necessarily dissuade an individual from the decision to blow the whistle, it may instigate an employee, as Caillier [20] asserts, to blow the whistle externally where he or she might hope to find a refuge from punitive measures by an organization. Nevertheless, assurances of anonymity and confidentiality may still encourage some employees to communicate their concerns internally and thus, we believe, should be carefully considered in organizational policies.

Finally, last in the ranking was ***the certainty that direct complaint to the noncompliant person or a peer will not correct the situation.*** Sometimes if a person approaches a wrongdoer, the latter might acknowledge an act of noncompliance '*by using excuses or justifications*', showing that it was rather an exception than a usual practice [29]; however, based on stories of professionals in the radiological sphere, those who raise the issue are quite frequently mocked or even threatened by a wrongdoer.

4.4 The ideal recipient of the report

Profound studies have been conducted on the variables that influence what path for reporting, internal or external, an individual who observed a wrongdoing will choose [see 4, 15, 16, 17, 22, 28, 37, 38, 41, 42]. This question can be inextricably linked to organizational factors such as perceived trust in an organization etc. On the macro-level, this can be restricted by the reporting mechanisms prescribed by law in a specific country. For example, data from the literature indicate that '*the UK legislation requires internal reporting in most circumstances. Australian and the large majority of US statutes favor external reports*' [14]. Therefore, employer-employee confidentiality plays a greater role in the UK, than in the USA or Australia [14]. The issue is complicated, however, by the specifics of legislative provisions at the state level in the USA, where some states '*require or encourage internal reporting before the whistle-blower goes outside the organization*' [17].

The question on the 'right recipient' of the report can be not only a reason for a collision at the workplace but can even escalate to the courtroom. A case that happened in the Los Alamos Laboratory illustrates how publicizing security and safety concerns allegedly led to retaliation against a whistle-blower. This person reported on the perceived lax security in the lab with regard to access to classified information '*on the timing, destination and security arrangements for transport of nuclear-weapons materials to the laboratory*' by uncleared employees [43]. After the perceived failure of Los Alamos managers to correct the '*systemic problems*', Gutierrez, the whistle-blower in question, decided to go public and reported the issue '*to federal lawmakers, to a nuclear watchdog group suing the laboratory and to three New Mexico newspapers*' [43]. Despite the fact that the court ruled that federal law, i.e. the Energy Reorganization Act, '*supersede*[s] *Los Alamos policies against lab workers having unauthorized communication with government officials and the media*', all Los Alamos lab employees were 'reminded' after the ruling on the prohibition on communicating with lawmakers '*on lab issues* [which] *could be construed as lobbying or could otherwise harm the lab*' [43].

Interested in whom the respondents would trust to accept their complaint, we asked them '***If you were to report security non-compliances, who is the ideal recipient of the report? Please list all options in the order of preferences from 1 to 6***'. By asking such a question, one could assess the reporting preferences and ultimately predict employees' behavior. Based on the answers we received, some tendencies might be discerned (see Figure **11**).

The responses indicate that in nuclear or radiological organizations, supervisors are looked upon most favorably as the recipients of the claim. This is followed in the list of preferences by the head of the organization or a special control body in the organization. This data coincides in findings with other research where informants refer the issue to the immediate supervisor first [44]. King [44] explains this by both, 1) established reporting channels within the organization and 2) relational distance between employees and upper management who might not be aware of the specifics of the problem. Overall, the results of our survey show the clear preference to contain an issue within the organization, as the first two preferred reporting channels belong to the organization. Then are followed by a specialized governmental body – 3rd place.

Figure 11: Preferred recipient of the report

Regarding a specialized governmental body one would mean a governmental organ intended to oversee security and receive complains on its violations. In Ukraine, for example, the National Anti-Corruption Bureau of Ukraine or NABU was established in 2014 as a law enforcement anti-corruption agency, which investigates corruption in Ukraine and prepares cases for prosecution. Therefore, a person in Ukraine can (anonymously if one so wishes) report to NABU an event where high officials abuse or misuse their authorities for gain or personal preferences. An international structure created under IAEA took the 4th place (we hypothesized about the existence of such to see if the responders would prefer this channel); the 5th place is occupied by an independent NGO or other public organization, and the last in the list of preferences is a private organization hired to serve the role of investigating non-compliances.

Miceli and Near [45], relying on previous research, condensed information on possible motives of internal reporting to two most powerful factors that can be explained by 'deviance' or 'differential association' theories. According to the first theory, choosing a supervisor or structure within an organization as the primary recipient for a complaint is explained by minimized risk associated with reporting to internal channels in comparison to an external one [45]. The latter theory accounts for '*norms of loyalty*', meaning that the climate prevalent in organizations '*is generally antagonistic toward exposing misconduct* [externally]' [45].

In that regard, Barnett [38] concluded that the consequences of external whistle-blowing are more severe '*both for organizations and whistle-blowers*'. For the former, it inflicts significant reputational damage, while for the latter it imposes risks of retaliation due to public enunciation of the violation [38]. In general, internal reporting hinders the employer-employee relationship the least and provides the opportunity for earlier correction of violations [30].

From an organizational policy perspective, Lavena [42] postulated that a supportive environment within an organization, where a supervisor is trusted by employees, contributes to a decrease in external reporting. If an organization does not tolerate dissent, thereby suppressing internal disclosure, whistle-blowers might speak out and report externally [17]. The finding that organizational size might be mediating the reporting paths - the bigger the organization, the greater the chances of external reporting, because, as suggested by Barnett [38], '*bureaucracies do not foster ideal environments for effective upward communication*' - should be considered in large research and development organizations, and in industrial settings dealing with nuclear and other radiological materials.

Summarizing what we have discussed before, there is a definite value in establishing clear reporting policies within an organization. At the same time, such policies should not be restrictive; instead of instilling fear in employees for escalating a complaint into a public domain, managers should treat security reports seriously and build a participatory work environment, characterized by solidarity, engagement and openness.

4.5 Attitude regarding those who report

By replying to the question, 'What, in your opinion, best describes to the profile of people who report security non-compliances?' 83% of the respondents have chosen an answer that says **'They are everyday people who really care about security in their organization and the nuclear community as a whole'** (see Figure **12**). Therefore, most of the respondents to our survey do not consider whistle-blowing to be a deviant behavior. Here we should admit that, although in our analysis here we use the term whistle-blowing, in the survey we have purposefully decided to avoid using this word and used the term report/reporting instead. This is partly because of the dramatism that could surround the term. Another reason is the

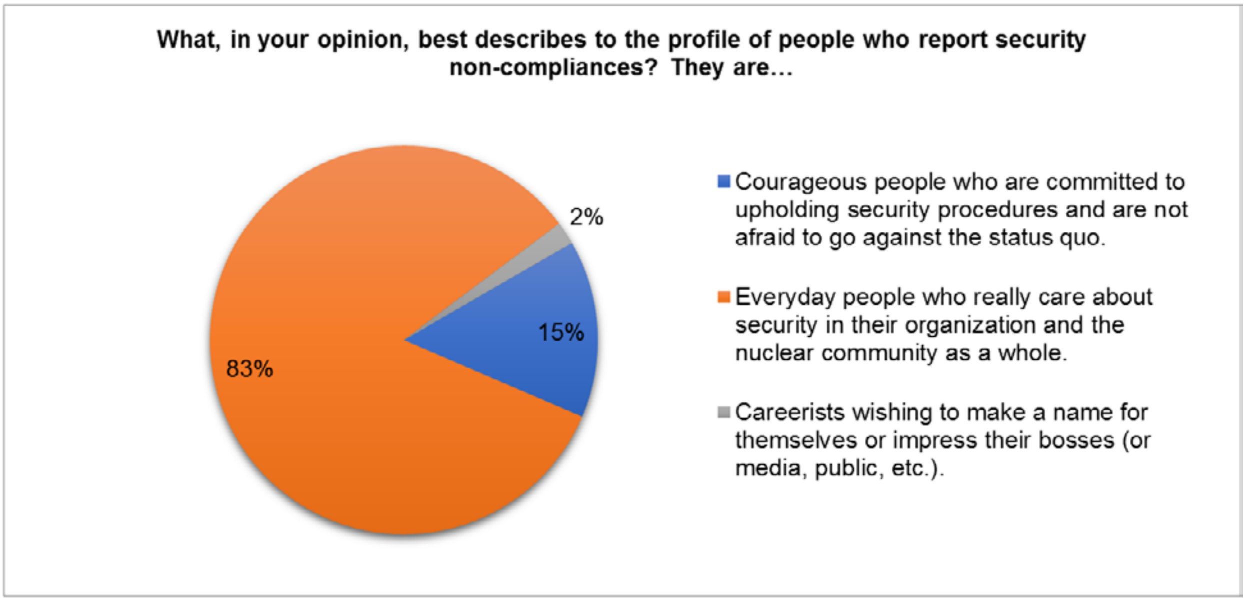


Figure 12: Attitude to those who report

controversy that exists in some countries with regard to translation of the word whistle-blowing and its ambivalent, often negative, meaning, which will be discussed in greater detail below. In that context, we used the word reporting, which is more of a neutral term and does not bring negative connotations. Thus, it allowed respondents to answer questions without being influenced by any painful historic narrative and to assess the act of reporting rather than an attitude towards the word.

Seeing reporting or whistle-blowing as a normal, not extraordinary behavior seems to be in line with one strand of argument in the literature, according to which scholars assert that '*whistle-blowers have not been found to be especially "moral" people, "religious" people, "political" people, or "socially responsible" people*' but ordinary people who decided to make it their business and to act '*regardless of their own good*' [46]. Another possible explanation for the fact that a majority of the respondents see whistle-blowing as an ordinary act could deal with a self-reporting bias, as noted repeatedly in the literature about whistle-blowing. Thus, respondents do not reveal their true feelings if asked directly on the subject but try to choose the answer which might seem rational to them or socially acceptable [34]. Therefore, in line with proclaimed values of being observant, the respondents might have chosen the answer that seemed right to them and would correspond to a strong security culture. To minimize the effect of self-bias, we would recommend asking additional and probing questions, including requests to the respondents to assess actions described in short vignettes. Doing this, however, would require large investments of time and not every employee in the nuclear sphere would consider dedicating his or her working time to filling out such a survey.

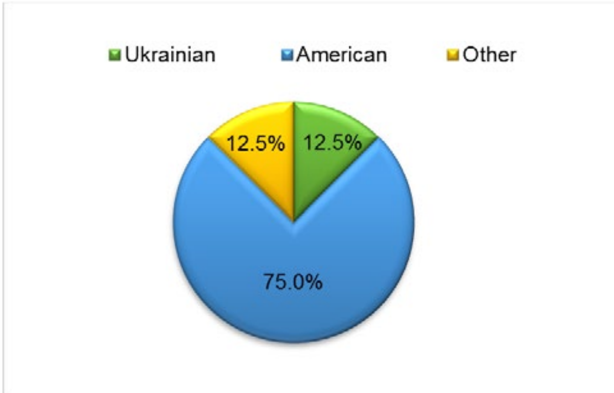


Figure 13: Nationalities of respondents who characterize people who report as "courageous people who are committed to upholding security procedures and are not afraid to go against the status quo"

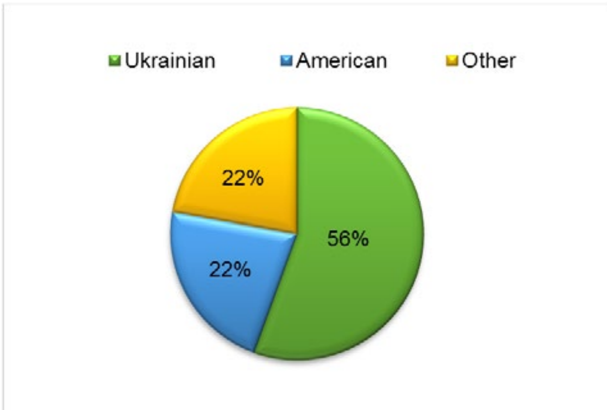


Figure 14: Nationalities of respondents who characterize people who report as "everyday people who really care about security in their organization and the nuclear community as a whole"

Interestingly, however, some eight respondents (15%) to the survey, six of which (constituting 75 percent) were Americans, one Ukrainian and one Lebanese, held the view that reporting security non-compliances is a characteristic that is attributed to ***'Courageous people who are committed to upholding security procedures and are not afraid to go against the status quo'*** (see Figure 13). Some possible explanations for this phenomenon might be as follows. First, we assume that the respondents might be influenced by the number of media coverages about whistle-blowers in the nuclear industry who became disadvantaged after reporting incompliances or revealing nuclear safety or security weaknesses. Due to the fact that the media, especially in the USA, had covered a lot of such dramatic and often unfortunate events (albeit retaliation is not as frequent as might seem from the media [26]), this might contribute to a feeling of self-sacrifice and courage that whistle-blowers allegedly need to have in order to challenge the existing situation. Thus, U.S. nationals and residents might have been more aware of retaliation against 'dissident' employees than, for example, Ukrainians, where media reports have not been that frequent, neither about whistle-blowing nor on the retaliation. Secondly, some court cases in the USA have exemplified that '*the law in the United States provides inadequate protection to whistle-blowers and gives organizations too little incentive to take corrective action, providing scant reason to believe that whistle-blowers will succeed in their quest to get wrongdoing stopped (e.g., Dworkin & Near, 1997; Miceli et al., 1999)*' [16]. Finally, some of the respondents had security functions as role-prescribed, i.e. security specialist (see Figure **1**), and most of them happened to be Ukrainians; therefore they (Ukrainians) might have considered reporting of violations as a duty rather than an act of courage or disloyalty (see Figure 14**.**).

One Ukrainian, however, did not hold a high opinion of whistle-blowers, as was shown by his choosing an answer that describes people who report violations as ***'Careerists wishing to make a name for themselves or impress their bosses (or media, public, etc.)'***. For him, reporting is not a negative deviant behavior but rather an unethical one which, as Appelbaum et al. [47] explain, '*deals with the breaking of societal rules*'. One might assume that a Soviet past, with its specific usage of reporting as an act to suppress civil disobedience and gain benefits from a regime, might have stigmatized reporting procedures in post-Soviet countries. This is especially true if one considers that whistle-blowing might have a negative connotation (it is translated in the Russian-speaking post-Soviet world as 'доносительство') related to reporting to the NKVD (abbreviated from *Narodnyi Komissariat Vnutrennikh Del*, meaning The People's Commissariat for Internal Affairs) – a ministry of the Soviet government responsible for security and law enforcement and which is associated with repression. Thus, receiving such an answer to the survey from a citizen of a country which was part of the Soviet empire is not something unusual or unexpected. In general, though, there might be in any country whistle-blowers who may seek '*self-aggrandizement and publicity*' [25].

To summarize, despite the overall optimistic picture regarding the prevalent attitude towards whistle-blowing as a normal, 'business-as-usual' act, we would like to point to the disturbing number of incidences when reporting was not taken seriously and those where retaliation did take place. Therefore, agreeing with the role of training on whistle-blowing policies, ethics and organizational procedures [32], we believe a set of generic templates for communications as well as dedicated training sessions will encourage reporting in the nuclear and radiological sphere and help minimize loopholes in security.

4.6 What is needed by employees to follow security procedures in their organizations

Figure **15** reports the things that subjects of the study indicated were the most important in helping them to follow security procedures in their organizations. We asked respondents to rank six factors in the order of importance. The mean results indicate that **additional training in security procedures and clear guidance from senior management to follow the rules and management's own adherence to them were the most appreciated factors**, getting on average 2.71 and 2.74 rankings, respectively.

An example that happened in Lithuania in 1992 might be brought to the attention of the reader to showcase the importance of training on security policies. A computer programmer named Oleg Savchuk, who placed a computer virus, was sentenced in court for trying to sabotage a nuclear reactor at the Ignalina Nuclear Power Plant [48]. Allegedly, this worker was trying, in such a way, '*to call attention to a weakness in the plant's control system and then may have hoped to be rewarded for his service*' [49]. Though his true motives remain unclear, the attempt (even though maybe with a benign final purpose) to damage the facility is apparent, which clearly demonstrates the need for training, both in security procedures and whistle-blowing practices. We suggest that the latter not only will help to improve security in organizations but also may reduce the number of non-legitimate claims and deter frivolous campaigns.

In our survey, we pointed to the need for practicing skills during exercises. Glynn and Bunn [2] also suggested that brainstorming on possible diversion scenarios and responding counter efforts of security personnel, which have proven to be beneficial in the pharmaceutical business, could be used to simulate security breaches in the nuclear industry. Although a lot of similar exercises are currently run in organizations that handle nuclear or radiological

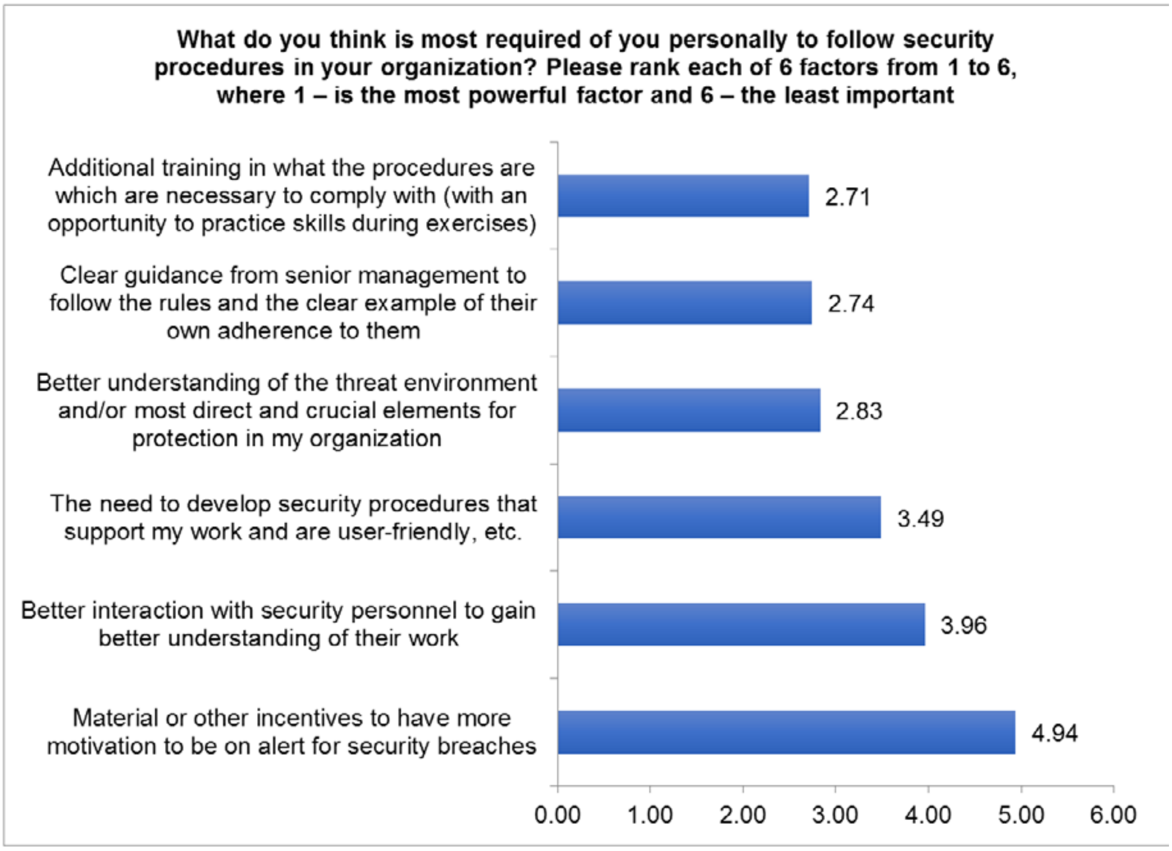


Figure 15: Things required to follow procedures

materials, the benefits of brainstorming on potential vulnerabilities and stimulation of vigilance are greater than simple usage of preset security scenarios. This will also fit into the need to give employees a ***better understanding of the threat environment and/or most direct and crucial elements for protection in their organization*** (placed at 2.84 in the ranking). Although the design basis threat is a classified document, secrecy should not hamper sharing some of the information with employees; it might lead to greater engagement from the side of workers in terms of protecting critical elements and complying with procedures. Also, engaging staff members in implementing the IAEA's self-assessment methodology of nuclear security culture [50] serves as an additional learning experience, enabling workers to apply generic principles to specific needs of their organizations' security regime.

The next factor in the ranking scale, with a score of 3.49, was ***the need to develop security procedures that support my work and are user-friendly, etc.*** One may suggest that this is especially relevant in the cyber-dimension of nuclear security. There was a case described by the media, where Edward McCallum, the former director of safeguards and security programs of the U.S. Department of Energy, was cited saying that many laboratories that deal with nuclear materials or perform research on them resist introduction of new network security architectures and procedures, since they perceive them as '*unnecessarily expensive and a hindrance to operations*' [51]. This attitude shows that a special effort should be made to explain the importance of newly-established procedures and increase their user-friendliness.

With regard to the ranking in our survey, the fifth place was taken by the **better interaction with security personnel to gain better understanding of their work.** This can be closely related to training activities and empowerment actions described above; however, it also implies narrowing of the communication gap between the security unit and other personnel. This is because a person who is not part of the facility's security contingent may think security is someone else's responsibility and that security successes and failures have little to do with anything that person does or fails to do. In this regard, a study to determine possible implications for whistle blowing in relation to the existence of professional subcultures will be useful.

Unlike Miceli and Near [22], we did not receive support for a statement that material incentives would encourage whistle-blowing. Our participants downplayed the role of ***financial rewards in encouraging them to follow security procedures and stay on alert when they are breached***. Financially rewarding security vigilance could have a negative effect, first, because people who have observed the violation may, *vice versa*, be discouraged by the financial incentives (in order not to be regarded as '*hunters for financial gain*', which might contradict their ethical principles and sometimes even cause ostracism by their colleagues); second, there also can be those who may misuse the system and reap financial gain by making illegitimate claims. With closer examination, Caillier [20] finds an explanation to the phenomenon of disregard for monetary benefits in public service motivation. The concept of public service motivation takes its roots from the *'special calling'* '*to pursue the common good and further the public interest*' [34]. Whether or not a low regard for monetary rewards is a defining feature of public-service motivation is not entirely clear, but it is definitely trivial, as Brewer and Selden [34] showed, among public employees.

5. Conclusions and recommendations

The issue of reporting on security breaches in the nuclear and radiological sector has been overlooked for a long time; the scarce discussions on the topic have usually been limited to an acknowledgement of the problem of operationalization of fair reporting in an organization's policies and lack of approaches in ensuring its effectiveness. At the same time, instances of retaliation against whistleblowers who report security incompliances or raise concerns about inappropriate security measures in the nuclear field are disturbingly frequent, as was shown by some anecdotal evidence in our work. Dworkin and Near [30] reasonably admit that '*the problem for organizations is not how to avoid whistle-blowing, but how to diminish its negative consequences and to maximize its positive aspects*'. We find that this statement extends to the nuclear and radiological sphere. Opportunities enclosed in reporting for boosting a strong nuclear security culture are huge, but so are the challenges. The task is to unearth drivers of reporting so that one can build on them, expose vulnerabilities and work on the elimination of impediments, promote raising good-faith concerns and decrease adverse factors associated with whistle-blowing. With the pursuit of this current study, we have contributed to an appreciation of the importance of this task and, hopefully, have provided some findings to be used further.

We have not encountered step-by-step guidelines on establishing reporting mechanisms on security matters in organizations that handle nuclear and radiological materials; therefore, one may conclude that they are not very common and thus need to be explored. Therefore, on the state-level, we would like to suggest looking at the areas where such exist. For example, in the chemical industry, Chemical Facility Anti-Terrorism Standards (CFATS) were adopted with the aim of improving security at high-risk chemical facilities in the USA. The recent report of the Government Accountability Office on *Critical Infrastructure Protection. Improvements Needed for DHS's Chemical Facility Whistleblower Report Process* [see 7] provides some useful suggestions on the regulatory level for fostering reporting procedures. We suggest familiarization with the lessons drawn from the chemical industry so as to avoid repetition of the same mistakes in the nuclear and radiological sphere when operating a reporting mechanism. In a long run, review of legislation in countries as well as analysis of prospects of passage in the countries where it is currently absent is warranted. Some work on the state level (for instance, for the USA) has been done by researchers [see 30], though it may require an update due to the development of legislation and subsequent changes since the time of the research.

On the organizational level, an overview of how reporting on safety matters in nuclear and radiological organizations is implemented could be beneficial. For example, in the Brookhaven National Laboratory, USA, people are given the authority to stop an activity that he or she believes constitutes an imminent danger for the environment or health. Every newcomer to the lab is trained in the Stop Work Procedure. If the threat is not an urgent one, a 24-hour hotline operates for reporting safety concerns; the voice mailbox is checked by the operator at least twice a day.

At the management level, we find it useful to consult the work of Miceli et al. [14] *A Word To The Wise: How Managers And Policy-Makers Can Encourage Employees To Report Wrongdoing*. It contains some guidance on how to encourage reporting. Apparently, effective and efficient reporting procedures will require a climate and culture change within an organization [42]. A nuclear security culture coordinator (referenced in the IAEA draft guidance) should embark on these efforts, but alone little can be done; a greater commitment from the highest level of management to the security personnel in organizations and their appreciation of importance of reporting procedures are imperative for progress in this area. All of the survey participants indicated the value of clear guidance from management on security provisions and their adherence of the latter to the rules in general. Thus, reporting procedures should be known by the staff of organizations that deal with sensitive materials and transcribed into the *modus operandi* of the organization.

Our study has shown that professionals in the field will trust their supervisors to be the recipient of their concern. Consequently, supervisors have responsibilities for actions and treating disclosed concerns with due regard. If the issue shared demonstrates a risk to other operational units or structures, the information should be transferred in confidence to the right recipient; there is no place for a silo mentality in security matters, where stakes are so high due to the danger associated with the risk of unauthorized possession of materials, their misusage or sabotage etc.

The study has shown that reporting typically is treated as an ordinary behavior of normal employees who have a strong conscience for security. However, some of the respondents still feel that this is a risky undertaking, although one with honorable intentions, whereas a minority does not believe in the good intentions of those who report but rather see it as committing an act aimed at gaining career benefits for themselves. This gives a reason to suggest that education on whistle-blowing might be well-perceived by the professional community in the nuclear and radiological area and might help address the questions they might have and ease their concerns. This is especially needed so that morale in organizations is not compromised by ineffective, incomprehensive introduction of reporting.

Our study agrees that *'finding the right encouragements or inducements for whistle-blowers might be problematic and certainly will require long term, concerted effort'* [22]. To add to this, our study has shown a rather low regard of professionals in the nuclear and radiological sphere to material incentives; the relation of this phenomena to public-service motivation theory needs to be tested.

Findings also suggest that security problems must be regarded in a complex manner. Our participants indicated on numerous occasions that imperfect legislation, the level of financing, inappropriate division of labor and blurred responsibilities etc. can compromise security in organizations.

6. Acknowledgement

We thank the Brookhaven National Laboratory Nonproliferation and National Security Department staff and Department Chair Susan Pepper, for their welcome and support during preparation and implementation of this research, and their valuable comments to the survey. The U.S. Department of Energy provided funding for the research internship, for which we are very grateful. We wish also to thank the World Institute for Nuclear Security for posting a news item on their website about the survey and all colleagues who helped with distribution of the survey. Finally, our sincere gratitude goes to all those who filled out the survey, for the dedication of their time and insights shared.

7. References

[1] Khripunov, I., *Nuclear Renaissance and Security Culture.* IFANS Review, 2010. **18**(2): p. 95-120.

[2] Glynn, K.M. and M.G. Bunn, *Preventing Insider Theft: Lessons from the Casino and Pharmaceutical Industries*. 2013, Harvard Kennedy School of Government.

[3] IAEA, *Establishing a Code of Ethics for Nuclear Operating Organizations*, in *IAEA Nuclear Energy Series*. 2007, INTERNATIONAL ATOMIC ENERGY AGENCY: Vienna. p. 4.

[4] Near, J.P. and M.P. Miceli, *Organizational dissidence: The case of whistle-blowing.* Journal of Business Ethics, 1985. **4**(1): p. 1-16.

[5] Kohn, S.M. and T. Carpenter, *Nuclear Whistleblower Protection and the Scope of Protected Activity Under Section 210 of the Energy Reorganization Act.* Antioch LJ, 1986. **4**: p. 73.

[6] Jubb, P.B., *Whistleblowing: A Restrictive Definition and Interpretation.* Journal of Business Ethics, 1999. **21**(1): p. 77-94.

[7] GAO, *Critical Infrastructure Protection. Improvements Needed for DHS's Chemical Facility Whistleblower Report Process. Report to Congressional Committees*. 2016, United States Government Accountability Office: Washington, D.C.

[8] nis2016.org. *Joint Statement of the 2012 Seoul Nuclear Industry Summit*. 2016 30.11.2016]; Available from: http://nis2016.org/agenda/documents/documents-joint-statement-of-the-2012-seoul-nuclear-industry-summit/.

[9] nis2016.org. *Joint Statement of the 2016 Nuclear Industry Summit*. 2016 30.11.2016]; Available from: http://nis2016.org/agenda/documents/documents-nuclear-industry-summit-2016-joint-statement/.

[10] IAEA, *Objective and Essential Elements of a State's Nuclear Security Regime: nuclear security fundamentals*. 2013, International Atomic Energy Agency: Vienna. p. 15.

[11] IAEA, *Nuclear Security Culture. Implementing Guide*. 2008. p. 37.

[12] IAEA, *Establishing the Nuclear Security Infrastructure for a Nuclear Power Programme. Implementing Guide.* 2013, International Atomic Energy Agency: Vienna. p. 73.

[13] Bunn, M. *Incentives for Nuclear Security (Conference Paper)*. 2005 30.11.2016]; Conference Paper]. Available from: http://live.belfercenter.org/publication/12712/incentives_for_nuclear_security.html

[14] Miceli, M.P., J.P. Near, and T.M. Dworkin, *A word to the wise: How managers and policy-makers can encourage employees to report wrongdoing.* Journal of Business Ethics, 2009. **86**(3): p. 379-396.

[15] Ahmad, S.A., M. Smith, and Z. Ismail, *Internal Whistle-Blowing Intentions: A Study of Demographic and Individual Factors.* Journal of Modern Accounting and Auditing, 2012. **8**(11): p. 16-32.

[16] Miceli, M.P., *Whistle-Blowing Research and The Insider Lessons Learned and Yet to Be Learned.* Journal of Management Inquiry, 2004. **13**(4): p. 364-366.

[17] Dworkin, T.M. and M.S. Baucus, *Internal vs. external whistleblowers: A comparison of whistleblowering processes.* Journal of Business Ethics, 1998. **17**(12): p. 1281-1298.

[18] IAEA. *Resources for Women*. [cited 2016 05.12.2016]; Available from: https://www.iaea.org/about/employment/women.

[19] WiN. *Welcome to Women in Nuclear Global*. [cited 2016 05.12.2016]; Available from: http://www.win-global.org/.

[20] Caillier, J.G., *Public Service Motivation and Decisions to Report Wrongdoing in US Federal Agencies Is This Relationship Mediated by the Seriousness of the Wrongdoing.* The American Review of Public Administration, 2016: p. 1-22.

[21] Keenan, J.P., *Comparing Indian and American managers on whistleblowing.* Employee Responsibilities and Rights Journal, 2002. **14**(2-3): p. 79-89.

[22] Miceli, M.P. and J.P. Near, *The relationships among beliefs, organizational position, and whistle-blowing status: A discriminant analysis.* Academy of Management journal, 1984. **27**(4): p. 687-705.

[23] U.S.DOE, *Special Report. Inquiry into the Security Breach at the National Nuclear Security Administration's Y-12 National Security Complex*, in *DOE/IG-0868*, O.o.I.G.O.o.A.a. Inspections, Editor. 2012.

[24] Schlosser, E. *Break-In at Y-12. How a handful of pacifists exposed the vulnerability of America's weapons-grade uranium.* The New Yorker 09.03.2015 01.12.2016]; Available from: http://www.newyorker.com/magazine/2015/03/09/break-in-at-y-12.

[25] Dozier, J.B. and M.P. Miceli, *Potential predictors of whistle-blowing: A prosocial behavior perspective.* Academy of Management Review, 1985. **10**(4): p. 823-836.

[26] Miceli, M.P. and J.P. Near, *Relationships among value congruence, perceived victimization, and retaliation against whistle-blowers.* Journal of Management, 1994. **20**(4): p. 773-794.

[27] Martin, B., *Nuclear Power and Antiterrorism: Obscuring the Policy Contradictions 1.* Prometheus, 2007. **25**(1): p. 19-29.

[28] Miceli, M.P. and J.P. Near, *Characteristics of organizational climate and perceived wrongdoing associated with whistle-blowing decisions.* Personnel Psychology, 1985. **38**(3): p. 525-544.

[29] Gundlach, M.J., S.C. Douglas, and M.J. Martinko, *The decision to blow the whistle: A social information processing framework.* Academy of Management Review, 2003. **28**(1): p. 107-123.

[30] Dworkin, T.M. and J.P. Near, *Whistleblowing Statutes: Are They Working?* American Business Law Journal, 1987. **25**(2): p. 241-264.

[31] Greenberger, D.B., M.P. Miceli, and D.J. Cohen, *Oppositionists and group norms: The reciprocal influence of whistle-blowers and co-workers.* Journal of Business Ethics, 1987. **6**(7): p. 527-542.

[32] Keenan, J.P., *Blowing the whistle on less serious forms of fraud: A study of executives and managers.* Employee Responsibilities and Rights Journal, 2000. **12**(4): p. 199-217.

[33] Barnett, T., K. Bass, and G. Brown, *Religiosity, ethical ideology, and intentions to report a peer's wrongdoing.* Journal of Business Ethics, 1996. **15**(11): p. 1161-1174.

[34] Brewer, G.A. and S.C. Selden, *Whistle blowers in the federal civil service: New evidence of the public service ethic.* Journal of public administration research and theory, 1998. **8**(3): p. 413-440.

[35] Park, H., M.T. Rehg, and D. Lee, *The influence of Confucian ethics and collectivism on whistleblowing intentions: A study of South Korean public employees.* Journal of Business Ethics, 2005. **58**(4): p. 387-403.

[36] TMC-NEWS, *Sandia Hacker Gets $4 Million: Analyst Fired For FBI Contact*, in *Albuquerque Journal (NM) (KRT) Via Thomson Dialog NewsEdge*. 14.02.2007, TMC NEWS.

[37] Seifert, D.L., et al., *The influence of organizational justice on accountant whistleblowing.* Accounting, Organizations and Society, 2010. **35**(7): p. 707-717.

[38] Barnett, T., *A preliminary investigation of the relationship between selected organizational characteristics and external whistleblowing by employees.* Journal of Business Ethics, 1992. **11**(12): p. 949-959.

[39] Mufson, S., *Video of Sleeping Guards Shakes Nuclear Industry*, in *Washington Post* 04.01.2008 Washington Post

[40] Bunn, M., et al., *Preventing Nuclear Terrorism: Continuous Improvement or Dangerous Decline?* 2016.

[41] Cho, Y.J. and H.J. Song, *Determinants of Whistleblowing Within Government Agencies.* Public Personnel Management, 2015. **44**(4): p. 450-472.

[42] Lavena, C.F., *Whistle-blowing individual and organizational determinants of the decision to report*

wrongdoing in the federal government. The American Review of Public Administration, 2016. **46**(1): p. 113-136.

[43] Hoffman, I. *Whistleblower Wins Case Against Lab. Ruling Protects Rights Of Workers To Report Concerns*. The Albuquerque Journal 02.07.1999; Available from: http://www.state.nv.us/nucwaste/news/nn10114.htm.

[44] King, G., *The effects of interpersonal closeness and issue seriousness on blowing the whistle.* Journal of Business Communication, 1997. **34**(4): p. 419-436.

[45] Miceli, M.P. and J.P. Near, *What makes whistle-blowers effective? Three field studies.* Human Relations, 2002. **55**(4): p. 455-479.

[46] Contu, A., *Rationality and Relationality in the Process of Whistleblowing Recasting Whistleblowing Through Readings of Antigone.* Journal of Management Inquiry, 2014. **23**(4): p. 393-406.

[47] Appelbaum, S.H., G.D. Iaconi, and A. Matousek, *Positive and negative deviant workplace behaviors: causes, impacts, and solutions.* Corporate Governance: The international journal of business in society, 2007. **7**(5): p. 586-598.

[48] Potter, W.C., *Less Well Known Cases Of Nuclear Terrorism And Nuclear Diversion In Russia*. 20.08.1997, Center for Nonproliferation Studies Unpublished Paper.

[49] Ferguson, C.D. and W.C. Potter, *The four faces of nuclear terrorism*. 2005: Routledge.

[50] IAEA, *Self-assessment of Nuclear Security Culture in Facilities and Activities*. IAEA Nuclear Security Series. 2017, Vienna: INTERNATIONAL ATOMIC ENERGY AGENCY.

[51] Verton, D. *DOE clamps down on whistle-blower for security leaks*. 10.06.1999 01.12.2016]; Available from: https://fcw.com/articles/1999/06/10/doe-clamps-down-on-whistleblower-for-security-leaks.aspx.

Open Source Analysis in support to the identification of possible undeclared nuclear activities in a State

Frank V. Pabian[1], Guido Renda[2], Giacomo G.M. Cojazzi[2]

1 Los Alamos National Laboratory Fellow, Retired
2 European Commission, Joint Research Centre, Directorate G – Nuclear Safety and Security, Via E. Fermi 2749, Ispra (Va), Italy

Abstract:

The revelation of Iraq's clandestine nuclear weapons program in the 1990s first made clear the necessity of bringing new tools to bear in the implementation of IAEA Safeguards. The adoption in 1997 of the Additional Protocol by the IAEA Board of Governors paved the way to the introduction of the State Level Approach. The IAEA's quest for such increased transparency regarding nuclear-relevant activities continues to evolve with the adoption of additional sources of safeguards-relevant information. A wide variety of information is available through official publications, academic and technical journals, and other media from which to glean insights on not just the capabilities of a nation-state, but also the direction of research and development along the varied paths to nuclear proliferation. Open source Information can also be found via various blogs that have particular areas of interest relevant to treaty monitoring and verification organizations. Commercial satellite imagery has, since the turn of the new millennium, become an increasingly valuable open source for IAEA Safeguards purposes.

Open source information can play an important role in several aspects related to the implementation of the Non-Proliferation Treaty provisions. The paper will briefly introduce the IAEA State-Level Concept (SLC), the role of Open source information in this context and then focus on some methodological considerations related to the use of open source information to investigate the existence of potential undeclared nuclear activities in a State. Finally, the paper will show how the combination of heterogeneous open source and geospatial information can lead to sig. nificant, but otherwise unknown, information for nuclear safeguards applications. In particular, the paper will present an exemplary application, in which, following-up on a single report from Iranian news, a review of commercial satellite imagery (including that available cost-free via Google Earth©) has made possible the identification of the location of the facility now known as the "Pasmangoor Nuclear Waste Storage and Stabilization Facility" in Anarak, Iran, and a new near-by small (likely pilot scale) ore processing facility that was completed near a previously abandoned, but recently reactivated, mine. The mine is known to contain copper, nickel, cobalt, arsenic, and uranium, making the new facility potentially of safeguards relevance.

Keywords: IAEA State-Level Concept, Open Source Analysis, Non-Proliferation.

1. Introduction

The revelation of Iraq's clandestine nuclear weapons program in the 1990s first made clear the necessity of bringing new tools to bear in the implementation of IAEA Safeguards. That program "*exposed all too clearly the limitations of a safeguards system focused exclusively on declared nuclear material*" [1] and which was only focused upon declared nuclear sites. The adoption in 1997 of the Additional Protocol [2] by the IAEA Board of Governors helped "*to detect deviations from what might be expected in a peaceful nuclear program sooner*" [3] and paved the way to the introduction of the State-Level Concept (SLC). The IAEA's quest for such increased transparency continues to evolve with the adoption of additional sources of information. The Agency "*carries out a comprehensive evaluation of all available safeguards-relevant information – including data provided by the state, the results of the agency's in-field activities and its extensive collection of safeguards-relevant information from open sources (such as scientific publications, conference records, and commercially available satellite imagery) – looking for consistency with the state's declarations*". [4]

Commercial satellite imagery has, since the turn of the new millennium, become an increasingly valuable open source for IAEA Safeguards purposes. Moreover, "satellite imagery is used routinely to evaluate information provided by States on their nuclear activities and to plan inspections, visits to facilities to verify design information and to conduct complementary access under the Additional Protocol." [5]

Open source information can play an important role in supporting several aspects of the SLC [6], integrating and supplementing the information retrieved from States' declarations and infield verification activities. In particular, Open source information has the potential of being the main key source of information for identifying potential undeclared nuclear activities in a State, an area that historically proved to be the most challenging of the entire non-proliferation regime.

The paper will briefly introduce the IAEA State-Level Concept (SLC), the role of Open source information in this context and then focusses on some methodological considerations related to the use of open source information to investigate the existence of potential undeclared nuclear activities in a State. Finally, the paper will show how the combination of heterogeneous open source and geospatial information can lead to significant, but otherwise unknown, information for nuclear safeguards purposes. In particular, the paper will present an exemplary application, in which, following-up on a single report from Iranian news, a review of commercial satellite imagery (including that available cost-free via Google Earth) has made possible the identification of the location of the facility now known as the "Pasmangoor Nuclear Waste Storage and Stabilization Facility" in Anarak, Iran and a new near-by small (likely pilot scale) ore processing facility that was completed near a previously abandoned mine. The area is known to contain copper, nickel, cobalt, arsenic, and uranium.

2. The IAEA State-Level Concept

To implement the NPT [7], International Atomic Energy Agency (IAEA) Safeguards saw significant changes in the last 20 years aimed at meeting the verification challenges posed by an evolving nuclear and geo-political context. Through the Additional Protocol (AP), introduced in 1997, verification activities on declared sites aimed at detecting the diversion of material and the misuse of declared facilities were supplemented by additional tools to assure the absence of undeclared nuclear material and activities on both declared sites and undeclared locations.

The IAEA's State-Level Concept (SLC) foresees a holistic approach to nuclear safeguards considering that the State is a whole greater than the sum of its (declared) nuclear-related facilities. The SLC "*applies to all States and involves a comprehensive State evaluation and State-level safeguards approach, including the identification of specific safeguards measures for each State, implemented through an annual implementation plan*." [8]

Within the SLC, the IAEA verification activities are carried out according to a tailored "State-Level Approach" (SLA), adapted to each State, to be able to detect diversion or misuse of declared nuclear material and facilities, as well as the existence of undeclared nuclear material or activities [9, 10].

Central to the State-Level Approach (SLA) for safeguards, the Acquisition Pathways Analysis (APA) allows the IAEA to estimate the possible routes and the time needed to achieve weapons-usable material. The estimates take into account all the available safeguards-relevant information on a State. On the basis of this analysis, the IAEA will then be able to design and plan the specific verification activities needed to reach the safeguards technical objectives in the State, with a schedule informed by acquisition pathways completion times and the effectiveness and efficiency of the available safeguards measures. The State-Level Approach follows an annual implementation plan, and is re-evaluated and adapted yearly [9-11].

3. Open Source Analysis in Non-proliferation

Surprisingly enough, there is no universally accepted definition of "open source information" and "open source analysis". For the purpose of this paper, in line with previous works (see e.g. [12, 13]), "open source information" will be defined as "*publicly available material that anyone can lawfully obtain by request, purchase, or observation*." [14] As it can be seen, it is a very general and inclusive definition, de facto excluding only explicitly classified information. Within the NPT safeguards regime, the IAEA includes in the "open source information" basket "*information generally available from external sources, such as scientific literature, official information, information issued by public organizations, commercial companies and the news media, and commercial satellite imagery*" [15] and trade data [16, 17].

The broadness of the open source information definition above gives room to some fuzziness: for instance it is not clear if grey literature (i.e. non-classified material not meant for unlimited public dissemination and therefore not available through the standard publication channels such as technical reports, working papers, ephemeral publications, etc.), which some studies consider to be open source [18], can be assumed to be included.

If the definition of open source information is not universally shared, the definition of open source analysis is even more fragmented. In this paper, open source analysis will be defined as a process of "*getting the right information (what) to the right people (who) at the right time (when) for the right purpose (why) in the right forum (where) and in the right way (how)*" [19] by "*merging openly available data and information coming from a wide variety of accessible sources into an overall comprehensive and cohesive picture*" [12].

Table 1 reports the four broad analytical areas in which open source analysis in support to non-proliferation can be grouped and presents some possible information sources.

Analytical Area	Sources
Technical/information analysis	Scientific Literature, official information, information from public entities, commercial companies [20]
Media monitoring	News, blogs, social networks [21, 22]
Imagery analysis	Commercial satellite imagery, photographs, video snapshots [23]
Import/export Analysis	Trade data [17, 24], legal/illicit procurement information

Table 1: Open source analytical areas potentially supporting non-proliferation. Adapted from [6].

One of the main problems with open source analysis in non-proliferation is the fact that only a very small part of the information collected can be considered to be relevant. Typically the analyst needs to collect and then filter out substantial quantities of information to assemble a sparse and incomplete set, not necessarily contributing to knowledge [25]. In addition, the quality of the information might result to be dubious and the risk of deliberate deception high [26-28]. Despite all these problems, open source analysis has the potential of being one of the most promising sources of discovery and detection of undeclared nuclear activities on undeclared sites in a State. The following sections will focus on the possibility to use open source analysis to derive new insights on potentially undeclared nuclear activities[1] in a Non-Nuclear Weapon State (NNWS) that signed the NPT.

4. Some Methodological Considerations

While no verification activity is without issues, the detection of undeclared activities on undeclared sites faces formidable technical and epistemological challenges that would require dedicated research. The following paragraphs will provide a brief overview of some of them, highlighting the complexity of the task the IAEA has to carry out.

4.1 IAEA Safeguards Implementation Statements

According to the NPT, The task of the IAEA is to implement safeguards on Non-nuclear-weapon States (NNWSs) "*for the exclusive purpose of verification of the fulfilment of its obligations assumed under this Treaty with a view to preventing diversion of nuclear energy from peaceful uses to nuclear weapons or other nuclear explosive devices*" [7]. Currently, NNWSs can be broadly categorized in four groups:

1. States with a Comprehensive Safeguards Agreement and the Additional Protocol in force, for which the Agency already drew a broader conclusion;
2. States with a Comprehensive Safeguards Agreement and the Additional Protocol in force, for which the Agency has not already drawn a broader conclusion;
3. States with a Comprehensive Safeguards Agreement but no Additional Protocol;
4. States without a Comprehensive Safeguards Agreement.

For group 1, in 2017 the Agency concluded that for 70 States "*all nuclear material remained in peaceful activities*" [29]. For groups 2 and 3 the Agency concluded that, "*declared nuclear material remained in peaceful activities*" [29]. For group 4 the Agency "*could not draw any safeguards conclusion*" [29]. As it can be seen, the IAEA was able to exclude the presence of undeclared nuclear material and activities only for group 1, limiting itself to a statement over the *declared* nuclear material for groups 2 and 3.

The pivotal difference between the States in group 1 and those in group 2 is the broader conclusion of absence of undeclared material and activities. To reach a broader conclusion, the IAEA "*must draw the conclusions of both the non-diversion of the nuclear material placed under safeguards (as described above) and the absence of undeclared nuclear material and activities for the State as a whole*" [30]. The broader conclusion allows the entry into force of the Integrated Safeguards regime.

The evaluation and verification of declared nuclear material and activities is conceptually straightforward (even though it might be extremely resource-intensive), and is mainly based on onsite verification activities and measurements. The confirmation of the termination of past nuclear activities and the dismantlement of the related facilities (e.g. a decommissioned civilian nuclear fuel cycle programme or part of it) is usually performed through information gathering and complementary accesses, foreseen under the Additional Protocol, and does not represent unsurmountable conceptual challenges as it is known that there was a programme and the conditions of its termination have been stated and could in principle be verified. The Agency concludes that there is no undeclared nuclear material or activity in a State when "*the activities performed under an additional protocol have been completed, when relevant questions and inconsistencies have been addressed, and when no indications have been found by the IAEA that, in its judgement, would constitute a safeguards concern*" [30].

The sentence there is "*no indication of undeclared nuclear material or activities*" [29] can be read as "given the verification activities planned and performed on the basis of our past and present knowledge of the State, we found

[1] For an overview on what are nuclear materials and activities to be declared (and how they should be declared), see e.g. [8].

'*no indication of undeclared nuclear material or activities'* [29]" and therefore it is possible to conclude that "*all nuclear material remained in peaceful activities*" [29]. The last *therefore* relies on the inductive inference: "The n verification activities performed did not find evidence of an undeclared activity, and they are considered to be sufficient to state that any additional verification activities would not find evidence of undeclared activities". Hence we can conclude that there is no undeclared activity. Since "*[t]he very nature of an inductive argument is to make a conclusion probable, but not certain, given the truth of the premises*" [31], it becomes extremely important to discuss how the strength of this conclusion can be characterized and made explicit, i.e. characterize its *dependability*.

While "*Truth with a capital T is an attribute of statements that correspond to facts in all possible contexts*", *dependability* is an attribute of statements that correspond to facts in a "*specified (but often not clearly identified) context*" [19]. A statement is considered to be more or less dependable subject to the degree to which it has been tested. Having to rely on dependability rather than on truth implies that the aspect of uncertainty management is important to make sure the decision-maker obtains an accountable message.

Although there are several ways of describing uncertainty (see e.g. [32]), here uncertainty will be set to include two main aspects: *aleatory* and *epistemic* [33] Aleatory uncertainty, also called randomness, is related to data exhibiting an intrinsic lack of a specific pattern, and is investigated by classical probability theory. This source of uncertainty is intrinsic and cannot be eliminated. Epistemic uncertainty describes the analyst's less than perfect knowledge of the data, and could theoretically (but de facto not practically) be reduced to zero. In every analysis entailing uncertainty, aleatory and epistemic uncertainties are intertwined and need to be dealt with [34].

4.2 Dependability of Open Source Analysis in identifying an undeclared nuclear activity is still an unknown

Excluding the provision of third-party information and assuming no open anomaly or discrepancy arising from declarations and inspections, the main source for discovering potential undeclared nuclear activities in a State at the Agency's disposal is open source information collection and analysis. To understand the degree of dependability of the statement "no undeclared nuclear activity exists in the State as a whole", one would need to identify, for each nuclear activity:

1. Which are the potential indicators one might aim at to identify its presence and what is their strength;
2. Which are the tools that would be able to identify the existing indicators and what is their efficiency in detecting them;
3. What is the effectiveness of the available detection methods in detecting an existing signal in a real world scenario. This implies the knowledge of, *inter alia*, the size of the search space and the share of the search space the safeguards staff can reasonably cover with a given technique.

Some information about which are the potential indicators of existence of a particular nuclear technology in a State is available to the IAEA in the Physical Model [35] (point 1), and, although to the authors' knowledge it has never been published in the open Literature, it is possible to create a catalogue of potential tools able to detect the various indicators (point 2). The task of producing this catalogue is not without challenges as, for a given technology, the scientific community has contrasting opinions about their difficulty of implementation and detection [36, 37].

Understanding the size of the search space and the share of the search space the safeguards staff can reasonably cover with a given technique still represents a challenge. To make things worse, deliberate deception and signal suppression by the proliferator could lower the detection efficiency (Table 2 reports examples of concealment techniques that a proliferator could adopt to suppress/scramble potential signals available in the open source).

Gathering evidence for the presence of indicators of a nuclear engineering programme in a State, especially when related to military aspects and therefore with active efforts to keep it concealed, means having to deal with additional epistemological issues [39]:

- The analyst is out to detect something whose existence is uncertain. Making a parallelism with any classic measurement performed by an inspector on declared nuclear material in a declared facility [40, 41], while the three postulates of the theory of measurement [42] inform the inspector that he will never know the absolute true value of the characteristics he is measuring, they also inform him that the true value does exist and it is -within the limit of the typical safeguards measurement campaigns - de facto constant. In contrast, when searching for indicators of a possible clandestine nuclear programme, the analyst does not know whether such a programme really exists.
- Assuming that the clandestine programme exists, the analyst does not know its characteristics, and therefore would not be in the position to choose the best detection method for finding indicators of its presence[2]. This not only impacts effectiveness and efficiency, but also adds considerable epistemic uncertainty about the outcome of the verification activities.

[2] One of the purposes of the Acquisition Pathways Analysis step of the State Level Concept is to guide the analyst to the most likely characteristics of a possible clandestine programme.

OS Analysis	Type of Signature	Possible Concealment
• Technical/official information analysis • Media monitoring	R&D activities	• Manage publication activities • Use widely available technical information • Claim legitimate applications • Cover stories
	Environmental monitoring, public health records	• Suppress effluents • Suppress reporting
Imagery Analysis	Security features of infrastructure	Conceal or place within other secure facilities
	Functional and Operational design features	Mask true use through signature suppression
Import/Export analysis	• Patterns of material acquisition • Special equipment acquisition • Imports of dual-use equipment	• Shuffle, divert acquisitions • Obtain from multiple suppliers/intermediaries • Mix with legitimate uses • Develop clandestine networks • Produce indigenously • Divert equipment from legitimate activities • Claim legitimate uses

Table 2: Examples of concealment techniques to suppress/scramble potential signals in the open source. Adapted from chemical and biological weapon program signatures and concealment actions [38] as presented in [6].

Actions	Possible Sources
1. Monitor the New Media for cueing insights (cast a wide net for new information)	Blogs, news aggregators, pushed emails from topical interest groups, paid subscription services
2. Search area of interest on virtual globes based on cueing from collateral information	Google Earth, Here, Bing Maps, Flash Earth, Yandex Maps, Arc GIS (online base map imagery)
3. Review all related geospatial labeling	Wikimapia, Google Earth Community forums
4. Review all available ground imagery on social media, photo–sharing sites, videos for additional possible insights	Lookr, Flickr, Worldflicks, Instagram, YouTube, etc. Some of them available as Google Earth layers
5. Follow-up with search of cues and cues derived from labeling, including imagery and news	Google, Bing, Yahoo, Yandex, Baidu, Armscontrolwonk, ISIS-online, etc
6. Review all available historical overhead imagery on Google Earth	Google Earth Historical Layer
7. Cross-reference images with commercial satellite imagery vendor archives to: a. Determine acquisition dates b. Review most recent imagery in archives for any significant changes c. Determine if additional imagery purchase is warranted	Digital Globe, Airbus, etc. Metasearch engines as Image Hunter are sometimes useful.
8. Determine if enough information is available to make an assessment	Make determination with appropriate caveat (definite, probable, possible, suspect, etc) or not enough information.

Table 3: Possible actions for promoting imagery-related discoveries in the Open Source [22].

- In addition to the above-mentioned issue, even in the case in which the analyst chooses the appropriate method for detecting the existence of a given indicator, as previously discussed the actual detection probability of the presence of an indicator given its existence when using a given detection/measurement technique is often not known as no dependable attempts to investigate this aspect are available.

As a consequence, it is very difficult to tell the real, effective degree of dependability of verification activities for undeclared activities on undeclared sites.

Despite the current efforts in trying to systematize the possibility of making new discoveries [43, 44], until the above aspects are exhaustively investigated, it is likely that the discovery of undeclared activities on undeclared sites will remain serendipitous [22].

4.3 Potential actions for Deriving New Insights from Open Source Imagery Information

Table 3 proposes a potential set of actions for promoting discoveries in the Open source that might reveal potential undeclared nuclear activities on undeclared sites as presented in [22].

With all the caveats previously expressed, open source analysis can provide insights on previously unknown aspects of a nuclear fuel cycle in a State, and the above actions can potentially enhance the possibility of making new discoveries combining open source information and high resolution satellite images. The next section will provide an example of a new discovery in the field of potential nuclear-related activities in a State. Where relevant, references to the above actions will be made.

While some of the above actions are usually performed in sequence (e.g. actions 2-3, sometimes actions 2-7), they should not be intended as a rigid ordered sequence: depending on the nature of the initial cue (e.g. a high-resolution image in a virtual globe or a blog post), the other actions will follow according to the most suited sequence.

5. Exemplary Application of Heterogeneous Open Source Fusion for Deriving New and Potentially Safeguards-Relevant Information

By following-up on a single media report of the construction of a national radioactive waste storage facility located near Anarak, Iran, it was possible to correctly locate and characterize the radioactive waste site with commercial satellite imagery (starting with Google Earth), but it also became possible through subsequent open source analysis of the geological setting of that radwaste site to discover that a nearby, previously abandoned, mine – known to be in an area containing copper, nickel, cobalt, arsenic, and uranium – that had been reactivated, and, moreover, that an ore processing facility (likely pilot-scale) had been newly-established nearby to process ore from that mine. This is illustrated in the following sections. Reference is made of the "Actions" mentioned in Table 3.

5.1 Background

Anarak, Iran, is identified in the open literature as having historically been the site of three nuclear-related sites. Two were former uranium mines (identified as Talmessi and Meskani) with the third a small interim solid radioactive waste site. The mines have long been considered to have been mined-out and abandoned. The interim solid radioactive waste site was decommissioned in 2004. The solidified radioactive waste previously stored there, was generated during operations on small amounts of imported UO_2 that had been prepared for targets at Jabr Ibn Hayan Multipurpose Laboratories (JHL), irradiated at the Tehran Research Reactor (TRR), and sent to a laboratory belonging to the Molybdenum, Iodine and Xenon Radioisotope Production Facility (MIX) in Tehran for separation of ^{131}I in a lead-shielded hot-cell. Iran had informed the IAEA that the remaining nuclear waste was solidified and eventually transferred to a waste disposal site at Anarak. Upon request by the IAEA, that waste was removed and transferred from Anarak to JHL in January 2004 for inspection. As of that time, no more nuclear material was known to be at the Anarak facility.[3] However, in October 2014, Iran media reported that "*[c]hief of the Atomic Energy Organization of Iran (AEOI) Ali Akbar Salehi paid a visit to a long-term nuclear waste storage facility in the central province of Isfahan…to get update in the on the construction of the nuclear waste stabilization and storage facility.*"[4] (Action 1)

5.2 The Pasmangoor Nuclear Waste Stabilization and Storage Facility

In early-2015, a search was conducted of commercial satellite imagery of the Anarak area, which made possible the identification of a likely candidate for the "Nuclear Waste Stabilization and Storage Facility" nearing completion near Anarak, Iran [45]. Having first located that candidate site on the most recent imagery available from Google Earth at that time (July 16, 2013) and having seen on the historical layer of Google Earth that the site was first underway by October 2011 (Action 2 and 6), a review of Digital Globe imagery archives revealed multiple acquisitions centered on that same site, indicating that this site first began attracting continuing interest by unknown others in early 2014 (Action 7). The facility exhibited the requisite features for a secure storage vault-type radwaste structure situated in a dry and stable area that is not susceptible to flash flooding (see Figure 1 and Figure 2) [46]. Available geological reports and a cross-section of the area, which highlight previously abandoned uranium mines located nearby, describe the geological setting as that of a "graben-syncline"[5], providing additional evidence for the site being appropriate for the storage of radwaste, as the subsurface geology is also stable. [47] (see Figure 3).

On April 7, 2016 (as part of the celebration of the tenth "National Nuclear Technology Day" in Iran), Iran inaugurated the "Pasmangoor Nuclear Waste Stabilization & Storage Facility"[6] and videos were presented by the Iranian government and subsequently posted on YouTube, which verified the above analysis (Action 4).[7] Those videos provided both aerial drone imagery of the site and interior views of the main storage vault building with radioactive waste storage canisters shown being off-loaded from a delivery truck by an overhead crane and into one of the concrete vaults. (see Figure 4 and Figure 5)

[3] http://www.globalsecurity.org/wmd/world/iran/anarak.htm

[4] http://www.tasnimnews.com/English/Home/Single/543275

[5] A 'graben syncline' is a concave fold of rock layers with its limbs lifted by faults, as a result of which the core of the fold becomes displaced downward relative to the rock layers on either side, as in a rift valley.

[6] http://www.tasnimnews.com/en/news/2016/04/07/1042135/iran-unveils-12-nuclear-achievements-includingnew-n-waste-facility

[7] https://www.youtube.com/watch?v=YbzzVeXlVOA and https://www.youtube.com/watch?v=_bQTYpz3ylg. The first mention of the Pasmangoor radioactive waste storage site was in an IAEA report from 2000, which reported that it was under preliminary site investigation. See: IAEA Waste Management Database: Report 2 - L/ILW-SL, March 28, 2000. http://www .pub.iaea.org/MTCD/publications/PDF/rwmp-3/Report_2.pdf. The plans called for "near surface disposal" in a "simple storage building."

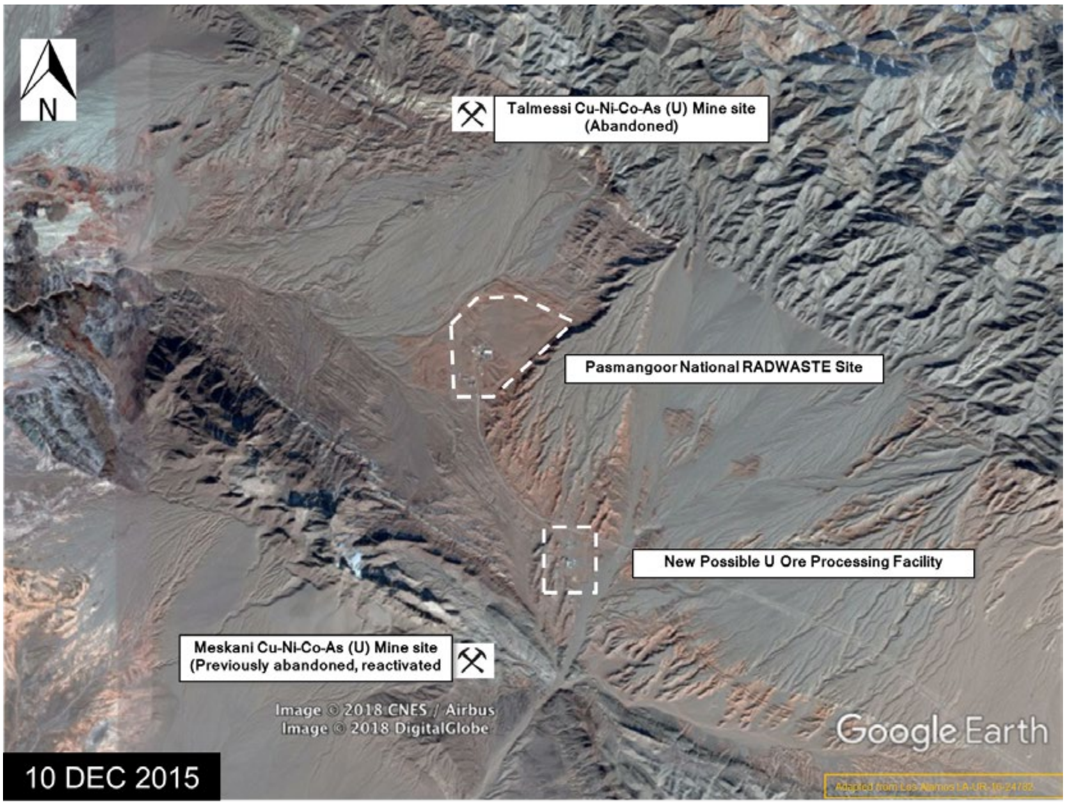


Figure 1: The Pasmangoor Nuclear Waste Stabilization and Storage Facility and the adjacent new possible uranium ore processing facility in relation to two known and previously abandoned uranium mines near Anarak, Iran. Given the proximity to two well-known uranium-mining areas, it would not be unreasonable to assume that this entire area is now under the authority the Atomic Energy Organization of Iran (AEOI).

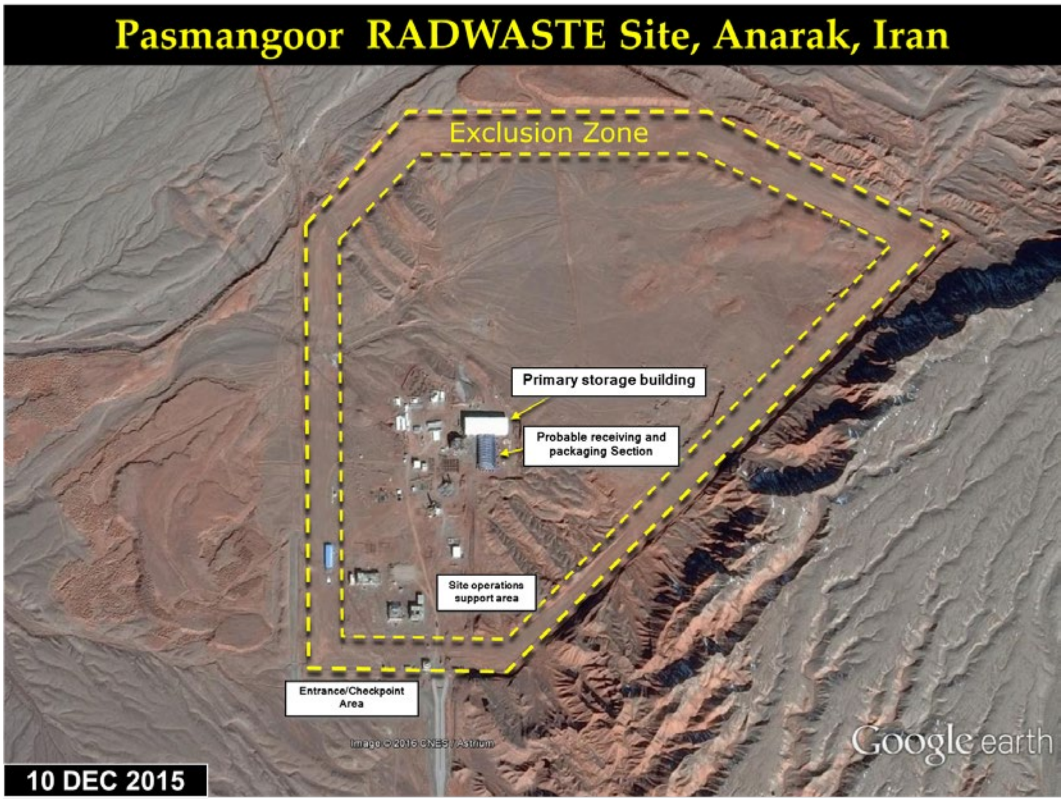


Figure 2: Overview of the new Pasmangoor Nuclear Waste Stabilization and Storage Facility near Anarak, Iran, as viewed on Google Earth. Note that the facility is double-perimeter-secured with a graded exclusion zone in between.

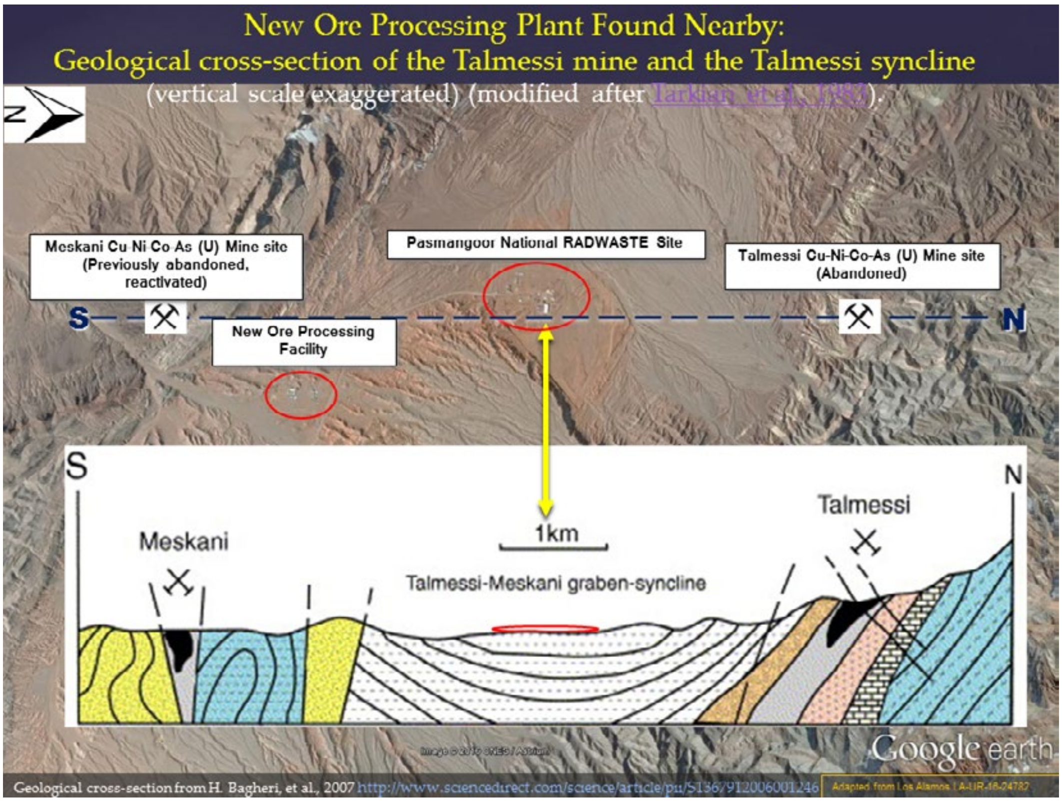


Figure 3: Geological cross-section running south to north (left to right) through the two former uranium mines and the new Pasmangoor national nuclear waste stabilization and storage facility. The vertical scale is exaggerated and this figure is modified after [48] as reported with the geological cross-section from [47].

Figure 4: The Pasmangoor Nuclear Waste Stabilization and Storage Facility near Anarak, Iran. The white-roofed building is the concrete vault radwaste storage building. See: https://www.youtube.com/watch?v=YbzzVeXlVOA

Figure 5: Interior views of the main concrete vault storage building as seen on an Iranian publicly posted video.
See: https://www.youtube.com/watch?v=_bQTYpz3ylg

5.3 The New Meskani Ore Processing Facility

The Pasmangoor Nuclear Waste Stabilization and Storage Facility identification started from a news media report, continued on high-resolution imagery and was corroborated by audiovisual posted on the web. The following example originated from the analysis of the high-resolution satellite images of the Pasmangoor Facility, continued with the search of scientific literature about the geology of the area and finally obtained additional corroboration over time by monitoring the evolution of the site's activities on high-resolution satellite imagery and researching of the site with open source tools.

In July 2018, Google Earth provided more current (July 4, 2018) commercial satellite imagery of higher resolution of both the Pashmangoor radwaste facility and a second facility (which during initial construction appeared as some type of "Entrance Facility" for the new radwaste site – Action 2). That second facility could be identified remotely on satellite imagery as a small "ore processing facility", and exhibited sufficient features to be assessed as a small (e.g., pilot-scale) "possible uranium ore processing facility" (See Figure 6). Among the noteworthy features of that facility which compared favorably with those found at the known Yellowcake Production Plant (YPP) located near Ardakan, Iran, include: truck scales and possible radiometry station on the road from the mine site (see Figure 7, top); and a segregated ore pile storage area (see Figure 7, bottom). Other identified features include: an ore receiving, crushing, grinding and conveyor circuit; five probable fine ore storage silos; a probable leaching building; a probable mixer-settler solvent extraction shed; reagent storage tanks; and a small concrete lined possible water-holding tank. More recent imagery from mid-2018 shows that an 11-meter diameter clarifier/settler tank had been added, along with what might be a growing processing waste pile (See Figure 8). Figure 9 provides another comparison of the ore processing related infrastructure observed at the Yellowcake Production Plant (YPP), Ardakan, Iran with that observed at the new ore processing facility associated with the Meskani copper-nickel-cobalt-uranium mine near Anarak, Iran.

Physical site security includes an entrance/exit checkpoint near the main road, and makeshift earthern barrier walls, which might also serve as visual obscuration berms for perimeter security along the road to the Pasmangoor nuclear waste stabilization and storage facility (see Figure 10**)**. Imagery from late May 2016 indicated that the ore processing facility could have become operational, as the ore piles had changed and liquid was visible in the formerly clean concrete lined holding tank (action 6). The nearby Meskani copper-nickel-cobalt-uranium mine site, which had been abandoned for decades, had evidently been reactivated (Action 6), and one large operations support building (with a blue roof) was constructed during February 2014 (See Figure 11). The mine and ore processing facility are serviced by a newly paved access road, which had just been built to support the Pasmangoor radwaste site.

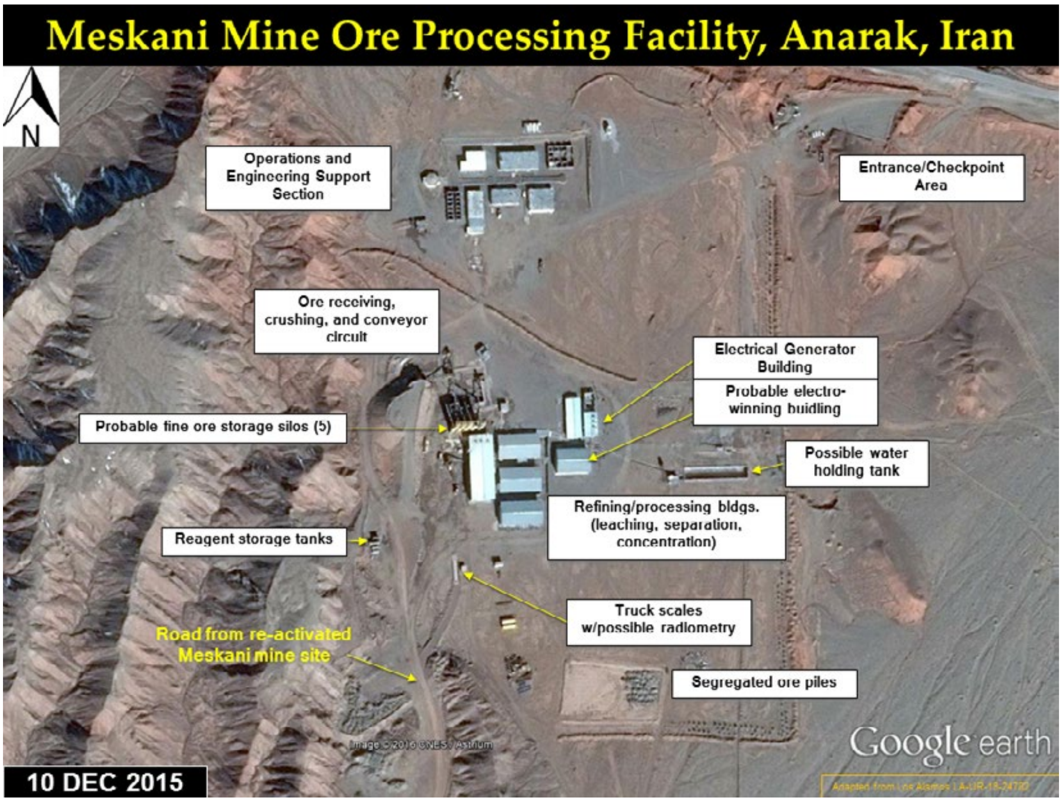


Figure 6: Late-2015 overview of the new small ore processing facility serving the newly re-activated, adjacent, Meskani mine. Labels identify the more likely roles of each part of the facility for illustrative purposes and are not meant to be definitive.

Figure 7: Comparison of features observed at the Yellowcake Production Plant (YPP), Ardakan, Iran with those observed at the new ore processing facility associated with the Meskani mine near Anarak, Iran. The top left shows an ore discrimination station with scales and radiometry for incoming ores from the Saghand uranium mine. The top right shows a similar appearing truck scale for the new small ore processing facility near the Meskani mine. The lower left shows the uranium ore on a storage pad near the ore crushing and grinding circuit at the YPP. The lower right shows mined ore piles on a storage pad at the new small ore processing facility near the Meskani mine.

Figure 8: Mid-2018 view of the new small ore processing facility serving the newly re-activated, adjacent, Meskani mine. Labels identify the more likely roles of each part of the facility and are not meant to be definitive.

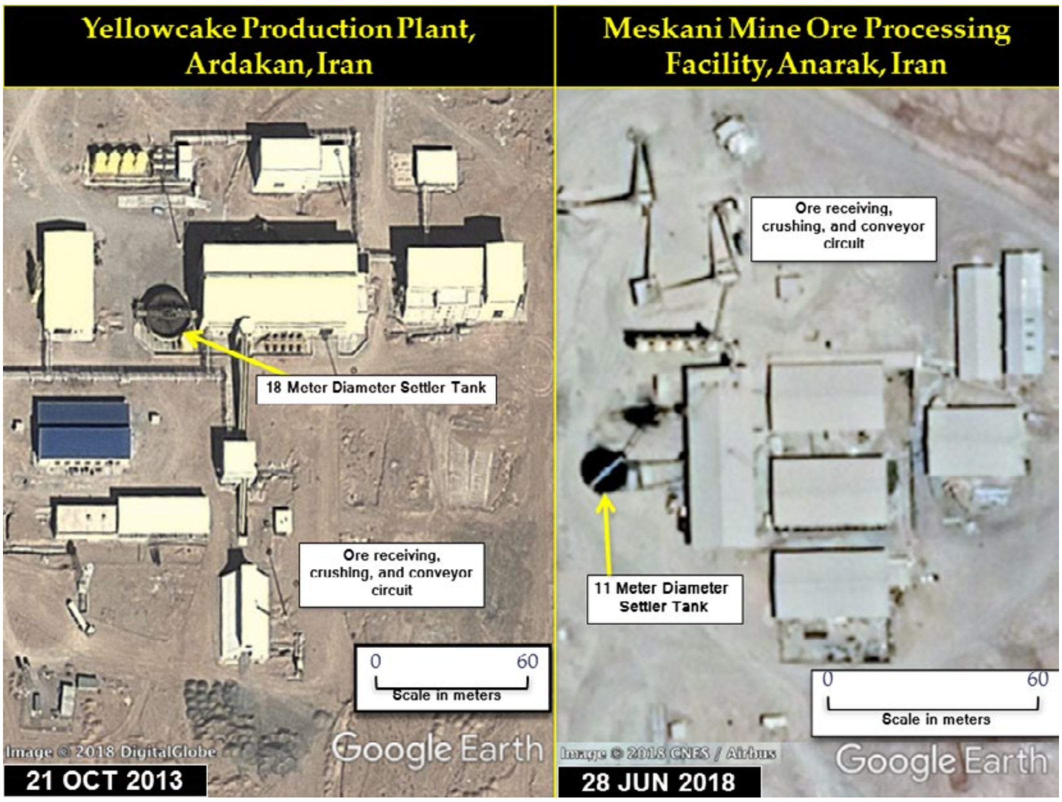


Figure 9: Comparison of ore processing related infrastructure observed at the Yellowcake Production Plant (YPP), Ardakan, Iran with that observed at the new ore processing facility associated with the Meskani mine near Anarak, Iran. An 11-meter diameter clarifier/settler tank was added to the processing circuit of the ore processing facility near Anarak post-2015 (right), smaller than the 18-meter diameter clarifier/setter tank located at Ardakan (left).

Figure 10: Operations and Engineering Section of the new ore processing facility showing the physical security as exemplified by an entrance/checkpoint area and visual obscuration berms that also serve as perimeter security barriers. Tree plantings (three rows of dot-like features) inside the berms are also evident, which will help with berm stabilization as well as providing additional future visual obscuration.

Figure 11: Close-up of the mining operations support area of the recently re-activated Meskani mine.

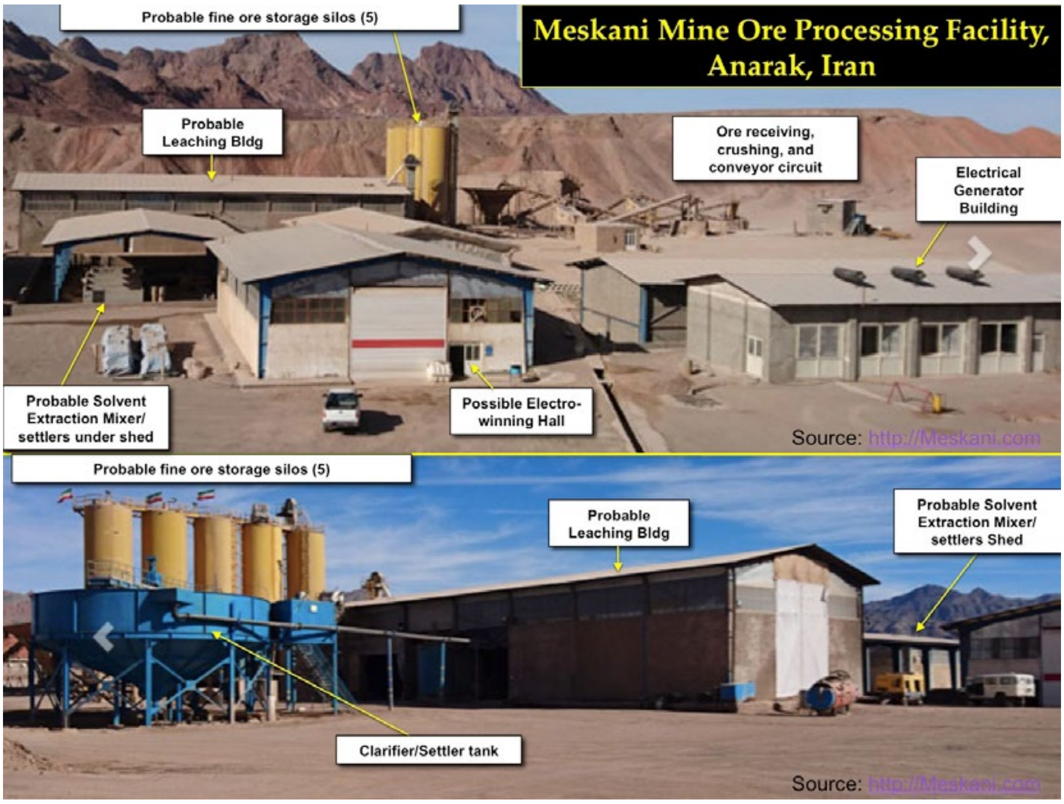


Figure 12: Ground Photos of the Meskani Ore Processing Facility. Labels identify the more likely roles of each part of the facility and are not meant to be definitive. The original images can be found at http://Meskani.com

5.4 Dependability of the Identifications

The identification of the radwaste facility is now certain, as it has been corroborated by official videos released by Atomic Energy Organization of Iran.

The identification of the processing facility as an "ore processing facility" has a very high degree of dependability, as the high-resolution satellite images provided by Google Earth over time allow a clear identification of the function of the installed infrastructure. Visual comparison with another well-known ore processing facility in the same country do not leave reasonable doubts on the correctness of the identification. Upon further open-source investigation, that included finding a Farsi labeled ground photo of the facility on the Google Earth "Photos" layer[8], it was possible to determine that the facility is indeed a facility for the processing ores from the Meskani mine. The identified operating company, Meskani, has a web site[9] providing a detailed history of the site, a discussion of ore process flow, an equipment list, along with two ground photos of the facility (actions 1, 4, 5). Figure 12 provides those two ground perspectives of the facility, with one showing the clarifier/settler tank, indicating that the images were acquired post-2015. All available information indicates that the ore processing facility is for the sole purpose of refining copper from those ores.

The Safeguards relevance of the ore processing facility cannot be ascertained easily from the high-resolution satellite images alone. While the area is known to contain uranium [47], the reporting found on the Meskani mine website claimed that the first cathode copper production of 99.99% was manufactured in February 2015. With respect to possible uranium extraction, the website specifically states that the AEOI (The Atomic Energy Organization of Iran) had, after 2005, "*begun extensive studies aimed at solidifying radioactive elements in the mine. Based on exploratory studies, the presence of radioactive elements in this mine was not proved, therefore, the organization returned the mine to the Industrial, Mine and Trade Organization of Iran in 2007.*" That information expands upon other available reporting that the AEOI had renewed uranium exploratory efforts in the Meskani area prior to 2007 [47]. The ability to remotely differentiate and characterize an ore processing facility that extracts copper vs uranium has been shown by other researchers [49, 50] to be somewhat difficult apart from the identification of a building housing the "Electro-winning" process step for copper extraction, which would not be part of any uranium processing flowsheet. The Meskani Mine website states that the facility includes such an electro-winning step, and the ground photos appear to provide support to that claim. What could be the electro-winning hall is located immediately adjacent to a building that can be clearly identified as an electrical power generator building (see Figure 8 and Figure 12). According to the research described above regarding differentiation between a typical copper ore processing facility and a typical uranium ore processing facility, "*[o]ne*

[8] The photo, dated December 10, 2015, provides a panoramic view of the ore processing facility with a label naming it as the "Meskani Copper Mine" https://plus.google.com/photos/photo/108081858346305578393/6365120356766431170

[9] http://meskani.com/

major difference between a dedicated Uranium Mill and a dedicated Copper mill has to do with the scale of operation. For economic viability Copper mills have to produce much larger outputs than Uranium mills. Thus invariably a Copper mill is at least 2 to 3 times larger than a Uranium mill" [50]. From an economic viability point of view, the size of the Meskani Ore processing facility seems to err on the small side for a copper processing facility.

No.	Criterion Description	H1 (copper mill)	H2 (uranium mill)
1	The area contains the material in object and is compatible with mining exploitation	The area is well known for containing copper [47, 48]. The quality of the copper ore in the area is reportedly sufficient for the operation of a copper mine	The area is known for containing uranium, but open literature reports the presence of U as a trace element with an abundance of ca. 10 ppm [47, 48], well below the inferior limit considered for commercial viability (ca. 1000 ppm[1,2]). One source mentions a literature study reporting the presence of "*[a] ggregate accumulations and single isolations of uranium-bearing (7-30%) hard bitumen*" [53]. According to another source the mine is among the oldest exploited uranium deposits in Iran with up to 200 tons of uranium potentially available [54] AEOI performed U explorations over the years that reportedly did not prove the presence of exploitable radioactive material[9].
2	The site hosts all the needed infrastructure	Most of the site infrastructure is compatible with a copper processing facility. The most important facility for discriminating between a copper and a uranium mill (the electro-winning facility) could not be identified dependably as the potential building hosting it is substantially different from the usual shape and characteristics [49, 50]	Most of the site infrastructure is compatible with a uranium processing facility. The most important facility for discriminating between a copper and a uranium mill (the precipitation facility) could not be identified dependably [49, 50]. The presence of a radiological portal at the weighing station cannot be corroborated by the available images (Figure 7).
3	The facility is consistent with other similar installations	The processing facility is erring on the small side for a copper mill to be considered as commercially viable [50]. The size of the CCD[3] seems however compatible with the throughput declared on one website dedicated to the facility[9]. The area has historically been of interest for Cu extraction.	The visual similarity with the Ardakan uranium ore concentrate processing facility is striking (Figure 9). The size of the Meskani facility is compatible with a small pilot scale uranium processing facility. Average U abundance in the Saghand ore (processed in Ardakan) is however two orders of magnitude higher than the one reportedly available at Meskani[4].
4	Collateral sources corroborate the end use	The processing facility has a dedicated website[9] identifying it as a copper mill, giving details about the company, the site, the type of infrastructure and processes. The site contains also a small history of the site and a couple of ground pictures (see Figure 12). The pictures available on the site do not provide conclusive evidence of the existence of an electro-winning facility.	There is no collateral reporting of uranium ore processing having occurred at this facility (but uranium has been reported to be one of the elements, in addition to copper, occurring in the ore that is being processed) One website[9] clearly identifies the site as a copper mill and explicitly states that the site was subject to AEOI explorations that reportedly did not find prove of radioactive material and therefore the site was released. The facility is close to an AEOI radiological waste disposal and storage site, but there is no evidence that the ore processing facility is under the purview of AEOI. The facility does not appear in the list of those declared to the IAEA. Its close proximity to a declared nuclear-related site makes the facility – easily revealed by satellite imagery - highly unsuitable for a covert, undeclared nuclear activity meant to remain secret, and a strong candidate for complementary accesses.

Table 4: Analysis of Competing Hypotheses (ACH) for the Safeguards significance of the Meskani mine. The two hypotheses against identified criteria. The analysis has the sole purpose of illustrating the method and should not be considered as a dependable analysis of the site.

1 http://www.world-nuclear.org/information-library/nuclear-fuel-cycle/mining-of-uranium/uranium-mining-overview.aspx
2 https://www.nr.gov.nl.ca/nr/mines/prospector/matty_mitchell/pdf/prospecting_for_uranium.pdf
3 "Counter Current Decantation" unit
4 https://www.iranwatch.org/iranian-entities/saghand-uranium-mine

No.	Criterion Description	H1 (copper mill)	H2 (uranium mill)
1	The area contains the material in object	++	+
2	The site hosts all the needed infrastructure	+/-	+/-
3	The facility is consistent with other similar installations	+/-	+
4	Collateral sources corroborate the end use	++	--

Table 5: Evaluation of the two hypotheses on the basis of the evidence supporting/disproving the identified criteria. The evaluation has the sole purpose of illustrating the method and should not be considered as a dependable analysis of the site.

The presence of a (possible) radiological portal at the weighing stations would represent a non-conclusive indicator, as the potential presence of radioactive material in the ores might justify a radiation monitor characterizing the truck loads leaving the facility for health and safety reasons.

Methods for the analysis of competing hypotheses (ACH), developed in the domain of intelligence analysis [51], are of potential interest for open source analysts that need to be able to "*consider inconsistent and anomalous information, develop competing hypotheses (which can include deceptions), and test hypotheses in a manner that reduces susceptibility to cognitive limits and biases*" [52]. Table 4 and Table 5 present a partial illustrative analysis of competing Hypotheses for the Meskani ore processing facility. The two hypotheses considered are H1: "The Meskani facility is for copper processing" and H2: "The Meskani facility is for uranium processing". The objective is to illustrate a possible use of the ACH and not to perform a non-proliferation analysis of the site used in the example. The analysis is illustrated making use of a subset of available information on the topic.

With this ACH illustration, while acknowledging that it is being derived from a limited subset of information, we can arrive at one valuable insight: Given the available collateral sources, H1 ("The Meskani facility is for copper processing") is largely favored over H2 ("The Meskani facility is for uranium processing").

In the purely hypothetical case in which the collateral sources were to be considered part of an elaborate deception scheme, the advantage of H1 over H2 would decrease considerably, and conclusive identification of the facility's purpose would require either additional evidence gathering or onsite access. It is therefore particularly important, for any open source analysis, to characterize thoroughly the completeness, coherence and quality of the information upon which the analysis is based, and make such characterization an explicit part of the message passed to the evaluation team.

6. Conclusion

The revelation of Iraq's clandestine nuclear weapons program in the 1990s first made clear the necessity of bringing new tools to bear in the service of the IAEA and other international non-proliferation efforts. The adoption in 1997 of the Additional Protocol by the IAEA Board of Governors paved the way to the introduction of the State-Level Concept (SLC). Open source information can play an important role in supporting several aspects of the SLC, integrating and supplementing the information retrieved from States' declarations and infield verification activities. In particular, Open source information has the potential of being the main key source of information for identifying potentially undeclared nuclear activities in a State, an area that historically proved to be the most challenging of the entire non-proliferation regime.

While no verification activity is without issues, the detection of undeclared activities on undeclared sites faces formidable technical and epistemological challenges; and until such sites are thoroughly investigated it is likely that the dependability in excluding their existence cannot come close to that for declared activities on declared sites, and any new discovery of such undeclared activities will remain serendipitous. Nonetheless, open source analysis can provide insights on previously unknown aspects of a nuclear fuel cycle in a State, and the actions presented in this paper can potentially enhance the possibility of making new discoveries combining open source information including high-resolution satellite images.

The serendipitous discovery of a new ore processing facility near an AEOI operated radiological waste site located near a previously abandoned mine in an area containing copper, nickel, and uranium provides an excellent analytical case study exemplar, combining heterogeneous open source and geospatial information to derive significant, but otherwise at that time unknown, information for nuclear safeguards applications. While all currently available open source information indicates that uranium is not currently being extracted as a byproduct of copper ore processing at the Meskani mine, the site has attracted the interest of AEOI and, as recently as 2007, was under study by the organization as a potential source of "radioactive elements"[9]. As a consequence, this exemplar is an instructive case in the combined use of open source analysis to identify potentially relevant industrial activities and analyze their safeguards significance.

7. Acknowledgements and Disclaimer

The reflections here presented stem from the activities performed within the activities “Innovative Concepts and Methodologies for Nuclear Safeguards” (INSAF) and “Open Source Information For Nuclear Security” (OSINS), later incorporated in the project “Methods, data analysis and knowledge management for Nuclear Non Proliferation, Safeguards & Security” (MEDAKNOW) funded within the European Commission (EC) Euratom Horizon 2020 Research and Training Programme. The authors are grateful to all the researchers that contributed to these projects, In particular Dr. Lance K. Kim, Dr. Rainer Jungwirth, Dr. Simone Cagno and Mr. Erik Wolfart. The discussions with them fertilized the ideas that are here expressed, even though any incorrectness or mistake is the authors’ sole responsibility. The views expressed in this article are those of the authors and do not necessarily represent the official views of the European Commission or other Institutions.

8. References

[1] Rockwood, L. Safeguards and Nonproliferation: The First Half-Century from a Legal Perspective. In ESARDA (Editor), *Nuclear Safeguards and Non-Proliferation: Course Syllabus*, pp. 79–94. Office for Official Publications of the European Communities, 2008. https://esarda.jrc.ec.europa.eu/images/files/ESARDA_Course/nuclear_safeguards_and_non-proliferation%20final.pdf.

[2] International Atomic Energy Agency. *Model Protocol Additional to the Agreement(s) Between State(s) and the International Atomic Energy Agency for the Application of Safeguards*. INFCIRC/540(Corrected). IAEA, 1997.

[3] Keeley, J.F. and Cameron, J.K. The need to know: commercial satellite imagery and IAEA safeguards. In *Non-proliferation, arms control and disarmament: enhancing existing regimes and exploring new dimensions*. 1998. http://dspace.ucalgary.ca/bitstream/1880/44991/1/Keeley_NPACD_1998_KJ018.pdf.

[4] Ferguson, M. and Norman, C. All-Source Information Acquisition and Analysis in the IAEA Department of Safeguards. In *IAEA Symposium on International Safeguards, Vienna*, IAEA-CN-184/048. 2010.

[5] IAEA. IAEA Safeguards - Serving Nuclear Non-Proliferation, 2015.

[6] Renda, G., Kim, L.K., Jungwirth, R., Pabian, F.V., and Cojazzi, G.G.M. The potential of open source information in supporting Acquisition Pathway Analysis to design IAEA State Level Approaches. In *Proceedings of the IAEA International Safeguards Symposium: Linking Strategy, Implementation and People, Vienna*. 2014.

[7] IAEA. Treaty on the Non-Proliferation of Nuclear Weapons, 1970.

[8] International Atomic Energy Agency. *Guidance for States Implementing Comprehensive Safeguards Agreements and Additional Protocols*. Technical Report IAEA Service Series 21, IAEA, 2012.

[9] International Atomic Energy Agency. The Conceptualization and Development of Safeguards Implementation at the State Level - Report by the Director General (GOV/2013/38). Accessed via ISIS - Nuclear Iran, 2013. http://www.isisnucleariran.org/assets/pdf/GOV201338.pdf.

[10] International Atomic Energy Agency. Supplementary Document to the Report on The Conceptualization and Development of Safeguards Implementation at the State Level (GOV/2014/41), 2014. https://armscontrollaw.files.wordpress.com/2014/09/iaea-state-level-safeguards-document-august-2014.pdf. Report to the IAEA Board of Governors.

[11] Renis, T., Yudin, Y., and Hori, M. Conducting Acquisition Path Analysis for Developing a State-Level Safeguards Approach. In *Proceedings of the 55th INMM Annual Meeting*. 2014.

[12] Renda, G., Cojazzi, G.G.M., and Kim, L.K. A Systems Thinking Approach to Open Source Analysis in Support of Nuclear Non-Proliferation. In *Advances in Nuclear Nonproliferation Technology and Policy Conference: Bridging the Gaps in Nuclear Nonproliferation*. American Nuclear Society, American Nuclear Society, 2016.

[13] Renda, G., Cojazzi, G.G.M., Cagno, S., and Kim, L.K. A Systems Thinking Approach to Open Source Information Analysis for Nuclear Non Proliferation. In *Proceedings of the 14th International Workshop on Separation Phenomena in Liquids and Gases (SPLG 2017), 4-9 June 2017, Stresa, Italy*, JRC105992. 2017.

[14] Best, Jr., R.A. and Cumming, A. *Open Source Intelligence (OSINT): Issues for Congress*. Technical Report RL34270, Congressional Research Service, 2007.

[15] International Atomic Energy Agency. The Safeguards System of the International Atomic Energy Agency. 2009.

[16] Versino, C., Cojazzi, G.G.M., Contini, F., Tsois, A., Chatelus, R., and Tarvainen, M. Global trade data to support IAEA safeguards. *ESARDA Bulletin*, volume 45:pp. 14–21, 2010.

[17] Marinova, E., Ardhammar, M., El Gebaly, A., Schot, P.M., and Ferguson, M. Analysis of nuclear relevant information on international procurement and industrial activities for safeguards purposes. In *Symposium on International Safeguards: Linking Strategy, Implementation and People, Vienna (Austria), 20-24 October 2014 . Book of Abstracts, Presentations and Papers*, CN–220-085. 2015.

[18] Holland, B.R. *Enabling Open Source Intelligence (OSINT) in private social networks*. Technical report, Iowa State University, 2012.

[19] Blockley, D.I. The Importance of Being Process. *Civil Engineering and Environmental Systems*, volume 27, no. 3:pp. 189–199, 2010.

[20] Wallace, R. and Lundy, A. *Using Open Source for Proliferation Analysis*, chapter 11, pp. 197–220. Elsevier, 2008.

[21] Cojazzi, G.G.M., van Der Goot, E., Verile, M., Wolfart, E., Rutan Fowler, M., Feldman, Y., Hammond, W., Schweighardt, J., and Ferguson, M. Collection and Analysis of Open Source News for Information Awareness and Early Warning in Nuclear Safeguards. *ESARDA Bulletin*, volume 50:pp. 94–105, 2013.

[22] Pabian, F.V., Renda, G., Jungwirth, R., Kim, L.K., Wolfart, E., and Cojazzi, G.G.M. Open Source Analysis in Support of Nonproliferation Monitoring and Verification Activities: Using the New Media to Derive Unknown New Information. In *Proceedings of the 2014 IAEA Symposium on International Safeguards: Linking Strategy, Implementation and People, Vienna, 20-14 October 2014*. 2014.

[23] Pabian, F. *Commercial Satellite Imagery: Another Tool in the Nonproliferation Verification and Monitoring Toolkit*, chapter 12, pp. 221–249. Elsevier, 2008.

[24] Versino, C., Tsukanova, M., and Cojazzi, G.G.M. *Catalogue of WEB Data Services on Global Trade*. EUR 24657 EN. Publication Office of the European Union, 2010.

[25] Few, S. *Signal: Understanding What Matters in a World of Noise*. Analytics Press, 2015.

[26] Tsikerdekis, M. and Zeadally, S. Online deception in social media. *Communications of the ACM*, volume 57, no. 9:pp. 72–80, 2014.

[27] Katz, J.I. Deception and Denial in Iraq: The Intelligent Adversary Corollary. *International Journal of Intelligence and Counter Intelligence*, volume 19, no. 4:pp. 577–585, 2006.

[28] Katz, J.I. *Deception and Denial in Iraq: The Limits of Secret Intelligence and the Intelligent Adversary Corollary*. Technical report, Washington University, 2006.

[29] International Atomic Energy Agency. Safeguards Statement for 2017. IAEA website, 2017. https://www.iaea.org/sites/default/files/18/06/statement-sir-2017.pdf.

[30] International Atomic Energy Agency. *IAEA Safeguards Glossary*. Number 3 in International Nuclear Verification Series. IAEA, 2002.

[31] Hales, S.D. Thinking tools: You can prove a negative. *Think*, volume 4, no. 10:pp. 109–112, 2005.

[32] Walley, P. *Statistical reasoning with imprecise probabilities*. Chapman and Hall, 1991.

[33] Hacking, I. *The emergence of probability: A philosophical study of early ideas about probability, induction and statistical inference*. Cambridge University Press, 2006.

[34] Ferson, S. and Ginzburg, L.R. Different methods are needed to propagate ignorance and variability. *Reliability Engineering & System Safety*, volume 54, no. 2-3:pp. 133–144, 1996.

[35] Liu, Z. and Morsy, S. Development of the Physical Model. In *Proceedings of the IAEA Safeguards Symposium*, volume 29. 2007. http://www-pub.iaea.org/MTCD/publications/PDF/SS-2001/PDF%20files/Session%2013/Paper%2013-07.pdf.

[36] Kim, L.K., Renda, G., and Cojazzi, G.G.M. Methodological Aspects on the IAEA State Level Concept and Related Acquisiton Path Analysis. In *Proceedings of the 37th ESARDA Annual Meeting*. 2015.

[37] Kemp, R.S. The Nonproliferation Emperor Has No Clothes. *International Security*, volume 38:pp. 39–78, 2014. http://hdl.handle.net/1721.1/89182.

[38] US Congress, Office of Technology Assessment. *Technologies Underlying Weapons of Mass Destruction*. Technical Report OTA-BP-ISC-115, US Congress, Office of Technology Assessment, 1993.

[39] Renda, G. and Cojazzi, G.G.M. *Open Source Information Analysis in Support to Non-Proliferation: A Systems Thinking Approach*. Technical Report JRC102286, EC Joint Research Centre, 2018.

[40] International Atomic Energy Agency. *International Target Values 2010 for Measurement Uncertainties in Safeguarding Nuclear Materials*. STR Series STR-368, IAEA, 2010.

[41] International Atomic Energy Agency. *IAEA Safeguards Technical Manual: part F - Statistical Concepts and Techniques*. IAEA-TECDOC TECDOC-261, IAEA, 1982.

[42] Rabinovich, S.G. *Measurement errors and uncertainties: theory and practice*. Springer Science & Business Media, 2006.

[43] Kornell, J., Gastelum, Z.N., and Goldblum, B.L. Informational Sensing for Nonproliferation. In *Advances in Nuclear Nonproliferation Technology and Policy Conference: Bridging the Gaps in Nuclear Nonproliferation*. American Nuclear Society, American Nuclear Society, 2016.

[44] Kim, L.K., Renda, G., and Cojazzi, G.G.M. How Much is Enough?: Preliminary Insights from Search Theory as Applied to Detecting Undeclared Activities. In *Proceedings of the INMM 57th Annual Meeting*. Institute of Nuclear Materials Management, 2016. http://www.inmm.org/source/proceedingssearch/.

[45] Pabian, F., Renda, G., Jungwirth, R., Kim, L.K., Wolfart, E., and Cojazzi, G.G.M. Recent Developments Promoting Open-Source Geospatial Synergy: Emerging Trends and Their Impact for Nuclear Nonproliferation Analysis. In *Proceedings of the 56th INMM Annual Meeting, Indian Wells, USA*. 2015.

[46] Pabian, F.V. *Commercial Satellite Imagery as an Evolving Open-Source Verification Technology: Emerging Trends and Their Impact for Nuclear Nonproliferation Analysis*. EUR Report EUR27687, EC Joint Research Centre, 2016. http://publications.jrc.ec.europa.eu/repository/bitstream/JRC97258/reqno_jrc97258_online version pdf.pdf.

[47] Bagheri, H., Moore, F., and Alderton, D. Cu–Ni–Co–As (U) mineralization in the Anarak area of central Iran. *Journal of Asian Earth Sciences*, volume 29, no. 5:pp. 651–665, 2007.

[48] Tarkian, M., D. Bock, W., and Neumann, M. Geology and mineralogy of the Cu-Ni-Co-U ore deposits at Talmessi and Meskani, central Iran. volume 32:pp. 111–133, 1983.

[49] Sundaresan, L., Chandrashekar, S., and Jasani, B. Monitoring Uranium Mining and Milling using Commercial Observation Satellites. *ESARDA Bulletin*, volume 53:pp. 73–82, 2015.

[50] Sundaresan, L., Chandrashekar, S., and Jasani, B. Discriminating Uranium and Copper mills using satellite imagery. *Remote Sensing Applications: Society and Environment*, volume 5:pp. 27–35, 2017.

[51] Heuer, R.J. *Psychology of intelligence analysis*. Center for the Study of Intelligence, 1999.

[52] Stech, F.J. and Elässer, C. Deception Detection by Analysis of Competing Hypothesis. In *Proceedings of the 2005 International Conference on Intelligence Analysis*, volume 8, pp. 12–15. Citeseer, 2005.

[53] Bagheri, H., Moore, F., Shamsipour, R., Noghrehian, M., and Romanko, A. Reactivation and polyphase mineralization in Anarak area, Central Iran. http://www.ginras.ru/struct/21/11/files/BagheriMooreShamsipourNoghrehianRomankoCenozoic_talm.pdf.

[54] Cordesman, A. and Seitz, A. *Iranian Weapons of Mass Destruction: The Birth of a Regional Nuclear Arms Race?*. Praeger Security International. Praeger Security International, 2009. https://books.google.it/books?id=vK8WU7OWKpwC.

Evolution of Verification Data Evaluation under the State-Level Concept

Claude Norman, Jacques Baute, Robert Binner, Mika Nikkinen, Agatha Walczak-Typke

International Atomic Energy Agency (IAEA)
Wagramer Strasse 5, A-1400 Vienna, Austria

Abstract:

Every year, thousands of days are spent by IAEA inspectors in nuclear fuel cycle (NFC) facilities and other sites around the world. A large portion of this time is used for carrying out in-field measurements by various non-destructive assay (NDA) techniques and for taking environmental samples (ES) and/or destructive analysis (DA) samples. Comparably intensive resources are needed to maintain continuity of knowledge (CoK) on the verification data collected through these activities by means of a range of sophisticated containment and surveillance (C/S) systems. The IAEA collects, authenticates, quality controls, maintains and evaluates a large body of verification data and compares them with State declarations to support two of the main objectives of safeguards under the State-level concept (SLC), that is: the detection of diversion of nuclear material and of undeclared production or processing of nuclear material at declared facilities and locations outside facilities (LOFs).

In recent years, the NFC Information Analysis Section of the Safeguards Department Division of Information Management (SGIM-IFC), which is in charge of the evaluation of verification data, has been faced with a number of challenges: the first and most demanding is the need to evolve facility-based evaluation concepts to innovative, consolidated concepts that can integrate different types of information and support credible State-level safeguards conclusions, the second is the increasing volume and diversification of verification data to be evaluated given static resources, and the third is the need to keep abreast of modern methodologies and technologies with a view to ensure optimal effectiveness and efficiency.

This paper reviews the conceptual and methodological issues associated with these challenges and the approach that was applied to address them while taking advantage of the corresponding development opportunities. It presents the overall strategy adopted as well as the supporting project plan and the progress made to date in the related project components, with a special emphasis on the implementation of data visualization tools.

Keywords: evaluation; State-level Concept; methodology; diversion detection; visualization.

1. Introduction

The main mission of the IAEA Department of Safeguards is to provide credible assurances that States are abiding by their safeguards obligations. Since the safeguards system was strengthened after the discovery of a clandestine nuclear weapon programme in Iraq in the early 1990s and its legal authority was subsequently reinforced by the additional protocol (AP) in 1997, the nature and sources of information collected and evaluated by safeguards experts have extensively diversified and the volume of material to be researched has considerably increased. The Division of Information Management provides the Department of Safeguards with services of data processing, secure information distribution, information analysis and knowledge generation and consists of teams of professionals specialized in the analysis of different types of information plus a team in charge of information integration. These specialists play a critical role in the work of the Division of Operations' State evaluation groups (SEGs) in identifying, analysing and consolidating safeguards-relevant information from all sources to draw independent, non-discriminatory and soundly based conclusions for all States having concluded a safeguards agreement (Fig. 1) [1,2,4].

All-source safeguards-relevant information falls in three broad categories:

- Information declared by States, which consists in nuclear material accountancy (NMA) reports and reports submitted to the IAEA pursuant to the AP to the States' safeguards agreements.
- Information resulting from verification activities, e.g. results of NDA measurements, DA samples and ES samples, seals verification, surveillance review and other verification activities.
- Other relevant information, e.g. from open sources (OS) or provided by third parties, such as, for example, media, scientific publications, IAEA and public databases, trade import/export information and commercial satellite imagery.

The organizational structure of the Division of Information Management reflects these categories, which correspond to different analytical competencies. Besides the Integration and Coordination Team, it comprises four specialized

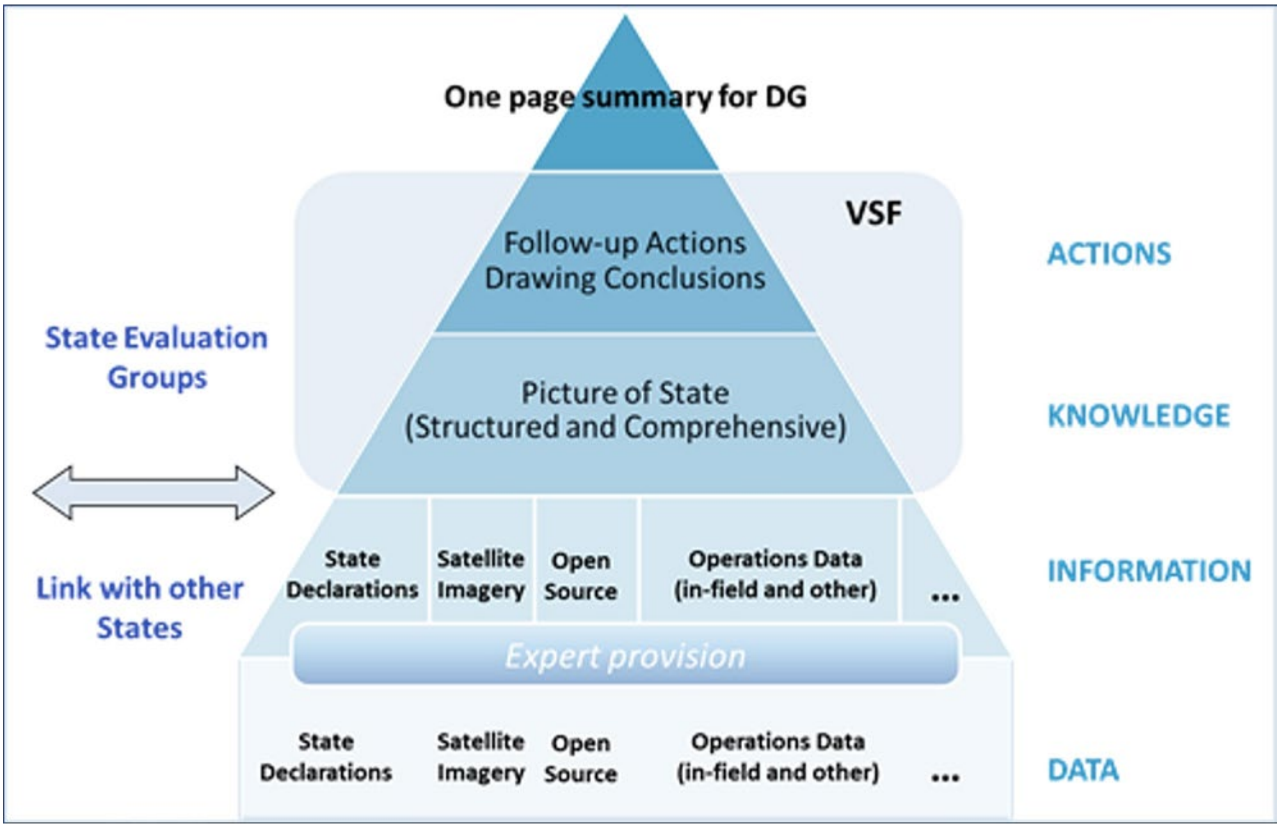


Fig.1: All Source safeguards-relevant information analysis – from data to actions [3]

Sections: the Declared Information Analysis Section whose role is self-explanatory, the State Factor Information Analysis Section in charge of general OS information analysis, the State Infrastructure Analysis Section specialized in geospatial information and satellite imagery analysis and the Nuclear Fuel Cycle (NFC) Information Analysis Section which collects, performs quality control of, stores, and evaluates results from in-field NDA measurements and from ES and DA samples to compare them with State declarations.

This paper will focus on the activities of the NFC Information Analysis Section. Its objective is to describe the challenges and opportunities encountered in this area from the evolution of the safeguards landscape and concepts [5], from the need for enhanced efficiency to cope with an ever increasing volume of data under static and sometimes reduced resource conditions, as well as from the progress made in information technology (IT) and data processing and evaluation methodologies. Section 2 below describes the strategy that was developed to address these challenges in a consistent, integrated and synergic manner, while utilizing state-of-the art IT tools and innovative data analysis and presentation. For each component of this strategy, it will review the progress accomplished to date as well as future development plans.

2. Verification data evaluation and its evolution under the State Level Concept

Every year, thousands of days are spent by IAEA safeguards inspectors in NFC facilities and other sites around the world. A large portion of this time is used for carrying out in-field measurements by various NDA techniques and for taking ES and/or DA samples. Comparably intensive resources are needed to maintain CoK on the verification data collected through these activities by means of a range of sophisticated C/S systems. The IAEA collects, authenticates, quality controls, maintains and evaluates a large body of verification data. In this context, the specific mission of the NFC Information Analysis Section, as illustrated in Fig. 2 below is defined as follows: *to contribute to the Department's provision of credible safeguards conclusions through the evaluation of verification data from samples (ES, DA) and in-field measurements (NDA) and their comparison with State declared information in order to detect and deter diversion and undeclared activities at declared facilities and sites.*

2.1 ES data evaluation – detection of undeclared nuclear material and activities

Fig. 2 shows that the role of ES data evaluation [12,13] is different from that of NDA and DA data evaluation and that

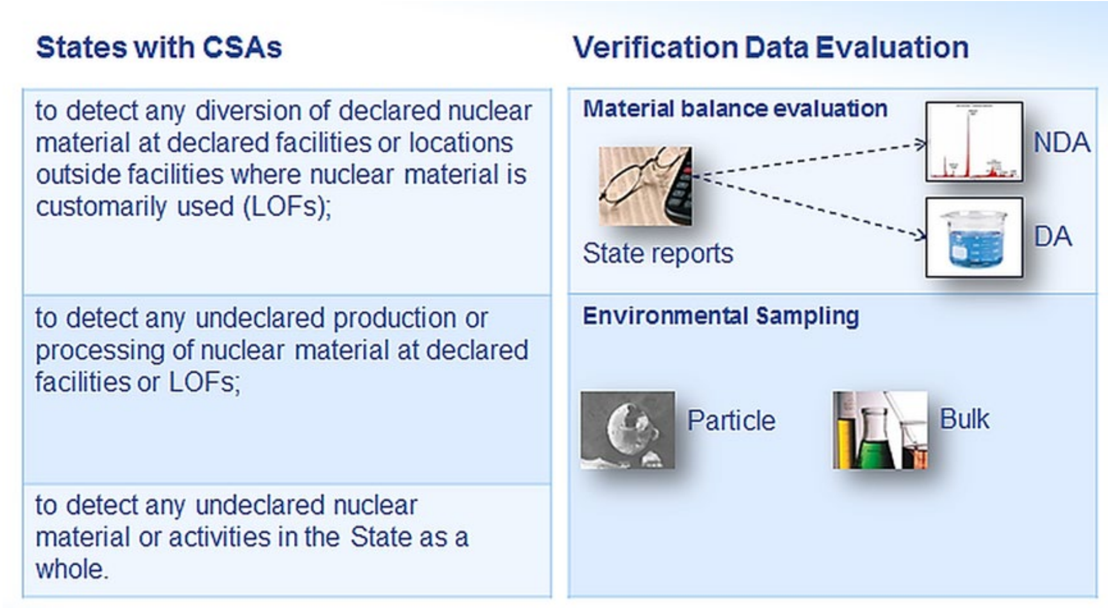


Fig.2: Role of verification data evaluation in supporting safeguards objectives under the State-level concept (example: States with a comprehensive safeguards agreement (CSA)).

it requires different expertise profiles. Its purpose is to confirm that NFC facilities are operated as declared, that there are no undeclared nuclear materials or activities in these facilities and, within the limits of its implementation modalities, that there are no undeclared nuclear materials or activities in the State as a whole. The principle of ES rests on the premise that nuclear processes release traces of nuclear and other material that constitute a signature of these processes and that can be transferred to samples collected at appropriate places. The characteristics of materials found on swipe samples (e.g. isotopic ratios, association with radionuclides or other elements) are compared with those predicted by specialized process-modelling tools. Particle analysis methods rely on the detection and measurement of individual nuclear material bearing particles on the sample. Bulk analysis methods involve the analysis of an entire swipe sample - in this case, the analytical results represent average values associated with the nuclear material contained within the sample [7].

ES was implemented in the context of strengthening the effectiveness of the safeguards system following the discovery of Iraq's clandestine nuclear weapons programme in the aftermath of the 1991 Gulf War. Its feasibility and detective power were established through a series of field trials in the context of the *Programme 93+2* with the support of Member States. Analytical laboratories that would later form the basis of the present international network of analytical laboratories (NWAL) demonstrated their capability to perform the extremely low-level radiochemical and isotopic measurements needed for the analysis of environmental samples. The field trials also showed that swipe sampling is the preferred method and it is now the standard, although other types of samples may be collected according to the technical objective pursued. For example, small quantities of ore or other compounds are regularly collected for material characterization as described below.

Since 1995, ES samples have been taken at locations where IAEA inspectors have access during inspections and design information verifications (DIV) and, following the approval of the Additional Protocol by the IAEA Board of Governors in 1997, ES can be taken at a broader range of locations in States where an AP is in force. ES has expanded over the years to include all NFC facility types and the number of ES collected increased steadily to reach the current number of up to ~400 samples per year. Sub-samples are distributed to the NWAL, which presently includes 21 laboratories in 8 States in addition to two European Commission Joint Research Centers and the IAEA safeguards analytical laboratory (SAL) in Seibersdorf, Austria.

ES continues to evolve through scientific and technical developments supported by Member States' laboratories in close collaboration with the IAEA. Technical Meetings are held every year, alternatively focusing on bulk or particle analysis, to review technological advances, among other objectives, and to discuss potential developments with representatives of the NWAL. For example, age dating [14,15,16] makes it possible to establish the chronology of certain processes based on the isotopic composition of plutonium bearing particles. Age dating of uranium bearing particles based on thorium in-growth would require an improvement of the sensitivity of laboratory analyses but is also of high interest for potential future applications. Another promising development field is nuclear material characterization (aka impurity analysis) [17, 18], which associates samples of ore and other uranium compounds with signatures in terms of the trace elements they contain (for example lanthanides). These signatures, compared with global databases currently being populated, can be used to determine the origin of these materials by applying specialized statistical algorithms. Trace element fingerprints can also provide information about processes the material may have undergone. More generally, stable chemical elements in nuclear material bearing particles could reveal chemical signatures associated to processes such as reprocessing or enrichment. The feasibility and technical requirements of such evaluation methods are currently being investigated. An existing routine application of impurity analysis is to determine if the purity of the material sampled is suitable for fuel fabrication or isotopic enrichment and hence, if it should be subject to nuclear material accountancy measures under article 34 (c) of INFCIRC/153 (Corr.).

Since they have been developed in the wake of the strengthened safeguards system and in synergy with the evolution of safeguards concepts in the last decades, ES evaluation processes and deliverables are well integrated in the present SLC system. ES evaluation reports are delivered to Operation Divisions at both sample and State level according to increasingly performant time targets. Weekly performance indicators are regularly issued to monitor the timeliness of the process and the number of ES samples evaluated in different categories. The ES evaluation processes are effectively and efficiently supported by a state-of-the art ES database, automated report generation tools and regularly upgraded expert NFC modelling tools. This advanced IT environment makes it possible to compare the characteristics of isotopic species found in samples with those predicted by theoretical models and with isotopic species observed at other facilities worldwide. However, the unique expertise necessary for ES evaluation is very rare and its application to safeguards requires a long on-the-job training period. Therefore, a well-thought-out long-term recruitment and training plan is needed to maintain an adequate level of professional capacity and capability in the ES evaluation area.

2.2 DA and NDA data evaluation – detection of nuclear material diversion

For their part, the NDA and DA data resulting from inspectors' verification sampling plans and combined with bulk measurements, i.e. weight and volume measurements, are compared with the State's NMA reports to detect diversion through the material balance evaluation (MBE) process. MBE is a complex analytical activity which assesses and combines all quantitative declared information and verification results. In particular, at bulk handling facilities (BHF) where material is processed in loose forms (gases, liquids, powders), complex measurement systems are needed to establish the flows and inventories of material. The conclusions regarding material balances rest on resource-intensive statistical and metrological analyses based on the estimation and propagation of measurement uncertainties into uncertainties associated to balance statistics. The objective of these analyses is to determine if the BHF operators' imbalances and the differences between nuclear material amounts declared by operators and measured by inspectors can plausibly be explained by legitimate measurement errors and, hence, to draw conclusions on the absence of diversion from these facilities.

In contrast with ES data evaluation, MBE was developed at a much earlier stage of the safeguards' history and is rooted in the criteria-driven, facility-based approach which has long underpinned the IAEA's conclusions. While MBE principles and methodologies remain generally valid in the framework of a State-level evaluation, their scope (previously restricted to material balance areas (MBA) within facilities) needs to be expanded to the analysis of the nuclear material flows, inventories and balances of the whole State, taking into account the increasing use of random inspection schemes in State level approaches (SLA) and the implications for the statistical analysis of data collected according to these patterns. In addition to this undertaking, which poses a number of methodological challenges, new approaches are needed to address increasingly large and diversified data flows, to optimize the distribution of limited MBE resources and to align them with the State-level technical objectives (TO) identified through the acquisition path analysis (APA) performed by the SEGs. In addition, MBE results need to be consolidated and compared with information from other sources. Last but not least, considerable progress was made in the field of IT and statistical methodologies since MBE was first developed several decades ago. The current migration of the safeguards Departmental IT platform under the Modernization of Safeguards Information technology (MoSaIc) project provides a unique opportunity to adapt and evolve methodologies and to integrate them into new software tools.

An additional and stringent practical challenge is to effectively address these development needs under a static budget with a small group of statistical analysis professionals whose primary mission is to deliver timely input to safeguards approaches, evaluations and conclusions for all States with extended NFCs. Priority mandates also include a substantial support to the IAEA verification activities under the Joint Comprehensive Plan of Action (JCPOA) in Iran. Furthermore, evolving evaluation approaches and processes make it necessary to regularly communicate and collaborate with stakeholders within and outside the Safeguards Department through the organization of training and liaison actions. A fruitful project to evolve safeguards verification data evaluation must therefore rest on a well-structured and synergic strategy, based on a clear long-term development plan and taking into account manpower limitations while making the best use of available extra-budgetary support, e.g. in the form of Member State Support Program (MSSP) human resources and expertise. The strategy implemented by the NFC information analysis Section since its creation in July 2011 and illustrated schematically in Fig. 3 is articulated around a set of components whose common objective is to promote and provide new types of evaluation reports designed to effectively support the work of SEGs in drawing sound safeguards conclusions:

Fig.3: Organization and components of the NFC Information Analysis Section strategy to evolve verification data evaluation under the State-level concept

Quite evidently, the starting point of any strategy, as represented at the top of the diagram is to ensure sufficient human resources (HR) both in terms of manpower and expertise. The first implementation phase of the project therefore consisted in rebuilding a team of competent statistical data evaluators after the Safeguards Department capability and capacity in this field had virtually vanished following retirements and rotation of long-standing specialized staff. This was achieved through an extensive recruitment and training campaign completed in 2013 and 2014. However, maintaining adequate staffing, based on a regularly reviewed succession plan, remains a continuous effort, given the current shortage of adequate expertise on the world market.

In order to address the methodological component of the project and to foster new ideas, a biennial Technical Meeting (TM) on Statistical Methodologies for Safeguards was initiated to establish an overview of the methodological landscape in this field, gather worldwide expertise in addressing current gaps and questions, draft recommendations around the high-level structure represented in Fig.4 below and build a network of specialists to remedy the lack of internal resources by identifying potential MSSP support tasks. The first TM was held in Vienna in October 2013.

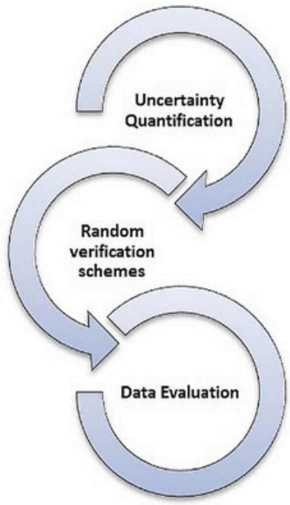
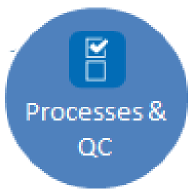
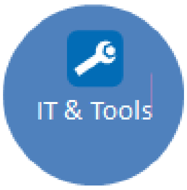


Fig.4: Three high-level interconnected methodological development areas as identified during the 1st TM on Statistical Methodologies for Safeguards (Vienna, October 2013).

Considerable progress, described in numerous publications [10], was made to date in the first two areas (uncertainty quantification and random verification schemes) and led to the preparation of several new safeguards technical reports (STRs), thanks to extensive MSSP support in the form of cost free experts (CFEs) and individual support tasks. The next phases planned include the harmonization of uncertainty quantification terminology between safeguards partners (evaluators, facility operators, laboratories) in preparation of the periodic review of international target values (ITV -2020) as well as a methodological consolidation of random inspections schemes. These topics will be the focus of the 3rd TM in October 2017. On completion of the prerequisite methodological work on uncertainty quantification and random verification schemes, the final phase will consist in reviewing and upgrading data evaluation methodologies which constitute the cornerstone of the overall project.



In parallel to the methodological review, evaluation processes and procedures are being adapted to the Departmental organisation which supports the work of the SEGs under the State-level concept. Process improvements were implemented in coordination with Operation Divisions in order to optimize both timeliness and quality based on available resources. Direct collaboration with inspectors in the framework of State-level approaches have significantly increased as well as in-field integration of evaluator expertise through their participation in inspections and design information verification (DIV) activities. This has greatly improved communication and collaboration between inspectors and evaluators and, in some cases, has allowed the resolution of long-standing issues. Quality control (QC) continues to be an essential component of the data evaluation activities and is now implemented at the level of the source data, of the evaluation process and of the resulting conclusions, by systematic peer-review, and by an additional review by inspectors in charge of facilities and States to ensure that all in-field and operational information has been taken into account.



In the context of the re-engineering and integration of safeguards databases and software under the MoSaIc project and their migration into the secure integrated safeguards environment (ISE), all legacy software that was developed over the last decades to support statistical analysis, e.g. sampling plans, verification performance evaluation, analysis of DA sample results, and MBE, are also being re-engineered and integrated under the Statistical Testing, Evaluation and Planning for Safeguards (STEPS) project. The STEPS project is designed to take into account both methodological and best practise developments and is expected to substantially increase the efficiency of the evaluation processes through the automation of calculations, QC checks and report generation.

In the framework of the State-level concept, operations inspectors and safeguards analysts need to understand and consolidate conclusions from many different sources of information. A structured programme of seminars is organised by the NFC Information Analysis Section to ensure effective communication with safeguards analysts from different areas and with Operation inspectors. These seminars address the mathematical rationales underlying safeguards verification strategies as well as the statistical treatment of the quantitative data declared by NFC facility operators and collected by operations inspectors. Their objective is to present the mathematical and statistical methodologies applied in safeguards in a clear and progressive way, using a minimum of formalism and with special emphasis on practical examples taken from everyday safeguards experience.

In addition to training and regular liaison with IAEA partners, a valuable measure in monitoring the quality of NMA and verification data is a trilateral liaison framework [11] with the SRA[10] and facility operators to discuss MBE results for the elapsed material balance period, review trends in material balance statistics, investigate their causes and agree upon recommendations and possible remedial actions. When available, DA sample results from three laboratories (IAEA, RSAC/SSAC[11], facility operator) are also examined to identify biases and compare analytical uncertainties. Using not only IAEA's and operators' measurement results but also the SRA's results can help to investigate the source of significant pairwise differences of DA sample results. The cooperation of SRAs and facility operators with the IAEA in the framework of trilateral liaison meetings provides a useful mechanism to remedy any issue related to the quality of the operator's measurement systems before it becomes a safeguards concerns, thereby promoting a proactive rather than reactive approach. This considerably enhances safeguards effectiveness and efficiency since the root cause of NMA issues may be difficult to establish at a later point, when their effects on the material balance have reached a safeguards significant threshold. In several instances, yearly trilateral liaison meetings organized between the IAEA, the SRA and plant operators have noticeably improved the operators' accounting procedures and/or measurement performance. In addition, trilateral meetings considerably increase the quality of communications between safeguards partners by fostering direct contacts between IAEA, SRA experts and facility staff specialized in NMA and by making it possible to maintain continuity of knowledge on complex technical files in case of rotation of responsible staff on all sides. Given their in-depth knowledge of industrial processes, operational conditions and accounting systems, nuclear fuel cycle facility operators are often the most knowledgeable when it comes to identifying the source of procedural or measurement issues. A regular dialogue with them is an important confidence building measure that improves their understanding of safeguards objectives and practices and engages them to willingly cooperate in ensuring the performance of the facility's accounting and measurement system.

As was commented above, the bases for evolving DA, NDA and MBE data evaluation reports and designing new report types were laid by the NFC Information Analysis Section as a keystone and convergence point since the strategy described in this paper was first implemented. However, the deployment of new reports is progressive and depends on the development stage of the project components described above. The central challenge is to design a concept addressing the complexity of MBE at State level while *optimizing its effectiveness* at detecting diversion and/or misuse at *key points* of the State nuclear fuel cycle. This paragraph describes some of the main guiding principles, i.e. a) evolution from a facility oriented approach to a State-level approach b) integration of

[10] State or regional authority responsible for safeguards implementation.

[11] State/regional Systems of Accounting and Control.

the *Physical Model* [8], as a backbone of the method, to support flow analysis and information consolidation; c) use of modern *visualization tools* to extract significant facts and patterns and identify potential inconsistencies in growing volumes of data [6].

The table in Fig. 5 compares the main features of the new data evaluation reports with the former facility-oriented concept:

BEFORE	NOW
• Facility based evaluation	• Facility and State evaluation.
• For BHF holding more than 1 SQ	• For any facility if relevant.
• Fixed criteria (e.g. detection of 1 SQ)	• Focus on technical objectives.
	• Enhanced analytical content.
• Not fully developed in SER	• Supports SIR and SER.
• SIR "crunch" – 60 facilities (~210 MBE) in 2.5 months. Needs:	• Integrated in SE cycle:
✓ All BHF CIR Part I completed.	✓ effort spread over the year.
✓ All NDA reported and QC.	✓ improved quality.
✓ All DA reported and QC.	✓ More sustainable under reduced resources.
✓ Uncertainties (RSD) actualized.	• Combined with enhanced liaison.

Fig.5: Evolution from a facility oriented approach to a State-level approach

In addition to providing a solution to resource limitations related to internal processes and timetables, the highlight of this new evaluation approach is that it is in line with one of the main tenets of the SLC, i.e. it addresses specific technical objectives (TO) resulting from the SEGs' APA and makes it possible to focus analytical resources on these TO as opposed to systematically checking a certain number of predetermined criteria. For example, while MBE evaluation was performed in the past for BHF holding more than one significant quantity (SQ) only, it can now be performed for any facility in agreement with the SEG if this is considered relevant to an identified acquisition path. Conversely, although it is important to mention that all large BHF will continue to be subject to MBE, the thoroughness of the evaluation may be adapted to prioritize analytical resources in case diversion during a given material balance period was covered by effective and conclusive measures (e.g. C/S), making MBE redundant, or in case the effectiveness of MBE is insufficient (e.g. low detection probabilities due to very large material flows/inventories).

The key principle of the method consists in visually representing nuclear material flows on a backdrop structure based on the Physical Model (PM) as shown in Fig 6. It can be outlined as follows:

- Facilities are represented by boxes grouped according to their function in the State nuclear fuel cycle (stages of the PM).
- For a period that can be customized by the user, nuclear material flows between facilities are visualized by solid curves whose colour represents material types and whose width is proportional to their magnitude (normalized in SQ), which can be read from the tick marks on the PM separation lines.
- Beginning and ending inventories are represented according to the same scale convention.
- Flows into and out of the States are symbolized by ellipses.

The APAs developed by SEGs identify paths, steps and the corresponding TOs which involve diversion or misuse of nuclear material at declared facilities. This makes it possible, as described above, to align data evaluation efforts with the results of the APA, taking into account the other safeguards measures foreseen by the SLA. In addition, operational links between facilities that can influence specific MBE statistics and their trends are emphasized and integrated in the data evaluation. Initial EXCEL-based prototypes (2011) and later automated trials (2013) performed in collaboration with SEGs demonstrated that the interest of the nuclear material flow diagrams underlying this method –referred to as Sankey diagrams[12] or "Snakeys" in reference to their sinuous appearance (Fig.6 below) - go beyond data evaluation and can usefully support the general work of SEGs, *inter alia*, the APA itself. The method has now evolved from the key elements described above to include a number of interactive features which support the current Departmental evolution from paper to electronic deliverables. In addition, the original concept is designed to integrate other types of relevant information (e.g. APA, SLA as well as ES, NDA and DA verification results). It is envisioned that, in future, it could serve as a possible portal to safeguards information in a State seen from a nuclear material perspective.

3. Conclusion

A structured, comprehensive and synergic long-term strategy is implemented by the Department of Safeguards' Division of Information Management NFC Information Analysis Section to evolve the evaluation of verification data in order to ensure the integration of its concepts, methods and processes into the SLC framework while optimizing its effectiveness in detecting undeclared nuclear material and activities and diversion of nuclear material at declared facilities. The present paper presents the complementary and mutually supporting components of this strategy, which converge towards the promotion and provision of new types of data evaluation reports designed to better support the work of SEGs.

An essential and innovative feature of this new generation of safeguards data evaluation reports is that it utilizes the power of modern IT, which allows interactivity, supports the Department's evolution to secure electronic deliverables and takes advantage of data visualization to complement the limited capacity of the human brain to extract useful and relevant information from large volumes of data.

[12] Sankey diagrams are named after Irish Captain Matthew Henry Phineas Riall Sankey, who used this type of diagram in 1898 in a classic figure (see panel on right) showing the energy efficiency of a steam engine (from Wikipedia)

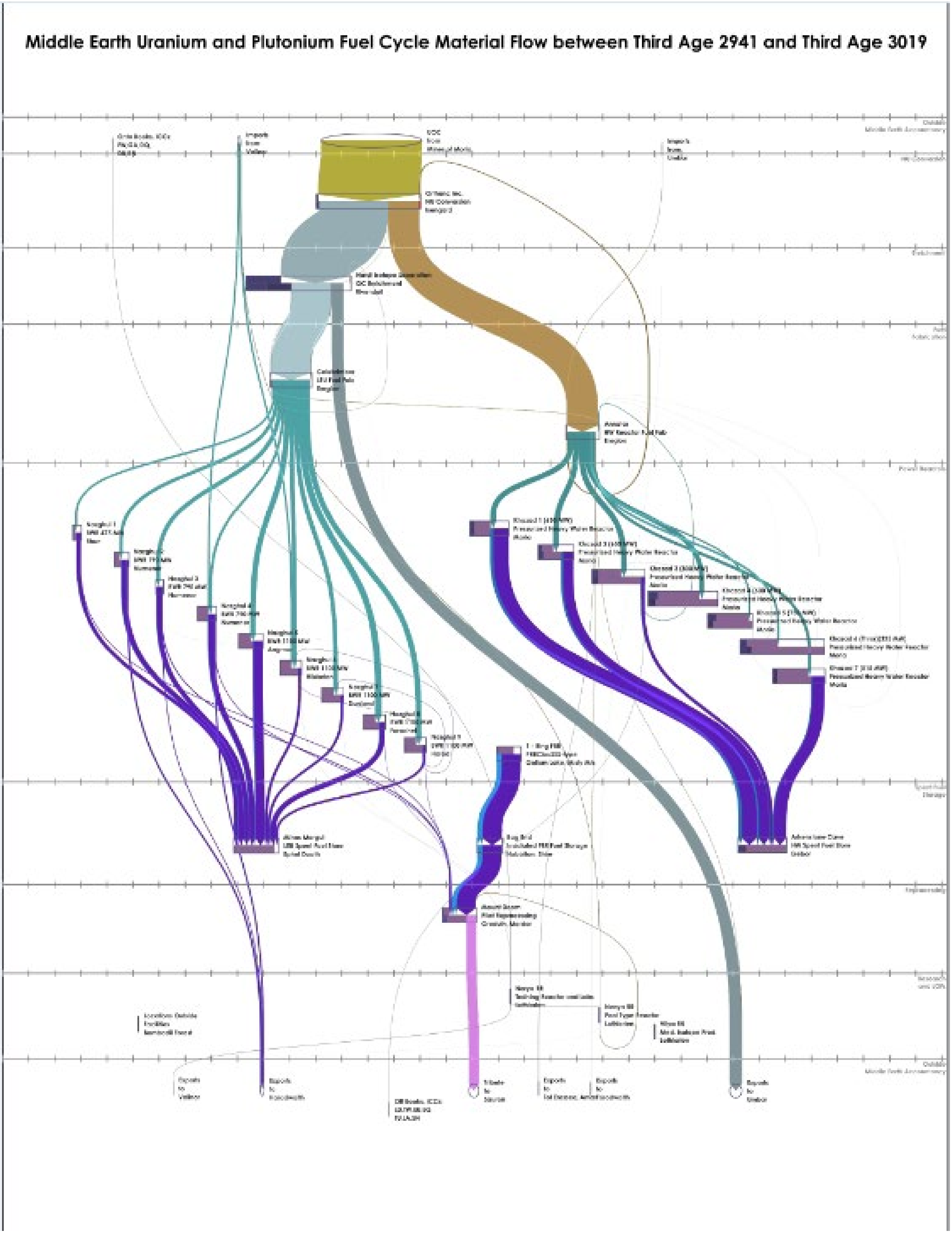


Fig.6: Snapshot of a nuclear material flow “Snakey” diagram for a hypothetical State

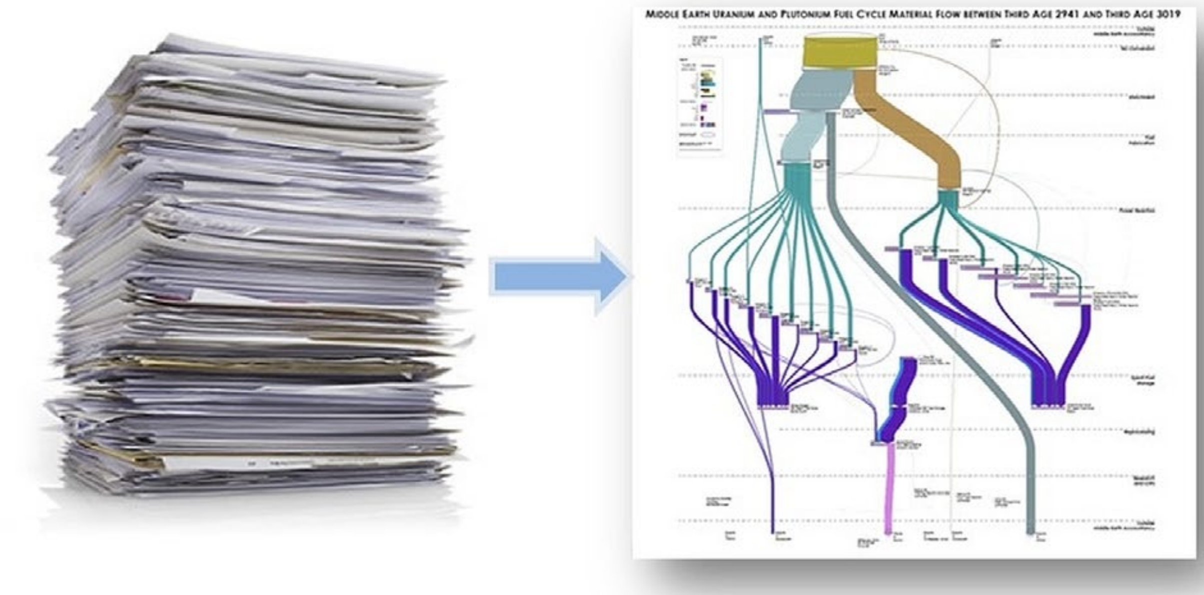


Fig.7: Data visualization can help analyze and understand large volumes of data

4. References

[1] Norman C, Barletta M, Ferguson M; *Collaborative Analysis in Support of State Evaluations*, *IAEA-CN-184/268*; *Proceedings of the IAEA International Safeguards Symposium: Preparing for Future Verification Challenges;* International Atomic Energy Agency; Vienna, Austria; 2010.

[2] Ferguson M, Norman C; *All-Source Information Acquisition and Analysis in the IAEA Department of Safeguards*; *Proceedings of the INMM 52nd Annual Meeting*; Palm Desert, CA, USA; 2011.

[3] Baute J; *Information Management Supporting the Evolution of the State-Level Concept; Proceedings of the INMM 53rd Annual Meeting*; Orlando, FL, USA; 2012.

[4] Norman C, Barletta M, Ferguson M; *Collaborative Analysis in Support of State Evaluations; Proceedings of the IAEA International Safeguards Symposium: Preparing for Future Verification Challenges;* International Atomic Energy Agency; Vienna, Austria; 2010.

[5] Norman C, Zhao K, Baute J; *Nuclear fuel cycle verification data and the state evaluation process: challenges and opportunities*; *Proceedings of the INMM 54th Annual Meeting*; Palm Desert, CA, USA; 2013.

[6] Baute J, Norman C, Binner R, Walczak-Typke A, Caillou F, Zhao K, Bonner E; *Dynamic Exploratory Visualization of Nuclear Fuel Cycle Verification Data in Support of the State Evaluation Process*; *Proceedings of the INMM 56th Annual Meeting*; Indian Wells, CA, USA; 2015.

[7] International Atomic Energy Agency; *STR-34: Environmental Sampling for Safeguards*; Vienna; 2011.

[8] International Atomic Energy Agency; *STR-325: Making use of the Physical Model*; Vienna; 2000.

[9] International Atomic Energy Agency; *The Structure and Contents of Agreements between the Agency and States in Connection with the Treaty on the Non-Proliferation of Nuclear Weapons: INFCIRC/153 (Corrected);* Vienna; 1972.

[10] Norman C, Krieger T, Binner R, Bonner E, Peter N, Portaix C, Richet S, Walczak-Typke A, Wüster J, Zhao K; *Outcome and Perspectives from the first IAEA International Technical Meeting on Statistical Methodologies for Safeguards*; *Proceedings of the IAEA International Safeguards Symposium: Linking Strategy, Implementation and People*; Vienna, 2014.

[11] Norman C, Iso S, Binner R, Zhao K, Portaix C; *Role and successes of trilateral liaison frameworks (IAEA – SSACs/RSACs – Nuclear Fuel Cycle Facility operators) in monitoring the quality of the operator's Measurement and Accounting Systems*; *Proceedings of the IAEA International Safeguards Symposium: Linking Strategy, Implementation and People*; Vienna, 2014.

[12] Fischer D, Ryzhinskiy M, Fuhr W, Higgy R, Nikkinen M, Vilece K; *Environmental sampling for detecting undeclared nuclear material/activities*; *Proceedings of the INMM/ESARDA workshop;* Santa Fe, USA; 2005.

[13] Bevaart L, Donohue D, Fuhr W; *Future requirements for the analysis of environmental samples and the*

evaluation of the results; *Proceedings of the ESARDA, 29th annual meeting: symposium on safeguards and nuclear material management;* Aix-en-Provence, France; 2007.

[14] Varga Z, Surányi G; *Production date determination of uranium-oxide materials by inductively coupled plasma mass spectrometry; Anal. Chim. Acta;* 599; 2007. 16–23.

[15] Shinonaga T, Donohue D, Ciurapinski A, Klose D; *Age determination of single plutonium particles after chemical separation*; *Spectrochim Acta B;* 64(1)*;* 2009. 95–98.

[16] Wallenius M, Tamborini G, Koch L; *The "age" of plutonium particles*; *Radiochim Acta;* 89; 2001. 55–58.

[17] Mayer K, Tushingham J, Boulyga S, Aregbe Y; *Report on the workshop on measurements of impurities in uranium*; *ESARDA Bulletin;* 43; 2009. 57–64.

[18] Bürger S, Boulyga SF, Peńkin MV, Bostick D, Jovanovic S, Lindvall R, Rasmussen G, Riciputi L; *Quantifying multiple trace elements in uranium ore concentrates: an interlaboratory comparison*; *J Radioanal Nucl Chem.* 301(3); 2014. 711–729.

Approximate Bayesian Computation Applied to Nuclear Safeguards Metrology

Tom Burr[1], Thomas Krieger[2], Claude Norman[1]

[1] Safeguards Information Management, IAEA, Vienna Austria
[2] Forschungszentrum Jülich, Inst. Energy and Climate Research, Jülich, Germany
E-mail: t.burr@iaea.org, t.krieger@fz-juelich.de, c.norman@iaea.org

Abstract:

Approximate Bayesian Computation (ABC) is an inference option if a likelihood for measurement data is not available, but a forward model is available that outputs predicted observables, such as gamma counts, for any set of specified input parameters, such as item mass. This paper reviews ABC and illustrates how ABC can be applied in safeguards metrology. A key aspect of metrology is uncertainty quantification (UQ), approached from physical first principles ("bottom-up") or approached empirically by comparing measurements from different methods and/or laboratories ("top-down"). Although ABC is not yet commonly used in metrology, an example using enrichment measurements is used to illustrate potential advantages in ABC compared to current bottom-up approaches. Using the same example, ABC is also shown to be useful in top-down UQ. And, the example shows good agreement between bottom-up and top-down measurement error relative standard deviation (RSD) estimates, while also allowing for the effects of item-specific biases. As a diagnostic, in applications of ABC, the actual coverages of probability intervals are compared to the true coverages. For example, if an ABC-based interval for the true measurement RSD is constructed to contain approximately 95% of the true values, then one can check whether the actual coverage is close to 95%. It is shown that one advantage of ABC compared to other Bayesian approaches is its apparent robustness to miss-specifying the model while maintaining good agreement between the nominal and the actual coverage.

Keywords: approximate Bayesian computation; metrology; non-destructive assay; uncertainty quantification

1. Introduction

Nuclear safeguards aim to verify that nuclear materials are used exclusively for peaceful purposes. To ensure that States honor their safeguards obligations, measurements of nuclear material inventories and flows are needed. Statistical analyses used to support conclusions require UQ, usually by estimating the RSD in random and systematic errors associated with each measurement method [1-9].

To monitor for possible data falsification by the operator that could mask nuclear material diversion, paired (operator, inspector) data are assessed. These paired data are declarations usually based on measurements by the operator, often using destructive assay, and measurements by the inspector, often using non-destructive assay (NDA). Statistical tests are applied one-item-at-a-time, and also to assess for a possible trend by computing the overall difference of the operator-inspector values using the

D statistic, one version of which is defined as $D = \frac{N}{n}\sum_{j=1}^{n}\frac{O_j - I_j}{O_j}$ where j indexes the sample items, O_j is the operator declaration, I_j is the inspector measurement, n is the verification sample size, and N is the total number of items in the stratum. The D statistic and the one-item-at-a-time tests rely on estimates of operator and inspector measurement error RSDs that are based on top-down UQ from previous inspections [1,2]. Inspector NDA measurements are made using portable neutron and gamma detectors taken into the facility, which involves challenges for UQ (Section 3). Such an assessment depends on the assumed measurement error model (for example, if the errors scale with the true value then a relative error model is appropriate) and associated uncertainty components, so it is important to perform effective UQ [2,3,4,8,9].

Another quantitative assessment in safeguards that requires UQ involves the material balance defined as $MB = T_{in} + I_{Begin} - T_{out} - I_{end}$, where T is transfers and I is inventory. The covariance Σ_{MB} of a sequence of n material balances is an n-by-n matrix with the MB variances on the diagonal and the covariances between pairs of MBs on the off-diagonals. The entries in Σ_{MB} are estimated using measurement error variance propagation applied to estimates of the RSDs in random and systematic error variances for each of the operator's measurement methods [1, 4-7].

MB evaluations and verification data assessments rely on estimates of measurement error RSDs. Historical paired (operator, inspector) data is used for top-down UQ, applying analysis of variance (ANOVA), to estimate RSDs. Bottom-up UQ propagates errors in all key steps of the assay to predict the RSD in the estimated nuclear material mass; this error propagation is similar to that used in the guide to expression of uncertainty in measurements [GUM, 10]. It is common for RSD estimates from bottom-up UQ to be smaller than those from top-down UQ [2, 9]. Currently, a gap between bottom-up and top-down RSD estimates does not directly impact inspectors' conclusions, because the top-down RSD estimates are used to set alarm thresholds in MB evaluations and verification data assessments. However, only when there is good agreement between bottom-up and top-down UQ can the potential to improve an NDA method be fully understood.

Because MB evaluations and verification data assessments rely on top-down estimates of random and systematic (see Section 2) measurement errors, top-down RSD estimates set a target for bottom-up UQ. One step to improve UQ is to improve bottom-up UQ so that its RSD estimates are in better agreement those from top-down UQ [2, 9]. Another step to improve UQ is to estimate the uncertainty in the RSD estimates so that any gap between bottom-up and top-down RSD estimates can be assessed for significance (which is not currently done in practice). Toward the goal of improving UQ, this paper introduces ABC for both bottom-up and top-down UQ. Any Bayesian approach provides a probability distribution for the unknown model parameters, which are the unknown random and systematic RSDs in this context, so the uncertainties in the RSD estimates are known. And, ABC has two potential advantages over other Bayesian methods in this context. First, ABC appears to be more robust to small or modest misspecifications of the data likelihood. Second, ABC can easily accommodate comprehensive bottom-up UQ, including effects such as uncertainties in nuclear data and model-based adjustment of test items to calibration items [9].

This paper is organized as follows. Sections 2 and 3 describe top-down and bottom-up UQ, respectively. Section 4 describes approximate Bayesian computation (ABC [11-13]). Section 5 applies ABC to top-down and bottom-up UQ for safeguards for NDA using the enrichment meter principle (EMP [14-16]). Section 6 is a summary.

2. Top-down UQ applied to paired (operator, inspector) data

An effective measurement error model must account for variation within and between groups, where a group is, for example, a calibration or inspection period. A typical model for relative errors for the inspector (I) (and similarly for the operator O) is

$$I_{jk} = \mu_{jk}(1+S_{Ij}+R_{Ijk}), \qquad (1)$$

where I_{jk} is the inspector's measured value of item k (from 1 to n) in group j (from 1 to g), is the true but unknown value of item k from group j, $R_{Ijk} \sim N\left(0,\delta_{RI}^2\right)$ is a random error of item k from group j, $S_{Ii} \sim N\left(0,\delta_{SI}^2\right)$ is a short-term systematic error in group j. To better understand Eq. (1), Fig. 1 plots 10 simulated values in each of 3 groups of $d=(O-I)/O$ values. Section 5.1 contains more information regarding Fig. 1.

The measurement error model sets the stage for applying ANOVA with random effects [17-19]. Neither $R_{I_{ij}}$ nor S_{I_i} are observable. However, for various types of observed data, one can estimate their respective variances δ_{RI}^2 and δ_{SI}^2. For the error model in Eq. (1), the standard deviation σ_D of D, is $\sigma_D = \sqrt{N^2\left(\frac{\delta_R^2}{ng}+\frac{\delta_S^2}{g}\right)}$ where $\delta_R^2=\delta_{RO}^2+\delta_{RI}^2$ and $\delta_S^2=\delta_{SO}^2+\delta_{SI}^2$, so alarm thresholds for D that correspond to user-specified false alarm probabilities can be selected. Similarly, the one-at-a-time tests also require estimates of δ_R^2 and δ_S^2, which are obtained by applying random one-way ANOVA to real paired difference data that are assumed to follow Eq. (1). Reference [3] evaluates impacts on alarm probabilities of using estimates of δ_R^2 and δ_S^2 instead of the true quantities In some safeguards contexts such as MB evaluation, the estimates of δ_R^2 and δ_S^2 must be partitioned into δ_{RO}^2 *and* δ_{RI}^2 and δ_{SO}^2 *and* δ_{SI}^2, respectively [2, 3]. Note from the expression for σ_D that δ_R^2 is divided by the number of observations ng, and that δ_S^2 is divided by the number of periods g, which makes sense according to the error model (1) and in view of Figure 1.

Error model (1) does not include long-term systematic error. The short-term systematic error is assumed to change between inspection periods [14,19] due to re-calibration and possibly other effects. In practice, there are sometimes tests for long-term systematic error, where long-term means as long as (or longer than) the data evaluation period, which is typically multiple inspection periods or years. Any long-term error is investigated and will be assumed in this paper to be zero.

3. Bottom-up UQ

NDA uses calibration and/or modelling to infer nuclear material (NM) mass using detected radiation such as neutron and gamma emissions. Three issues in UQ for NDA are:

1. NDA is applied in challenging settings because the detector is brought to the facility where ambient conditions can vary over time, and the items are often heterogeneous in some way. Because of such challenges, dark uncertainty [20] can be large, as is evident whenever bottom-up UQ predicts smaller RSD than is observed in top-down UQ.
2. There is no UQ guide for NDA that is analogous to the GUM. But, the GUM is typically followed for the error variance propagation steps in UQ, and each NDA method has a specific and documented implementation of UQ (for example, ASTM C1514 [15] for the EMP).
3. NDA is often used when test items differ substantially from calibration items; therefore, the concept of item-specific bias is important, and is addressed in Section 5.

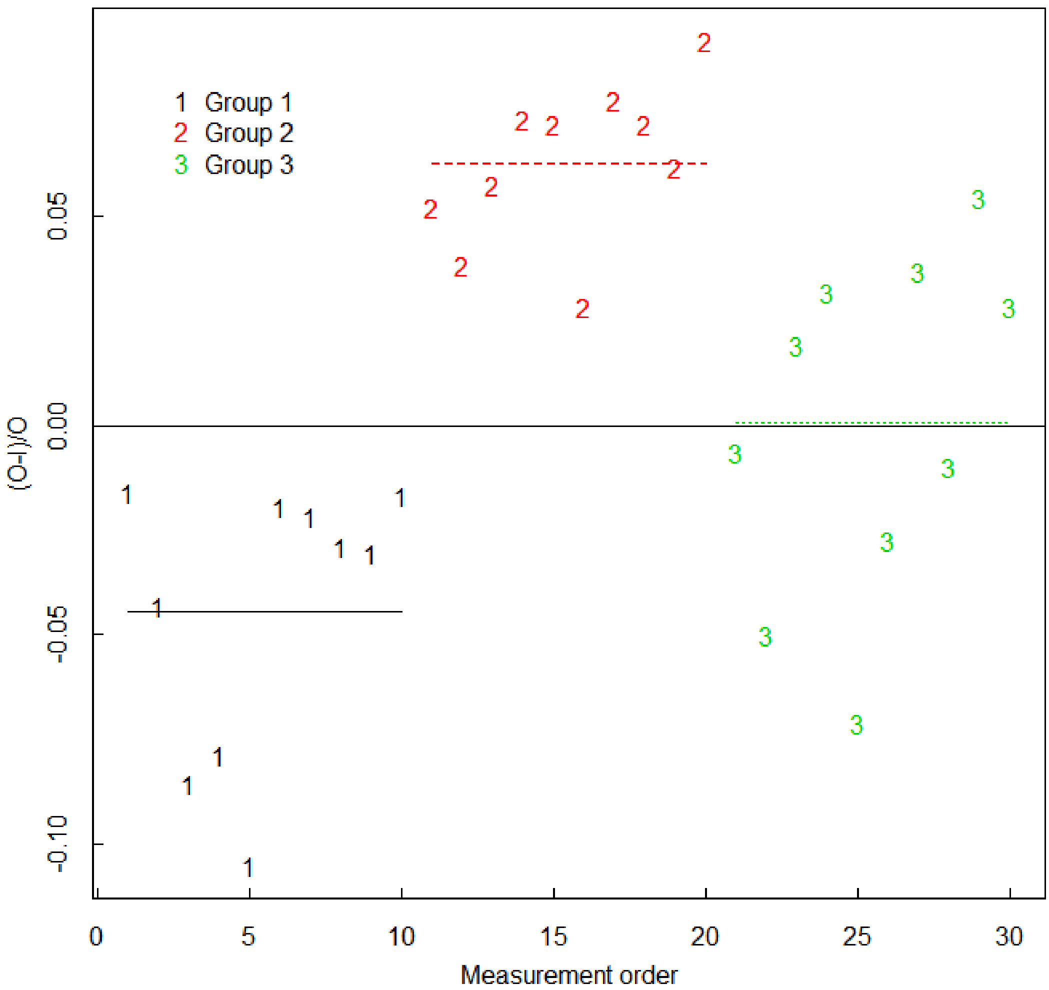


Figure 1: Example (simulated) of 10 $d = (O - I) / O$ values in each of 3 groups.

In NDA, error variance propagation is used as a component of bottom-up UQ by propagating errors in inputs. Bottom-up UQ is often approached by using the GUM's measurement equation, expressed as

$$Y = f(X_1, X_2, \ldots, X_N) \tag{2}$$

for measurand Y and inputs $X_1, X_2, \ldots, X_N$. The GUM applies the delta method to Eq. (2) to propagate error variances in the X_i to estimate the standard deviation in Y. The input quantities can include, for example, measured count rates, estimates of calibration parameters or other measurands, such as measured values in steps an assay method. The delta-method assumes that $f(X_1, X_2, \ldots, X_N)$ in Eq. (2) can be well approximated by a first-order Taylor series expansion around the mean values of each, and then the linear approximation to $f(X_1, X_2, \ldots, X_N)$ can be used to estimate σ_Y^2 given estimates of the variances for each X_i (and, correlations between the can be accommodated). If the first-order Taylor series is not sufficiently adequate, the GUM recommends Monte Carlo simulation. Note that Eq. (2) implies that Y is random, so the GUM implicitly adopts a Bayesian viewpoint (Section 4) without explicitly stating a prior distribution for Y [21, 22].

Recently, the NDA community is recognizing a need for more comprehensive bottom-up UQ that thoroughly addresses uncertainty in model-based adjustments of test items to calibration items [2,9]. Toward that goal, several US national laboratories are collaborating on a multi-year project to improve UQ for NDA and the standard committee ASTM C26.12 is another group also working on UQ for NDA. One possible outcome of these collaborations is better guidance on bottom-up UQ for calibration data that allows for both errors in predictors and for item-specific bias. It is also possible that approaches for better bottom-up UQ will be provided in the next version of the GUM [21, 22].

4. ABC

Bayesian ANOVA such as could be applied to data generated from Eq. (1) has been studied [17], and Bayesian methods are slowly being adopted in metrology [9,10,21,22]. However, Bayesian ANOVA using ABC has not been well studied. In any Bayesian approach, prior information regarding the magnitudes and/or relative magnitudes of δ^2_{RI} and δ^2_{SI} can be provided [21-23]. If the prior is "conjugate" for the likelihood, then the posterior is in the same likelihood family as the prior, in which case analytical methods are available to compute posterior prediction intervals for quantities of interest. In order that a wide variety of priors and likelihoods can be accommodated, modern Bayesian methods do not rely on conjugate priors, but use numerical methods to obtain samples of δ^2_{RI} and δ^2_{SI} from their approximate posterior distributions [23]. For numerical methods such as Markov Chain Monte Carlo [23], the user specifies a prior distribution for δ^2_{RI} andδ^2_{SI}, and a likelihood (which need not be normal). ABC does not require a likelihood for the data (but this section provides clarification regarding the need for a likelihood in this NDA context), and, as in any Bayesian approach, ABC accommodates constraints on variances through prior distributions [11-13, 24-26].

The "output" of any Bayesian analysis is the posterior distribution for each model parameter, and so the output of ABC for data generated from Eq. (1) is an estimate of the posterior distributions of δ^2_{RI} and δ^2_{SI}. No matter what type of Bayesian approach is used, a well-calibrated Bayesian approach satisfies several requirements. One requirement is that in repeated applications of ABC, approximately 95% of the middle 95% of the posterior distribution for each of δ^2_{RI} and δ^2_{SI} should contain the respective true values. That is, the actual coverage should be closely approximated by the nominal coverage. A second requirement is that the true standard deviation of the ABC-based estimates of δ^2_{RI} and δ^2_{SI} should be closely approximated by the standard deviation of the ABC-based posterior distributions of δ^2_{RI} and δ^2_{SI}. Inference using ABC can be briefly summarized as follows:

ABC can be described using this high-level algorithm description:

ABC Inference

For i in 1, 2, …, N

1. Sample θ from the prior, $\theta \sim f_{prior}(\theta)$.
2. Simulate data y' from the model $y' \sim P(y \mid \theta)$.
3. Denote the real data as y. If distance $d(S(y'),S(y)) \le \varepsilon,$ *accept* θ *as an observation from* $f_{posterior}(\theta \mid y)$.

Experience with ABC suggests that the ABC approximation to $f_{posterior}(\theta \mid y)$ improves if step (3) is modified to include a weighting function so that values of $\theta \sim f_{prior}(\theta)$ that lead to very small values of the distance $d(S(y'),S(y))$ are weighted more heavily in the estimated posterior [24,25].

In ABC, the model has input parameters θ and outputs data $y(\theta)$ and there is corresponding real data y_{obs}. For example, the model could be Eq. (1), which specifies how to generate synthetic *I* (or *O*) data, and does require a likelihood; however, the true likelihood used to generate the data need not be known to the user. Synthetic data is generated from the model for many trial values of θ, and trial θ values are accepted as contributing to the estimated posterior distribution for $\theta| y_{obs}$ if the distance $d(y_{obs},\ y(\theta))$ between y_{obs} and $y(\theta)$ is reasonably small. Alternatively, for most applications, it is necessary to reduce the dimension of y_{obs} to a small set of summary statistics $S(y_{obs})$ and accept trial values of θ if $d(S(y_{obs}),S(y(\theta)) < \varepsilon$, where ε is a user-chosen small threshold near 0. Here, for example, $y_{obs} = d = \dfrac{O-I}{O}$ data in each inspection group, and $S(y_{obs})$ includes within and between groups sums of squares Specifically, the ANOVA-based estimator of δ^2_{RI} in Eq. (1) is $\hat{\delta}^2_R = \dfrac{1}{n-g}\{\sum_{j=1}^{g}\sum_{k=1}^{n}(d_{jk}-\bar{d}_j)^2\}$,and the usual estimate of δ^2_{SI} is $\hat{\delta}^2_S = \dfrac{\sum_{j=1}^{g}(\bar{d}_j-\bar{\bar{d}})^2}{g-1} - \dfrac{\hat{\delta}^2_R}{n}$. The quantities $\hat{\delta}^2_R$ and $\hat{\delta}^2_S$ are therefore good choices for summary statistics for ABC. Recall that because trial values of θ are accepted if $d(S(y_{obs}),S(y(\theta)) < \varepsilon$, an approximation error to the posterior distribution arises that several ABC options attempt to mitigate. Recall also that such options weight the accepted θ values by the actual distance $d(S(y_{obs}),S(y(\theta))$ (abctools [25] in R [26]).

To summarize, ABC applied to data following Eq. (1) consists of three steps: (1) sample parameter values of δ^2_R and δ^2_S from their prior distribution $p_{prior}(\theta)$; (2) for each simulated value of θ in (1),simulate data from Eq. (1); (3) accept a fraction of the sampled prior values in (1) by checking whether the summary statistics computed from the data in (2) satisfy $d(S(y_{obs}),S(y(\theta)) < \varepsilon$. If desired, aiming to improve the approximation to the posterior, adjust the accepted θ values on the basis of the actual $d(S(y_{obs}),S(y(\theta))$ value. ABC requires the user to make three choices: the summary statistics, the threshold ε, and the measure of distance *d*. Reference [11] introduced a method to choose summary statistics that uses the estimated posterior means of the parameters based on pilot simulation runs. Reference [12] used an estimate of the change in posterior $p_{posterior}(\theta)$ when a candidate summary statistic is added to the current set of summary statistics. Reference [13] illustrated a method to evaluate whether a candidate set of summary statistics leads to a well-calibrated posterior, in the same sense that is used in this paper; that is, nominal posterior probability intervals should have approximately the same actual coverage probability, and the posterior variance should agree with the observed variance in testing.

5. EMP Example

The mass of ^{235}U in an item can be estimated by using the measured net weight of uranium U in the item and the measured ^{235}U enrichment (the ratio ^{235}U/U). Enrichment can be measured using the 185.7 keV gamma-rays emitted from ^{235}U by applying the EMP. The EMP aims to infer the enrichment by measuring the count rate of the strongest-intensity direct (full-energy) gamma from decay of ^{235}U, which is emitted at 185.7 keV [14-16]. The EMP assumes that the detector field of view into each item is identical to that in the calibration items (the "infinite thickness" assumption), that the item is homogeneous with respect to both the ^{235}U enrichment and chemical composition, and that the container attenuation of gamma-rays is the same as or similar to that in the calibration items so that empirical correction factors have modest impact and are reasonably effective. If these three assumptions are met, the known physics implies that the enrichment of ^{235}U in the U is directly proportional to the count rate of the 185.7 keV gamma-rays emitted from the item. It has been shown empirically that under good measurement conditions, the EMP can have a random error RSD of less than 0.5 % and a long term bias of less than 1 %, depending on the detector resolution, stability, and extent of corrections needed to adjust items to calibration conditions. Some bottom-up UQ examples for the EMP in [14,16,19] have estimated random error RSD ranging from less than its 0.5% target to approximately 1.0% (because of item-specific biases arising due to container thickness variations and other effects,) but less than the 2% to 4% reported from corresponding top-down UQ for the ^{235}U mass in UO_2 drums. Also, top-down UQ reports total error RSD (random and short-term systematic) of 4% to 20 % for some items analyzed in [19] (the RSD tends to be larger for smaller values of enrichment).

The known nominal enrichment in each of several standards can be fit to observed counts in a few energy channels near the 185.7 keV energy as the "peak" region and to the counts in a few nearby energy channels below and above the 185.7 keV energy but outside the peak area to estimate background (two-region EMP method), expressed as

$$Y = \beta_1 N + R_Y \tag{3,}$$

where Y is the enrichment, N is the peak count rate near 185.7keV, R_Y is random error and β_1 is a calibration constant. Figure 2 is an example low-resolution (NaI detector) gamma spectrum near the 185.6keV region. The gross count and the two background ROI counts can be combined into one net count, resulting in one predictor as in Eq. (3). For example, if the same number of energy channels are used for both the peak and background ROI, then Net count rate = Peak count rate – Background count rate. There is usually non-negligible error in N, so errors in predictors cannot be ignored when fitting Eq. (3) to calibration data [14]. Alternatively, both peak and background counts can be used as predictors [14-16]. There will be measurement errors in the gross and background count rates and there will often be correction factors applied, for example, to adjust test item container thickness to calibration item container thickness. There is much literature regarding errors in predictors and whether to fit Y as a function of N (reverse calibration) or to fit N as a function of Y and invert to solve for Y (inverse calibration). Both options should be investigated using simulation, because analytical approximations have been shown to not be sufficiently accurate either to decide between options or to assess the uncertainty in the chosen option [14,27]. However, the root mean squared prediction error (RMSE) of reverse calibration (Eq. (3) is an example of reverse calibration) has been generally found to be the same as or smaller than that of inverse calibration.

Calibration data is used to compute the estimate $\hat{\beta}_1$ of the model parameter β_1 in Eq. (3). The variance of $\hat{\beta}_1$ is not necessarily well approximated by the usual least squares expression because of errors in N. Therefore, [14,27] suggest that the RMSE in $\hat{Y}$ be estimated by simulating the calibration procedure, which allows for errors in N arising from Poisson counting statistics, and also arising from other sources, such as container thickness (with or without an adjustment for the measured container thickness) varying among test items. Errors in N due to imperfect adjustment for container thickness can manifest as item-specific bias. The ABC strategy below illustrates how item-specific bias can be understood and estimated. The RMSE in $\hat{Y}$ is defined as usual, as $E((\hat{Y} - Y_{true})^2) = E(\hat{Y} - E(\hat{Y})^2) + (E\hat{Y} - Y_{true})^2 =$ = variance + bias^2.

Note that one can express the calibration Eq. (3) as in Eq. (2), where X_1 is $\hat{\beta}_1$ and X_2 is N, with var($\hat{\beta}_1$) estimated by simulation, so GUM's Eq. (2) could be used to estimate var($\hat{Y}_1$) and cov($\hat{Y}_1$, $\hat{Y}_2$) , although [22] points out that GUM's Eq. (2) is not actually designed to be applied to calibration applications, regardless of whether there are errors in the predictors.

In general, item-specific bias can arise due to item-specific effects, expressed as

$$CR/M = g(X_1, X_2, \ldots, X_N), \tag{4,}$$

where CR is the item's neutron or gamma count rate, M is the item NM mass, g is a known function, and $X_1, X_2, \ldots, X_N$ are N auxiliary predictor variables such as item density, source NM heterogeneity, and container thickness, which will generally be estimated or measured with error and so are regarded as random variables . To map Eq. (4), to GUM's Eq. (2), write

$$M = CR \,/\, g(X_1, X_2, \ldots, X_N) = h(X_1, X_2, \ldots, X_M) \tag{5,}$$

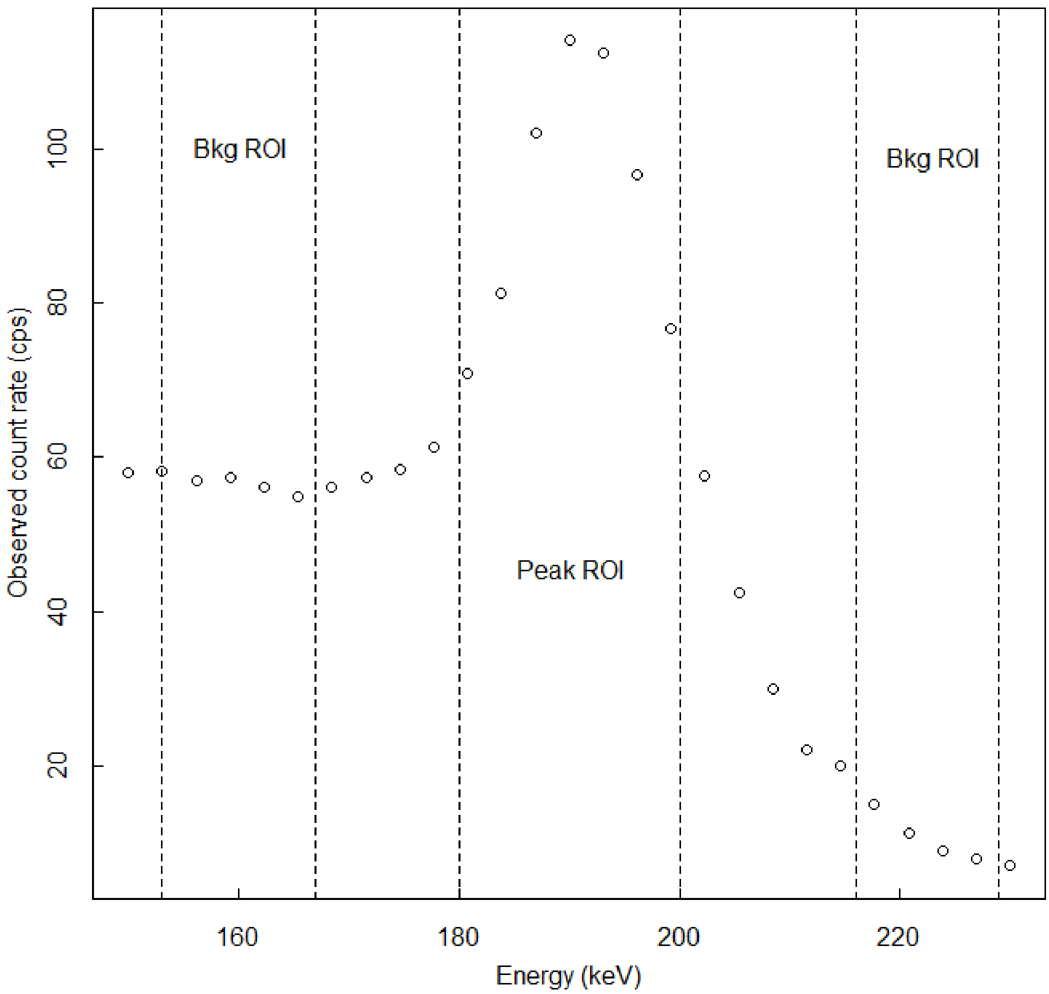


Figure 2: Example low-resolution (NaI detector) gamma spectrum near the 185.6keV peak with two background regions (one region below the 185.7 keV peak and one region above the 185.7 keV peak).

where the measured CR is now among the $M = N$+1 inputs. Note that Eq. (5) is the same as Eq. (4), but some of the X_i account for item-specific departures from reference items used for calibration. More specifically, Eq. (3) can be re-expressed as

$$Y = \beta_1(item)N + R_Y \qquad (6),$$

where the calibration constant $\beta_1(item)$ varies across items and R_Y is the random error in Y. Equation (6) is a random-coefficient regression equation, and real and/or simulated data generated from Eq. (6) can be used to estimate the average value of $\beta_1(item)$. Eq. (6) is a model that can explain item-specific bias, which is usually regarded as a random error (across items). Many NDA examples adjust test items to calibration items using some type of modelling [2,14]. In the EMP, an additional input variable X_3 could be an adjustment for container thickness to be applied to the detected net count rate in Eq. (6). And, one way to model the effect of imperfect adjustment for each item's container thickness is to include another random error in simulated net count rates used as synthetic calibration data, rather than to modify β_1. In practice, net count rates are sometimes adjusted to account for the measured container thickness, using Beer's law, which states that the gamma intensity after passing through a container with density ρ, attenuation coefficient μ and thickness t is multiplied by $\exp(-\mu\rho t)$). Note that errors in N have the same impact as errors in $\beta_1(item)$ because the term $\beta_1(item)N$ appears in Eq. (6).

5.1 ABC applied to the EMP

The purpose of this bottom-up example is to show how to apply ABC and to show how ABC makes a bottom-up estimate of random and systematic RSDs such as those illustrated in Figure 1, and how ABC includes uncertainty in the estimated RSDs. ABC applied to the EMP can be implemented in the following 7 steps.

1) Estimate the average regression coefficient $\hat{\beta}_1$ in Eq. (6) using available real calibration data, typically consisting of approximately 3 to 5 (Y,N) pairs. The real calibration data

used here are Y = 0.355, 0.80, 2.175, 3.305, 5.0 (^{235}U enrichments of 5 standards) and the corresponding N= 0.062, 0.139, 0.37, 0.575, 0.866 net count rates.

2) Use the estimate $\hat{\beta}_1$ from (1) to generate many (S = 10^5 or more) synthetic calibration runs using $Y = \beta_1(item)N + R_Y$ to generate synthetic sets of 5 paired (Y,N) values, with run i producing the estimate $\hat{\beta}_{1,i}$. This example generated the $\beta_1(item)$ values randomly and uniformly from 0.85 to 0.95.
3) Specify a prior distribution for the true enrichment μ_Y. If little is known about the true enrichment values, then, for example, specify a uniform prior ranging from the lowest possible true enrichment to the highest possible true enrichment. This example used wide a uniform distribution from 0.355 to 5.0, which avoids extrapolating outside the range of the true enrichments.

4) Specify a background count rate μ_B (this example used μ_B = 0.05) and use the estimated regression coefficient $\hat{\alpha}_1$ from the regression equation $N = \alpha_1(item)Y + R_N$ to generate a net count rate μ_N that corresponds to a value of μ_Y sampled from its prior distribution. This example used an RSD in Y of 0.1% and in R_N of 5%.

5) Specify a count time (this example used 600 seconds) t, simulate $B \sim Poisson(\mu_B t)$, $G \sim Poisson(\mu_G t)$, and compute a net count rate (assuming the same number of energy channels for the peak and background ROIs) $N = \frac{G}{t} - \frac{B}{t}$.

6) Repeat (4) and (5) many (10^5 or more) times to construct a large collection of simulated true enrichments μ_Y and corresponding net count rates N, which is an effective summary statistic.

(7) For each simulated test case, simulate a value of μ_Y from its prior, use steps (4) and (5) to generate N_{test}, and compute the distance $d(N_{test}, N_i) = |N_{test} - N_i|$ from N_{test} to each of the i = 1, 2, …, 10^5 realizations from step (6), and accept those μ_Y generated in step (6) that correspond to $|N_{test} - N_i| \le \varepsilon$ as observations from the posterior $\mu_Y | N$ (which in this case is somewhat complicated to specify analytically) weighting inversely by the distance $|N_{test} - N_i|$ if desired. Linear regression was not used in this ABC implementation for predicting μ_Y for each simulated test value of N, although it could have been, and note that regression is used in step (2) to generate the 10^5 pairs of (μ_Y,N) in the training data for ABC.

The result in applying steps 1-7 is an estimate of the posterior distribution for the true enrichment μ_Y, similar to that in Fig. 3, as explained below. To assess ABC performance, the two criteria mentioned can both be used: the estimated standard deviation of the posterior should be in good agreement with the observed standard deviation across test items, and the nominal probability interval coverage should also be in good agreement with the actual coverage. The data plotted in Fig. 1 were generated using the steps just given to apply ABC for both operator and inspector data, assuming for simplicity that both used the EMP and both recalibrated at the beginning of periods 1, 2, and 3. The estimated standard deviation of $d_{rel} = (O - I)/O$ (which includes both within- and between-group standard deviations) from top-down data such as that in Fig. 1 (also using ABC as outlined in Section 4) is 0.11, which is very close to that predicted from the bottom-up ABC (0.12 as explained in the next paragraph) posterior standard deviations for O and I.

Recall from Section 4 that the usual ANOVA-based estimator of σ_{Rd}^2 (using the multiplicative form of Eq. (1) for both operator and inspector) is $\hat{\sigma}_{Rd}^2 = \frac{1}{n-g}\{\sum_{j=1}^{g}\sum_{k=1}^{n}(d_{jk} - \bar{d}_j)^2\}$,

and the usual estimate of σ_{Sd}^2 is $\hat{\sigma}_{Sd}^2 = \frac{\sum_{j=1}^{g}(\bar{d}_j - \bar{\bar{d}})^2}{g-1} - \frac{\hat{\sigma}_R^2}{n}$.

The quantities, $\hat{\sigma}_{Rd}^2$ and $\hat{\sigma}_{Sd}^2$ are therefore good summary statistics for ABC, and were used to implement ABC for the top-down analysis of data such as that in Fig. 1.

The 0.12 bottom-up prediction for the standard deviation of $d_{rel} = (O - I)/O$ is illustrated by plotting the posterior for O for a particular N value in Fig. 3, which has a total (random plus systematic) RSD of 0.08 (from the 7-step procedure). Because this example assumes both O and I made the same type of EMP measurements, the bottom-up prediction of the RSD for $d_{rel} = (O - I)/O$ is given by $\sqrt{(0.08^2 + 0.08^2)} = 0.11$ (from bottom-up). The 0.12 top-down estimate of the RSD of $\delta_{d_{rel}}$ (see Fig. 4, using data such as the data in Fig. 1) is the RSD of the ABC-based posterior distribution for $\delta_{d_{rel}}$ from top-down UQ, with g = 3 groups and n = 10 paired measurements per group (as in Fig.1). The 0.12 estimate has an associated 14% RSD, and an approximate 95% probability interval for $\delta_{d_{rel}}$ is 0.086 to 0.15.

One advantage of having a probability interval for both the bottom-up and top-down estimate of $\delta_{d_{rel}}$ is that one can assess whether differences between the top-down and bottom-up estimates of $\delta_{d_{rel}}$ are significant. In this example, bottom-up UQ using ABC agrees very well with corresponding top-down UQ using ABC that used simulated O and I values as in Fig. 1; which means that in this application, ABC is well-calibrated. Trial and error was used to select $\varepsilon = 0.01$ to obtain good agreement between the ABC-based predicted standard deviation and the observed standard deviation. Coverages of the ABC-based probability intervals were checked and, as mentioned, excellent agreement between nominal and actual was observed. Specifically, the 99%, 95%, and 90% probability intervals contained approximately 99%, 95%, and 90%, respectively of the true values of μ_Y.

Because bottom-up RSD estimates are often compared to top-down RSD estimates to look for un-modelled effects ("dark uncertainty" [20]), it is important for RSD estimates to include information regarding uncertainty in the estimated RSDs. In this example, ABC provides estimates of the uncertainty in the parameter estimates (in this case, the estimated RSDs) in the same manner that any Bayesian analysis does, by providing a posterior distribution for each parameter. Because the top-down and bottom-up RSD estimates are essentially the same in this example, there is no evidence of dark uncertainty (and there should not be, because no dark uncertainty was simulated).

Assuming a normal distribution is not always a good approximation for the actual distribution of (*O-I*)/*O* values used in top-down UQ. So, regarding robustness of ABC in top-down UQ, it has been found that the actual coverages are essentially the same (to within simulation uncertainty) as the nominal coverages, at 90%, 95%, and 99% probabilities, for a normal distribution and all of the non-normal distributions investigated (uniform, gamma, lognormal, beta, *t*, and generalized lambda with thick or thin tails) for the distribution of the random error term R_Y in Eq. (6). Regarding robustness of ABC in the bottom-up context, a key aspect of ABC is the ease with which different forward models linking model parameters (such as the true RSDs in Eq. (2)) to model output and corresponding summary statistics. For example, the Poisson model used in the ABC implementation for the EMP can be easily replaced with an overdispersed Poisson model if exploratory analysis of real data suggests overdispersion.

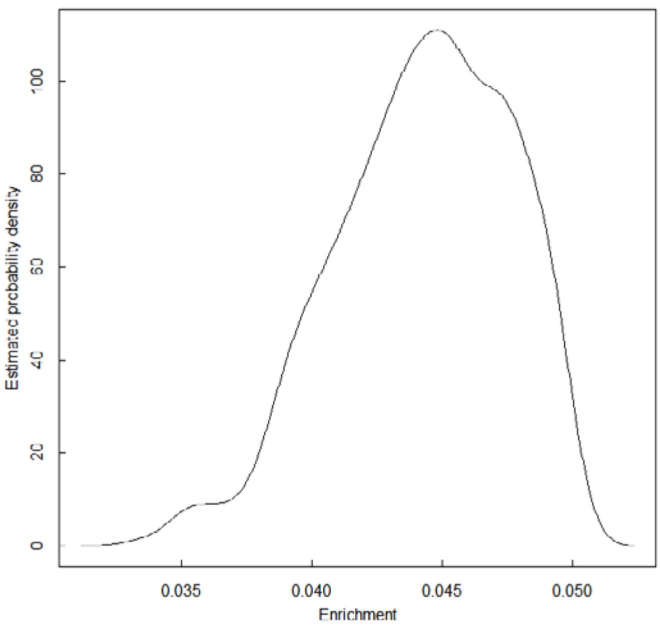


Figure 3: The bottom-up ABC-based estimate of the posterior δ_{TI} (or δ_{TO}).

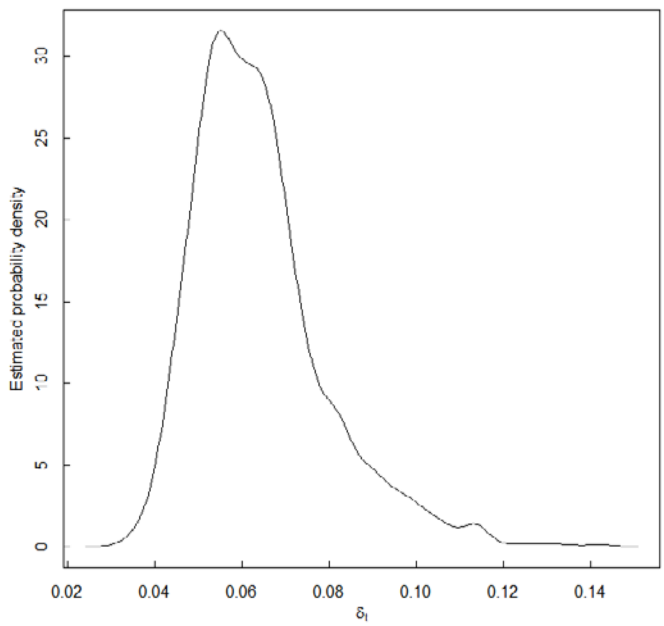


Figure 4: The top-down ABC-based estimate of the posterior for δ_T with RSD of 14%.

6. Discussion and Summary

ABC was used for both bottom-up and top-down RSD estimation in simulated EMP data (using a calibration set of 5 real EMP data pairs). ABC provided robust estimates of the posteriors for model parameters (the RSD values), so bottom-up RSD estimates could be compared to top-down estimates while accounting for parameter uncertainty (as defined by the width of the posterior).

ABC is very well-suited for bottom-up UQ in more challenging NDA applications, for example, when the measurement data is summarized using higher-dimensional summary statistics, such as the estimated net areas in peak regions of interest in gamma spectrometry [28,29], using microcalorimetry. Current microcalorimetry algorithms fit approximately 15 peak areas (associated with gamma ray energies) associated with different isotopes of Pu, U, and Am. These 15 peak areas are the summary statistics used in an ABC approach that requires a sophisticated forward model relating known isotope abundances to detected peak area [28,29]. The nuclear data that enter any analysis approach (ABC or other methods) include gamma emission energies, branching ratios, and half-lives. The branching ratios and half-lives determine the relative intensities of each peak for a given Pu isotopic fraction. Reference [29] indicates that uncertainties in emission energies are not as important in microcalorimetry as they are in lower resolution gamma spectroscopy such as that obtained in high-purity Germanium detectors, where spectral deconvolution is more challenging. ABC is compelling in spectrometry because ABC requires user-chosen summary statistics such as estimated peak areas, ABC can easily accommodate uncertainty in nuclear data, ABC can provide an estimate of the posterior distribution of each unknown parameter, including the unknown isotopic abundances. However, ABC requires a good-quality forward model linking the summary statistics to the isotopic abundances as well as to fundamental nuclear data that has recognized uncertainties.

Only when there is good agreement between bottom-up and top-down UQ can the potential to improve an NDA method be fully understood. Many NDA methods require calibration, so one type of bottom-up UQ involves calibration data. Although calibration might appear to be a simple application of regression, [9,14] illustrate that simulation is needed for effective bottom-up UQ in NDA because sample sizes are small, a ratio of random variables in the calibration analysis is used, and there are non-negligible error variances in predictor and response. In addition, calibration data should include item-specific effects that will be present in testing data. As illustrated for the EMP, ABC is a good tool for bottom-up UQ. Once improved bottom-up UQ is implemented, any remaining disagreement between bottom-up and top-down UQ could indicate, for example, that there are missing sources of uncertainty in bottom-up UQ [20], that the data and/or error model are not what are assumed, or that correlations among inputs in the measurand equation (Eq. (2)) are not adequately estimated.

ABC is also effective for top-down UQ, for example, in paired (*O*,*I*) data. The advantages of a modern Bayesian approach applied to paired (*O*,*I*) data include the facts that one can: (1) accommodate any prior and any likelihood; (2) enforce any type of constraint, such as ratios of variances, with appropriate choice of prior, and (3) assess whether an implementation is well calibrated; for example, simulation can assess what fraction of 95% posterior probability intervals actually contain the true parameter such as σ^2_{Rd}. Disadvantages of a Bayesian approach include: (1) bias has to be assessed by sensitivity studies that vary the true and assumed likelihood and/or prior, and (2) numerical approaches such as Markov Chain Monte Carlo are easy to implement, but the user must perform convergence diagnostics to check whether one is really sampling from the correct posterior. ABC does not avoid such convergence issues, but the illustrated simulation strategy allows one to assess whether the chosen summary statistics, the distance measure, and the acceptance threshold lead to a well-calibrated approach.

7. References

[1] Burr, T., Hamada, M.S., Revisiting statistical aspects of nuclear material accounting Science and Technology of Nuclear Installations, Vol. 2013, Article ID 961360, 15 pages, doi:10.1155/2013/961360, 2013.

[2] Bonner, E., Burr, T., Guzzardo, T., Norman, C., Zhao, K., Beddingfield, D., Geist, W., Laughter, M., Lee, T., Improving the effectiveness of safeguards through comprehensive uncertainty quantification, Journal of Nuclear Materials Management, 44(2), 53-61, 2016.

[3] Burr, T., Krieger, T., Norman, C., Zhao, K., The impact of metrology study sample size on verification samples calculations in IAEA safeguards, European Journal Nuclear Science and Technology, 2, 36, 2016.

[4] Speed, T., Culpin, D., The role of statistics in nuclear materials accounting: issues and problems, Journal of the Royal Statistical Society B, 149(4), 281–313, 1986.

[5] Goldman, A., Picard, R., Shipley, J., Statistical methods for nuclear materials safeguards: an overview, Technometrics, 24(4), 267–275, 1982.

[6] Picard, R., Sequential analysis of materials balances, Nuclear Materials Management, 15(2), 38–42, 1987.

[7] Burr, T., Hamada, M.S., Bayesian updating of material balances covariance matrices using training data, International Journal of Prognostics and Health Monitoring, 5 (1) 006, pages: 13, 2014 .

[8] Walsh, S., Burr, T., Martin, K., The IAEA error approach to variance estimation for use in material balance evaluation and the international target values, and comparison to metrological definitions of precision, Journal of Nuclear Materials Management, 45(2):4-14, 2017.

[9] Burr, T., Croft, S., Jarman. K., Nicholson, A., Norman, C., Walsh, S., Improved uncertainty quantification in nondestructive assay for nonproliferation, Chemometrics, 159, 164-173, 2016.

[10] JCGM 104:2009, Evaluation of measurement data-an introduction to the “Guide to the Expression of Uncertainty in Measurement,” 2009.

[11] Fearnhead P., Prangle, D., Constructing summary statistics for approximate Bayesian computation: semi-automatic approximate Bayesian computation, Journal of the Royal Statistical Society B, 74(3), 419–474, 2012.

[12] Joyce, P., Marjoram, P., Approximately sufficient statistics and Bayesian computation, Statistical Applications in Genetics and Molecular Biology, 7(1), article 26, 2008.

[13] Burr, T., Skurikhin, A., Selecting summary statistics in approximate Bayesian computation for calibrating stochastic models, BioMedical Research International, Vol. 2013, Article ID 210646, 10 pages, 2013. doi:10.1155/2013/210646.

[14] Burr, T.,Croft, S., Krieger, T., Martin, K., Norman, C. Walsh, S., Uncertainty quantification for radiation measurements: bottom-up error variance estimation using calibration information, Applied Radiation and Isotopes, 108,49-57, 2015.

[15] ASTM C1514, Standard test method for measurement of ^{235}U fraction using the enrichment meter principle, 2008.

[16] Walton, R., Reilly, T., Parker, J., Menzel, J., Marshall, E., Fields, W., Measurements of UF6 cylinders with portable instruments, Journal of Nuclear Technology 21 (2), 133-148, 174.

[17] Miller, R., Beyond ANOVA: basics of applied statistics, Chapman & Hall, 1998.

[18] ISO 21748:2010 Guidance for the use of repeatability, reproduciblity, and trueness estimates in measurement uncertianty estimation.

[19] Zhao, K., Penkin, P., Norman, C., Balsely, S., Mayer, K., Peerani, P., Pietri, P., Tapodi, S., Tsutaki, Y., Boella, M., Renha G., Kuhn, E., STR-368 International target values 2010 for measurement uncertainties in safeguarding nuclear materials, IAEA, Vienna, 2010, www.inmm.org.

[20] Thompson, M., Ellison, S., Dark uncertainty, Accreditation and Quality Assurance 16, 483–487, 2011.

[21] Bich, W., Revision of the guide to the expression of uncertainty in measurement. Why and how, Metrologia 51, S155-S158, 2014.

[22] Elster, C., Bayesian uncertainty analysis compared to the application of the GUM and its supplements, Metrologia 51, S159-S166, 2014.

[23] Carlin, B., John, B., Stern, H., Rubin, D., Bayesian Data Analysis (1st ed), Chapman and Hall, 1995.

[24] Blum, M., Nunes, M., Prangle, D. Sisson, S., A comparative review of dimension reduction methods in approximate Bayesian computation. Statistical Science 28(2), 189–208, 2013.

[25] Nunes, M. Prangle, D., abctools: an R package for tuning approximate Bayesian computation analyses. The R Journal 7(2), 189–205, 2016.

[26] R: A language and environment for statistical computing, R Development Core Team, R Foundation for Statistical Computing, Vienna, Austria, 2008.

[27] Burr, T., Croft, S., Dale, D., Favalli, A., Weaver, B., Williams, B., Emerging applications of bottom-up uncertainty quantification in nondestructive assay, ESARDA Bulletin 53, 54-61, 2015.

[28] Burr, T., Croft, S., Hoover, A., Rabin, M., Exploring the impact of nuclear data uncertainties in microcalorimetry using ABC, Nuclear Data Sheets 123, 140-145, 2015.

[29] Hoover, A., Winkler, R., Rabin, M., Vo, D., Ullom, J., Bennett, D.,Doriese, W., Fowler, J., Horansky, R., Schmidt, D., Vale, L., Schaer, K., Determination of plutonium isotopic content by microcalorimeter gamma-ray spectroscopy, IEEE Transactions on Nuclear Science 60(2), 681-688, 2013.

Peelle's Pertinent Puzzle in Nuclear Safeguards Measurements

Elisa Bonner[1], Tom Burr[2], Claude Norman[2]

[1] Colorado State University
[2] Safeguards Information Management, IAEA, Vienna Austria
E-mail: ebonner@gmail.com,t.burr@iaea.org, c.norman@iaea.org

Abstract:

In nuclear safeguards, two measurement methods are sometimes used to infer nuclear material mass. Suppose that the method 1 and 2 estimates are 1.0 kg and 1.5 kg, respectively. Using generalized least squares (GLS) to combine two estimates has a long history dating to its development by Gauss in 1795. In some settings, GLS exhibits curious behaviour, as described in Peelle's Pertinent Puzzle (PPP) where the GLS estimate to combine the 1.0 and 1.5 estimates is 0.88. PPP was introduced in 1987 in the context of combining two or more estimates of fundamental parameters that arise in nuclear interaction experiments. When PPP occurs, the GLS estimate is outside the range of the data, which has led to concerns that GLS estimation is flawed. This paper describes GLS estimation and PPP and points out that PPP can only occur if the two estimates are highly correlated and have different variances. Next, this paper shows that PPP can arise in an example from safeguards, in which the goal is to estimate the average nuclear material mass in N items. A sample of n_1 items from the population of N items is measured by a lower-quality assay method; a subsample n_2 of the n_1 sampled items is also measured by a higher-quality assay method. This paper shows that PPP can arise in applying GLS to combine the estimates from the lower-quality and higher-quality assay methods, for any of three different measurement error models. Model A is the same as that used by a conventional safeguards model. Model B is a modification of model A. Model C arises when both assay methods are calibrated using reverse regression, which in recent uncertainty quantification studies has been shown to outperform classical regression followed by inversion.

Keywords: combining two measurements, generalized least squares, Peelle's pertinent puzzle

1. Introduction

Nuclear safeguards aim to verify that nuclear materials are used exclusively for peaceful purposes. To ensure that States are honouring their safeguards obligations, measurements of nuclear material inventories and flows are needed. Statistical analyses used to support conclusions require uncertainty quantification (UQ), usually by estimating the relative standard deviation (RSD) in random and systematic errors associated with each measurement method [1-10].

This paper uses a safeguards quantitative verification measurement example to show the importance of accurate UQ of measurement errors and to show that although PPP can arise, GLS is still an effective option to combine two or more measurements of the same unknown true quantify.

The safeguards example modified slightly from [1] is as follows. The average nuclear material mass in *N* items is to be estimated by selecting n_1 items at random and measuring these items with measurement method #1, a non-destructive assay (NDA) device, such as a neutron multiplicity counter. The NDA device is then re-calibrated by randomly selecting a subset n_2 of the n_1 items and measuring them by measurement method #2, a destructive assay (DA) method, a balance and mass spectrometer. The problem is to estimate the population mean using the $(n_1 + n_2)$ measurement results and to determine the variance of the estimate. As a specific example, the population may be *N* containers of U. The quality characteristic is the average mass of U-235.

Suppose in this example that the method 1 estimate is 1.0 kg and the method 2 estimate is 1.5 kg. For a particular covariance matrix [2] that contains the variances of the two estimates on the diagonal (0.1134 and 0.0505) and the covariance between the two estimates on the off-diagonal (0.06), the GLS estimate that combines the 1.0 and 1.5 estimates is 0.88. Under what conditions is it reasonable for the GLS estimate to be less than the smaller of the two measurements of nuclear material (NM), or greater than the larger of the two measurements? As [3] explains, the 0.88 estimate is reasonable if the two methods have large positive correlation and method 1 has smaller variance. Note that if the GLS estimate fell between 1.0 and 1.5, it would appear that the two methods have negative correlation. Because 0.88 is smaller than both 1.0 and 1.5, it appears that the two methods have a strong positive correlation, which is indeed the case. The unequal variances of method 1 and 2 provide information regarding whether the population NM mass is more likely to be less than the minimum or greater than the maximum of the two estimates.

In this case, the 0.88 estimate is closer to the method 1 estimate, which has smaller variance than the method 2 estimate.

This paper is organized as follows. Section two reviews GLS and PPP. Section three describes the safeguards example from [1], and modifies the measurement error assumptions from the example. Section four presents simulation results and shows that PPP can arise in the safeguards example. Section five summarizes and emphasizes the importance of accurate UQ.

2. Generalized Least Squares (GLS) and Peelle's Pertinent Puzzle (PPP)

2.1 GLS

GLS for parameter estimation has a long history dating to its development by Gauss and Legendre in the early 1800s [11]. PPP was introduced in the context of estimating fundamental parameters that arise in nuclear interaction experiments [2]. In PPP, the GLS estimate is outside the range of the data, eliciting concerns that GLS is flawed [4,5]. Reference [3] defended GLS in the PPP context and provided an example when PPP can occur. Although PPP examples remain relatively rare, the present paper illustrates that PPP can occur in the example from [1], and also defends GLS as an effective option to combine two (or more) estimates of the same quantity, regardless of whether PPP occurs.

To illustrate GLS, denote the results of two assay methods on the same item as X_1 and X_2. GLS applied to X_1 and X_2 provides the best linear unbiased estimate (BLUE) $\hat{\mu}$ of μ, regardless of whether PPP occurs [6]. Here, "best" means minimum variance and unbiased means that the average of $\hat{\mu}$ across many realizations of the same procedure is the true value μ. Note that one can write $\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} \mu \\ \mu \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \end{pmatrix}$ where $\begin{pmatrix} e_1 \\ e_2 \end{pmatrix} = \begin{pmatrix} R_1 \\ R_2 \end{pmatrix}$, or if there are also systematic errors, $\begin{pmatrix} e_1 \\ e_2 \end{pmatrix} = \begin{pmatrix} S_1 \\ S_2 \end{pmatrix} + \begin{pmatrix} R_1 \\ R_2 \end{pmatrix}$, where S_1 is the systematic error of method 1, R_1 is the random error of method 1, and similarly for method 2 [7-10]. This paper uses either μ or T, depending on the context, to denote the true NM mass in an item. Denote the 2-by-2 covariance of $\begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ as Σ with the method variances $\sigma_{1,1}^2$ and $\sigma_{2,2}^2$ as the diagonal entries and the method covariance $\sigma_{1,2}^2 = \sigma_{2,1}^2$ as the off-diagonal entries. The well-known GLS estimate is $\hat{\mu} = cG^T\Sigma^{-1}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, where $G^T = (1,1)$ $c = (G^T\Sigma^{-1}G)^{-1}$, and the variance of $\hat{\mu}$ is $\sigma^2 = (G^T\Sigma^{-1}G)^{-1}$.

In the example, μ is the unknown and writing $\hat{\mu} = a_1X_1 + (1-a_1)X_2$, note that $\sigma_{\hat{\mu}}^2 = a_1^2\sigma_{X_1}^2 + (1-a_1)^2\sigma_{X_2}^2 + 2a_1(1-a_1)\sigma_{X_1,X_2}^2$. Then the GLS solution $\hat{\mu} = cG^T\Sigma^{-1}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ arises from standard calculus (by setting the derivative of $\sigma_{\hat{\mu}}^2$ with respect to a_1 to zero and solving for a_1) or from projection matrix results in linear algebra. The result is $a_1 = c_1$, where $c = (c_1, c_2) = (c_1, 1-c_1) = (G^T\Sigma^{-1}G)^{-1}G^T\Sigma^{-1}$.The estimate $\hat{\mu}$ is a weighted average of the two estimates, with weights summing to 1. In the case of uncorrelated measurements, with zeros on the off-diagonals of Σ, the weights are proportional to the inverse of the respective variances, so $a_1 = c_1 = \sigma_2^2 / (\sigma_1^2 + \sigma_2^2)$. If the measurements are uncorrelated, then the GLS estimate is guaranteed to be between the two estimates.

2.2 PPP

PPP is defined as either $\hat{\mu} > max(x_1, x_2)$ or $\hat{\mu} < \min(x_1, x_2)$. Sivia [12] gave a condition on Σ for which PPP cannot occur, expressed as: if $\rho \le min(\frac{\sigma_1}{\sigma_2}, \frac{\sigma_2}{\sigma_1})$ then PPP cannot occur. In practice, entries in Σ are estimated, and so [3] shows that there are situations where it appears that PPP occurs when it does not, and vice versa. A theorem in [3] shows that if a_1 and $a_2 = 1 - a_1$ have opposite signs, then PPP occurs:

Theorem 1. Suppose a_1 and $a_2 = 1 - a_1$have opposite signs. Then either $\hat{\mu} > max(x_1, x_2)$ or $\hat{\mu} < \min(x_1, x_2)$. That is, $\hat{\mu}$will always fall outside the range of (x_1, x_2). The simple proof from [3] of Theorem 1 is given here.

Proof. First assume $a_1 > 1$ and $a_2 < 0$. If $x_1 < x_2$ then $\hat{\mu} = a_1x_1 + a_2x_2 < a_1x_1 + a_2x_1 = x_1$ because $a_2 < 0$. Similarly, if $x_1 > x_2$ then $\hat{\mu} = a_1x_1 + a_2x_2 > a_1x_1 + a_2x_1 = x_1$ because $a_2 < 0$. The proof is completed by next assuming $a_1 < 0$ and $a_2 > 1$, and following similar steps.

3. The Safeguards Example

Jaech [1] used the following model, Eq. (1) for the better (DA) measurement and Eq. (2) for the worse (NDA) measurement:

$$X_{1i} = T_i + S_1 + R_{1i} \quad (1),$$

$$X_{2i} = \beta T_i + R_{2i} \quad (2),$$

where T_i is the true value (in kg) of item i, S_1 is the systematic error of method 1, R_{1i} is the random error of method 1, β is a constant that is estimated from calibration data. Estimation error in $\hat{\beta}$ leads to systematic error in method 2. In this context, $\hat{\beta}$ is estimated using $\hat{\beta} = \sum_{i=1}^{n_2} X_{2,i} \Big/ \sum_{i=1}^{n_2} X_{1,i}$, which is a ratio of random variables. In many applications, including this one, the variance of a ratio of random variables must be estimated by simulation because the estimate of variance based on the linear first-term Taylor series approximation is not accurate [8-10]. Then, the method 1 estimate of the population mean is $\hat{\mu}_1 = \sum_{i=1}^{n_2} X_{1,i} / n_2$ and the method 2 estimate is $\hat{\mu}_2 = \frac{1}{\hat{\beta}} \sum_{i=1}^{n_1} X_{2,i} / n_1$. This example involves measurement error in both X_1 and X_2, so the literature on "errors-in-predictors" is relevant [13,14] , and the variance in $\frac{1}{\hat{\beta}}$ is estimated by simulation in Section 4.

Rather than the way that GLS was presented in Section 2.1, GLS is often presented in the context of estimating β in a linear regression relating response Y to, for example, predictors X_1 and X_2, denoted $Y = \beta X + e$ [6], where X is a matrix with n rows containing X_1 values in column 1 and X_2 values in column 2. Perhaps this is why [1] did not recognize this safeguards example as one for which known GLS results apply (as shown in Section 2.1). So, instead of applying known GLS results, [1] re-derived the GLS solution, by setting the derivative of an approximate expression for $\sigma^2_{\hat{\mu}}$ with respect to a_1 equal to zero to solve for the value a_1 that minimizes the approximate expression for $\sigma^2_{\hat{\mu}}$. The approximation result from [1], which is evaluated in Section 4, is $\sigma^2_{\hat{\mu}} \approx \{\frac{na^2 + n_1 - 2an_1}{n_1(n-n_1)} - 1/N\}\sigma^2_{\mu} + \sigma^2_{S1} + \frac{\sigma^2_{R1}}{n_1} + \frac{n(1-a)^2\sigma^2_{R2}}{n_1(n-n_1)\beta^2}$ (3), where $\sigma^2_{\mu} = \sum_{i=1}^{N} (T_i - \bar{T})^2 / (N-1)$ (4).

To arrive at Eq. (3), reference [1] ignored estimate error in $\hat{\beta}$, assumed $\hat{\beta} = \beta$, and applied standard error variance propagation to a linear Taylor-series approximation of $\hat{\mu}$. Simulations in Section 4 show that estimation errors in the covariance matrix Σ can lead to the belief that PPP occurs when it does not, and vice versa.

This paper uses three distinct error models for example 1. Jaech's [1] Equations (1) and (2) will be referred to as model A. As model B, instead of Equations (1) and (2), one could use the more common error models [7]:

$$X_{1i} = T_i + S_1 + R_{1i} \qquad (5),$$
$$X_{2i} = T_i + S_2 + R_{2i} \qquad (6),$$

where Eq. (5) is the same as Eq. (2), and Eq. (6) explicitly provides the systematic error for method 2. Again, T_i is the true value of item i, $S_{X_1} \sim N(0,\sigma_{S_1})$,is the short-term systematic error of method 1, $R_{X_{1i}} \sim N(0,\sigma_{R_1})$ is the random error of method 1, and similarly for method 2 in Eq. (6). Note that model B is not the same as model A unless $S_2 = T(\beta - 1)$, which is a relative error model for S_2

As for model C, the data that were used to calibrate methods 1 and 2 prior to measuring the sampled items could be used. Recent numerical evaluations of four calibration options have led a recommendation to use reverse calibration [8-10], using n (X_{1i},T_i) pairs to fit $T_i = \beta_{1,0} + \beta_1 X_{1i} + R_{1_i}$ and n (X_{2i},T_i) pairs to fit $T_i = \beta_{2,0} + \beta_{2,1} X_{2i} + R_{2_i}$ for method two. The calibration options evaluated in [8] are to apply classical regression, fitting $X_{1i} = \alpha_0 + \alpha_1 T + R_{1_i}$, and then inverting to solve $\hat{T} = (X_{1i} - \hat{\alpha}_0)/\hat{\alpha}_1$ (and similarly for method two), or to apply reverse calibration, directly fitting $T_i = \beta_0 + \beta_1 X_{1i} + R_{1_i}$. Both options can adjust for errors in predictors or not, so there is a total of four calibration options. The reverse calibration option without adjusting for errors in predictors (but using simulation with errors in predictors to accurately evaluate the behaviour of the estimate) has been found to have the same or smaller estimation error, so it is the only option evaluated in Section 4. Figure 1 plots the observed bias in 1 (of 10^5) simulation with 3 standards, and as shown in [8], model C can be expressed as:

$$X_{1i} = T_i + S_{1,1} + S_{1,2}(T_i - \bar{T}) + R_{1i} \qquad (7),$$
$$X_{2i} = T_i + S_{2,1} + S_{2,2}(T_i - \bar{T}) + R_{2i} \qquad (8),$$

with both additive and multiplicative systematic errors. The additive systematic error arises from estimation error in the intercept. The multiplicative systematic error arises from estimation error in the slope, increasing from 0 at the middle of the calibration data to large positive or negative values near the ends of the calibration data. Some of the well-known results for least squares regression are relevant for evaluating calibration data; however, reverse calibration is not a straight-forward application of regression because of the errors in predictors, and [8-10] recommend simulation for accurate model fitting and uncertainty quantification arising from calibration data.

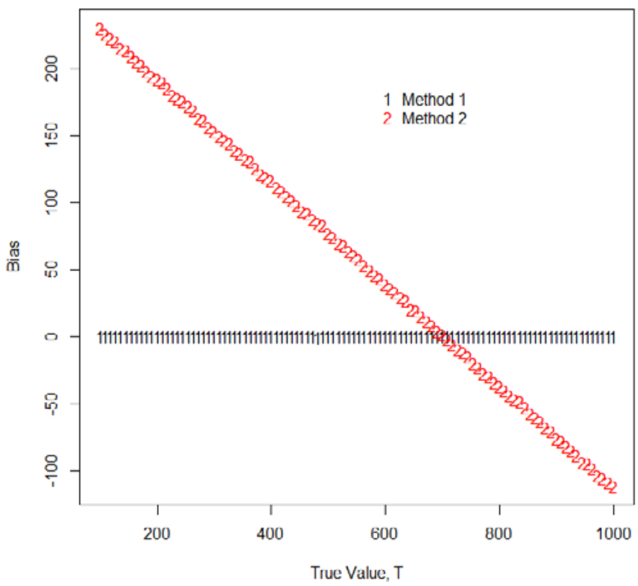


Figure 1: Bias versus true value for method 1 and method 2 in one calibration.

4. Simulation Results for the Safeguards Example

Recall that PPP was introduced in [2] for $\Sigma = \begin{pmatrix} 0.1134 & 0.06 \\ 0.06 & 0.0506 \end{pmatrix}$ for which the values $a_1 = 1.22, a_2 = -0.22$, minimize $\sigma^2_{\hat{\mu}}$ and $\hat{\mu} = 1.22 \times 1 - 0.22 \times 1.5 = 0.88$. Estimation errors in the sample covariance matrix $\hat{\Sigma}$ to estimate Σ can make it appear that PPP does not occur. For example, in 10^5 simulations in R with n = 10, 100, and 1000 (X_1, X_2) pairs, the relative frequency that $\hat{\Sigma}$ leads to the wrong conclusion that PPP does not occur is 72%, 40%, and 0%, respectively. Estimation errors in the sample covariance matrix $\hat{\Sigma}$ to estimate Σ can also make it appear that PPP does occur when it does not. For example, with $\Sigma = \begin{pmatrix} 0.1134 & 0.0506 \\ 0.0506 & 0.06 \end{pmatrix}$, $a_1 = 0.13, a_2 = 0.87$, so PPP does not occur, but in 10^5 simulations in R [15] with n = 10, 100, and 1000 (X_1, X_2) pairs, the relative frequency that $\hat{\Sigma}$ leads to the wrong conclusion that PPP does occur is 32%, 14%, and 0%, respectively.

It was found using 10^6 simulations in R [14] that PPP can occur for models A, B, and C. It was also found that Eq. (3) is not sufficiently accurate for $\sigma^2_{\hat{\mu}}$ (see results in Sections 4.1-4.3). The values in the covariance matrices given below are repeatable to the number of digits shown across sets of 10^6 simulations. In all the results below, $N = 100, \overline{T} = 100, \sigma_T = 50$.

4.1 Model A

The example from [1] was evaluated by simulation in R [15] using $n_1 = 30, n_2 = 5, \sigma_{S1} = 1, \sigma_{R1} = 0.1, \beta = 1.1, \sigma_{R2} = 0.1$. The estimated covariance matrix is $\hat{\Sigma} = \begin{pmatrix} 384 & 399 \\ 399 & 838 \end{pmatrix}$, which implies that the correlation between Method 1 and 2 is 0.70, and that $a_1 = 1.04, a_2 = -0.04$, so by Theorem 1, PPP occurs. Figure 2 plots the root mean squared estimation error in $\hat{\mu}$ versus a_1.

Figure 3 plots the observed and predicted $\sigma_{\hat{\mu}}$ using Eq. (3) from [1] versus a_1. Note that Eq. (3) from [1] is not an accurate approximation. It is noted here that reference [1] did not use $\hat{\mu}_2 = \frac{1}{\hat{\beta}} \sum_{i=1}^{n_1} X_{2,i} / n_1$, but instead used $\hat{\mu}_2 = \frac{1}{\hat{\beta}} \sum_{i=1}^{n_1 - n_2} X_{2,i} / (n_2 - n_1)$, which uses only those NDA measurements that were not used to estimate β.
All simulations evaluated both options. It was found that both options can lead to PPP, but the first option has smaller estimation error, so the reported results all used $\hat{\mu}_2 = \frac{1}{\hat{\beta}} \sum_{i=1}^{n_1} X_{2,i} / n_1$. However, Figure 3 used $\hat{\mu}_2 = \frac{1}{\hat{\beta}} \sum_{i=1}^{n_1 - n_2} X_{2,i} / (n_2 - n_1)$ because Eq. (3) from [1] was for $\hat{\mu}_2 = \frac{1}{\hat{\beta}} \sum_{i=1}^{n_1 - n_2} X_{2,i} / (n_2 - n_1)$.

4.2 Model B

The modified example from [1] was evaluated by simulation in R [15] using n_1=20, n_2=5, $\sigma_{S1} = 1, \sigma_{R1} = 1, \sigma_{S1} = 2, \sigma_{R2} = 2$. The estimated covariance matrix is $\hat{\Sigma} = \begin{pmatrix} 333 & 343 \\ 343 & 784 \end{pmatrix}$, which implies that the correlation between Method 1 and 2 is 0.67, and that $a_1 = 1.02, a_2 = -0.02$. so by Theorem 1, PPP occurs.

4.3 Model C

The modified example from [1] was evaluated by simulation in R [15] using N = 100, n_1=20, n_2=10, $\overline{T} = 100$, $\sigma_T = 50$, $\beta_0 = 1, \beta_1 = 100$, $\sigma_{T1} = 0.003, \sigma_{R1} = 0.003, \sigma_{T2} = 0.04$, $\sigma_{R2} = 0.04$. There were 3 calibration items, with true values of 100, 550, and 1000 grams. The estimated covariance matrix is $\hat{\Sigma} = \begin{pmatrix} 835 & 850 \\ 850 & 1357 \end{pmatrix}$ which implies that the correlation between Method 1 and 2 is 0.80, and that $a_1 = 1.03, a_2 = -0.03$, so by Theorem 1, PPP occurs.

Models A, B, and C can all exhibit PPP and for the numerical examples chosen, models A, B, and C have $a_1 = 1.04, a_2 = -0.04$, $a_1 = 1.02, a_2 = -0.02$, and $a_1 = 1.03, a_2 = -0.03$, respectively.

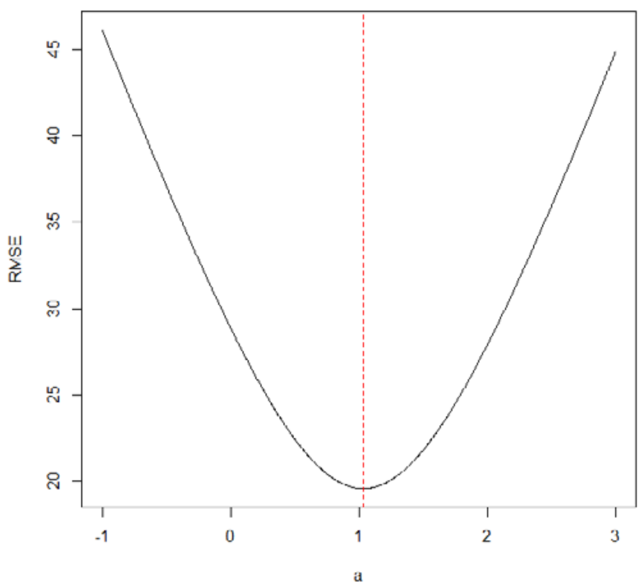


Figure 2: The RMSE versus a_1 for model A. The minimum RMSE occurs at $a_1 = 1.03, a_2 = -0.03$.

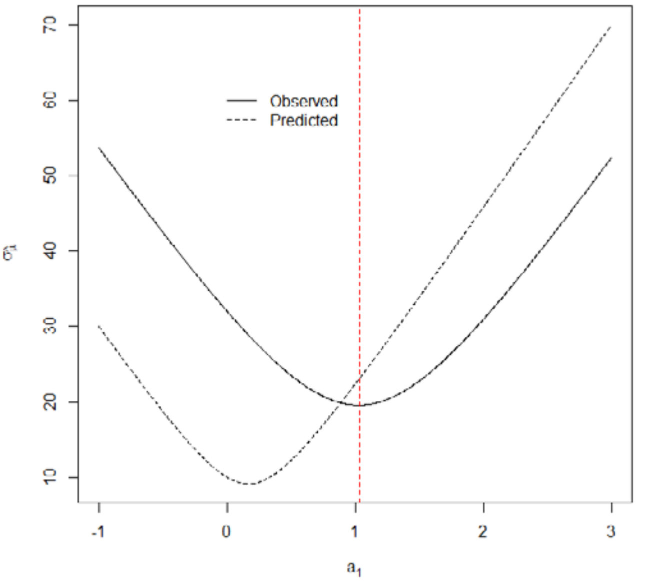


Figure 3: Observed and predicted value for Model A (from Eq. (3) from reference [1]) of $\sigma_{\hat{\mu}}$ versus a_1.

5. Summary

Recent work on uncertainty quantification [8-10] for NDA has found that simulation is needed for high-quality uncertainty quantification. This paper provides another example where simulation is needed for high-quality estimation of the 2-by-2 covariance matrix Σ of two assay methods. The example was a safeguards measurement example from [1] in which a sample of items was assayed using both a lower uncertainty (DA) method and a higher uncertainty (NDA) method was re-evaluated. First, it was shown that generalized least squares can be applied to optimally combine the resulting two estimates, $\hat{\mu}_1$ and $\hat{\mu}_2$ of the population mean. μ. Second, it was shown using simulation for any of three measurement error models, that there is large positive covariance between the two estimates, $\hat{\mu}_1$ and $\hat{\mu}_2$, and one estimate has much larger variance than the other. Third, it was shown that PPP can occur for all three models. Because PPP is a somewhat rare phenomenon, this finding is of interest. However, safeguards analysts need not be concerned if PPP occurs in such an example; because it is an understandable behaviour of GLS [2-4,16] in examples with large positive covariance matrices. Analysts are advised to use simulation to ensure high-quality estimates of Σ so that analysts know when PPP does occur. Reference [16] considers alternatives to PPP when there is non-negligible estimation error in $\hat{\Sigma}$.

6. References

[1] Jaech, J., Consistency check approach to quantitative verification, Journal of Nuclear Materials Management 13 (2), 30-33, 1984.

[2] Peelle, R., Peelle's Pertinent Puzzle; Oak Ridge National Laboratory Memorandum: Washington DC, USA, 1987.

[3] Burr, T., Kawano, T., Talou, P., Pen, F., Hengartner, N., Defense of the Generalized Least Squares Solution to Peele's Pertinent Puzzle, Algorithms 4(1), 28-39, 2011.

[4] Chiba, S., Smith, D., A Suggested Procedure for Resolving an Anomaly in Least-Squares Data Analysis Known as Peelle's Pertinent Puzzle and the General Implications for Nuclear Data Evaluation, Report ANL/NDM-121; Argonne National Laboratory: Argonne, IL, USA, 1991.

[5] Hanson, K.; Kawano, T., Talou, P., Probabilistic Interpretation of Peelle's Pertinent Puzzle. In Proceeding of International Conference of Nuclear Data for Science and Technology, 26 September–1 October 2004.

[6] Christensen, R., Plane Answers to Complex Questions, The Theory of Linear Models; Springer: New York, 1999; pp. 23–25.

[7] Walsh, T., Burr, T., Martin, K., Discussion of the IAEA Error Approach to Producing Variance Estimates for use in Material Balance Evaluation and the International Target Values, and Comparison to Metrological Definitions of Precision, Journal of Nuclear Materials Management, 45(2),4-14, 2017.

[8] Burr, T., Croft, S., Jarman. K., Nicholson, A., Norman, C., Walsh, S., Improved Uncertainty Quantification in NonDestructive Assay for Nonproliferation, Chemometrics, 159, 164-173, 2016.

[9] Burr, T., Croft, S., Krieger, T., Martin, K., Norman, C. Walsh, S., Uncertainty Quantification for Radiation Measurements: Bottom-Up Error Variance Estimation Using Calibration Information, Applied Radiation and Isotopes, 108,49-57, 2015.

[10] Bonner, E., Burr, T., Krieger, T., Martin, K., Norman, C., Comprehensive Uncertainty Quantification in Nuclear Safeguards, Science and Technology of Nuclear Installations, 1-16, 1155/2017/2679243, 2017.

[11] Stigler, S., Gauss and the Invention of Least Squares, Annals of Statistics 9(3), 465-474, doi:10.1214/aos/1176345451, 1981.

[12] Sivia, D., Data Analysis—A Dialogue With the Data. In Advanced Mathematical and Computational Tools in Metrology VII, Ciarlini, P., Filipe, E., Forbes, A.B., Pavese, F., Perruchet, C., Siebert, B.R.L., Eds.; World Scientific Publishing Company: Lisbon, Portugal, 108–118, 2006.

[13] Burr,T., Knepper, P., A Study of the Effect of Measurement Error in Predictor Variables in Nondestructive Assay, Applied Radiation and Isotopes 53(4-5): 547-555, 2000.

[14] Ripley, B., Thompson, M., Regression Techniques for the Detection of Analytical Bias, Analyst 112, 377-383, 1987.

[15] R: A language and environment for statistical Computing, R Development Core Team, R Foundation for Statistical Computing, Vienna, Austria, 2008.

[16] Burr, T., Kawano, T., Talou, P., Pen, F., Hengartner N., Graves T., Alternatives to the Generalized Least Squares Solution to Peele's Pertinent Puzzle, Algorithms *4*(2), 115-130, 2011.

Investigating the Cherenkov light production due to cross-talk in closely stored nuclear fuel assemblies in wet storage

Erik Branger, Sophie Grape, Peter Jansson, Erik Andersson Sundén, Staffan Jacobsson Svärd

Uppsala University, Uppsala, Sweden
E-mail: erik.branger@physics.uu.se

Abstract:

The Digital Cherenkov Viewing Device (DCVD) is one of the tools available to a safeguards inspector performing verifications of irradiated nuclear fuel assemblies in wet storage. One of the main advantages of safeguards verification using Cherenkov light is that it can be performed without moving the fuel assemblies to an isolated measurement position, allowing for quick measurements. One disadvantage of this procedure is that irradiated nuclear fuel assemblies are often stored close to each other, and consequently gamma radiation from one assembly can enter a neighbouring assembly, and produce Cherenkov light in the neighbour. As a result, the measured Cherenkov light intensity of one assembly will include contributions from its neighbours, which may affect the safeguards conclusions drawn.

In this paper, this so-called near-neighbour effect, is investigated and quantified through simulation. The simulations show that for two fuel assemblies with similar properties stored closely, the near-neighbour effect can cause a Cherenkov light intensity increase of up to 3% in a measurement. For one fuel assembly surrounded by identical neighbour assemblies, a total of up to 14% of the measured intensity may emanate from the neighbours. The relative contribution from the near-neighbour effect also depends on the fuel properties; for a long-cooled, low-burnup assembly, with low gamma and Cherenkov light emission, surrounded by short-cooled, high-burnup assemblies with high emission, the measured Cherenkov light intensity may be dominated by the contributions from its neighbours.

When the DCVD is used for partial-defect verification, a 50% defect must be confidently detected. Previous studies have shown that a 50% defect will reduce the measured Cherenkov light intensity by 30% or more, and thus a threshold has been defined, where a ≥30% decrease in Cherenkov light indicates a partial defect. However, this work shows that the near-neighbour effect may also influence the measured intensity, calling either for a lowering of this threshold or for the intensity contributions from neighbouring assemblies to be corrected for. In this work, a method is proposed for assessing the near-neighbour effect based on declared fuel parameters, enabling the latter type of corrections.

Keywords: DCVD; partial defect verification; Cherenkov light; Geant4; Cross-talk

1. Introduction

Irradiated nuclear fuel assemblies are commonly stored in water for radiation protection, as well as for decay heat removal. As a result of the interactions of the radiation emanating from the fuel assemblies with the surrounding water, Cherenkov light is produced This Cherenkov light has frequently been assessed by safeguards inspectors, using the presence, characteristics and intensity of the Cherenkov light to verify that the object under study is an irradiated nuclear fuel assembly, and not some other non-radioactive item.

The dominant path of Cherenkov light production is that gamma rays emitted in the decay of fission products enter the water, and are either photo-electrically absorbed or Compton-scatter off an electron. If the electron receives sufficient energy from the gamma ray, it will radiate Cherenkov light. In addition, high-energy beta-decay electrons can pass through the cladding and enter the water to produce Cherenkov light directly, though this contribution will be minor compared to the Cherenkov light produced by gamma decays. Neutrons cannot directly produce Cherenkov light since they have no electric charge, but radiation following a neutron interaction, such as e.g. inelastic scattering or fission of a uranium nuclei, can contribute. However, due to the low intensity of neutron emissions compared to gamma emissions, this contribution is expected to be negligible.

The Digital Cherenkov Viewing Device (DCVD) is one of the tools available to safeguards inspectors to measure the Cherenkov light emissions from irradiated nuclear fuel assemblies in wet storage. The DCVD can be used for gross- as well as partial-defect verification [1]. The type of partial defect analysis under study in this paper relies on comparisons of the measured intensities to predicted intensities, where removal or replacement of a fraction of the fuel rods will result in a lowered Cherenkov light intensity.

One of the main advantages of the DCVD is that the fuel assemblies do not have to be moved to an isolated area for measurement. A downside of measuring the assemblies where they are stored is that gamma radiation from closely stored assemblies can enter neighbouring

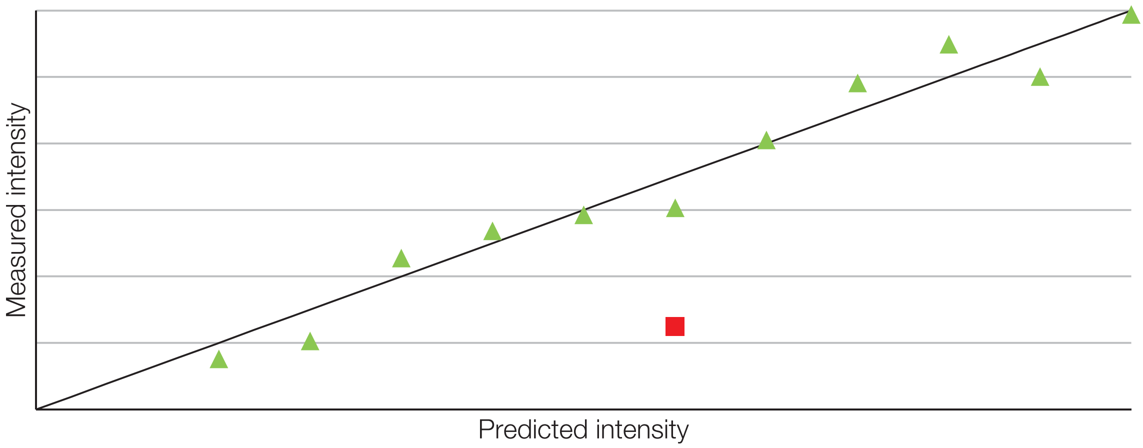


Figure 1: Illustration of the calibration procedure and partial defect verification method using the DCVD. For each fuel type, a linear fit is made between the predicted and measured intensity, where the fitted slope relates the predicted and measured intensity values. If any measured value deviates by more than 30% from the predicted (red square), a partial defect may be suspected.

assemblies and cause Cherenkov light emission there. This cross-talk, referred to as the near-neighbour effect, introduces a measurement error that is not compensated for in the currently deployed inspection procedure. The aims of this paper are: (i) to characterize and quantify the near-neighbour effect under selected fuel storage conditions, (ii) to identify how the near-neighbour effect affects the partial-defect verification procedure currently used, and (iii) suggest a method for its compensation.

1.1 Partial defect verification of used nuclear fuel using the DCVD

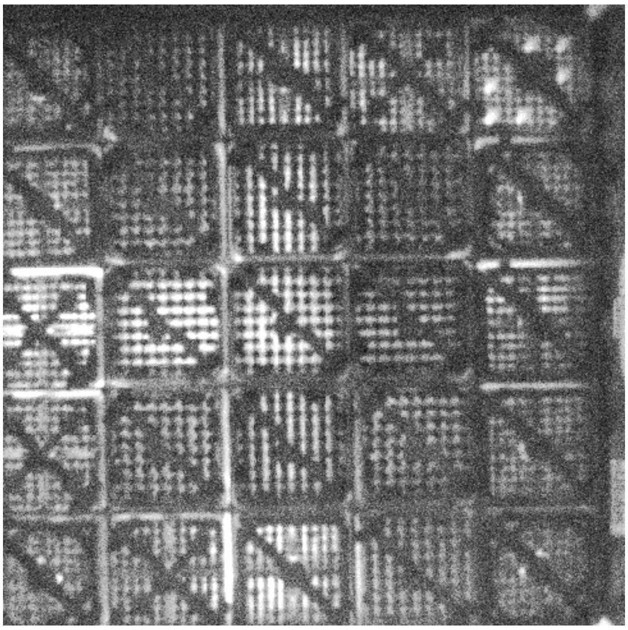
There are two methods used to detect partial defects in nuclear fuel assemblies with the DCVD. The first method uses image analysis to detect empty rod positions, and can be used to detect any removed rods in visible positions, as seen from the measurement position above the fuel. The second method is used to detect possible substitution of 50% of the fuel rods in an assembly. This method relies on the comparison of the measured intensity to a predicted intensity, based on operator-provided fuel declarations. In this analysis, the measured fuel assemblies are grouped by fuel type, so that each group contains fuels with the same physical design. For calibration within each group, the measured and predicted intensities are related by a linear fitting, as illustrated in Figure 1. As a result of this calibration, the predicted intensity values do not correspond to absolute measured intensity, but to a relative intensity of all fuel assemblies of the same type, and deviations from the group's linear fit call for further investigations of possible reasons. It is known from simulations that if 50% of the rods in an assembly are substituted with non-radioactive rods, the Cherenkov light intensity will be reduced by at least 30% [2]. Thus, if any measured intensity of an assembly is more than 30% lower than expected, a partial defect may be suspected.

Up until recently, the prediction method used was based on a parameterization of the Cherenkov light intensity as a function of burnup and cooling time in a BWR 8x8 configuration [3]. This method is currently being replaced by a new method [4], which more accurately considers the fuel irradiation history by calculating the inventory of fission products using ORIGEN [5], by considering the geometry of the fuel assemblies, and by including Cherenkov light intensity contributions from both gamma and beta decays [6].

1.2 DCVD measurements and the near-neighbour effect

During a measurement, the DCVD is typically mounted on the railing of a moveable bridge, looking down on the fuel storage pond. The fuel assemblies are typically stored densely enough that radiation from one fuel assembly may enter neighbouring assemblies and create Cherenkov light there. Due to the relatively long distance that the radiation must travel to reach a neighbour, only gamma-ray emissions are expected to contribute to the near-neighbour effect. The intensity of neutron emissions is too low in comparison to gamma emissions to contribute significantly, and the ranges of alpha and beta particles are too short to contribute. This work hence considers only Cherenkov light produced due to gamma-decays of fission products. The magnitude of the near-neighbour effect is a function of the distance between the fuels, the amount of storage rack material present in between the assemblies and the energy spectrum of the gamma-ray emissions, which depend on the fuel cooling time.

In Figure 2, an example is shown of the storage situation at the Swedish Central Interim Storage Facility for Spent Nuclear fuel (Clab), where 25 BWR fuels are stored in one fuel basket. The fuels are stored very close to each other, being separated by 4 mm of borated steel. At a reactor fuel pond, there is typically more distance in between the fuels for criticality safety reasons, and it is also more likely that fresh or low-burnup fuel is stored close to high-burnup fuel, which in turn may cause a significant near-neighbour



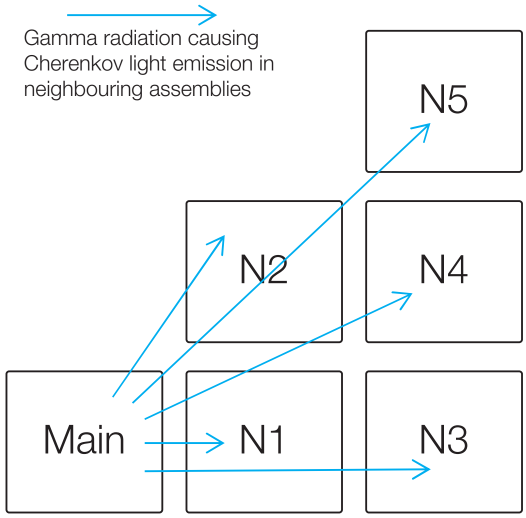


Figure 2 : Left: DCVD image of 25 BWR fuels stored at the Swedish Central Interim Storage Facility for Spent Nuclear Fuel (Clab). Image courtesy of Dennis Parcey, Clab, and the Canadian Nuclear Safety Commission (CNSC). **Right**: For an active assembly emitting gamma radiation, called main, this paper analyses the Cherenkov light produced in the neighbouring assemblies, labelled N1 to N5, by gamma radiation originating from the main assembly. For symmetry reasons, all surrounding assemblies in a 5x5 grid may be defined using labels N1 to N5.

intensity in the low-burnup neighbours. Low-burnup fuel will give rise to relatively low levels of gamma emission and consequently low levels of Cherenkov light, in comparison to high-burnup, short-cooled fuel. Accordingly, a large fraction of the gamma radiation in a low burnup fuel may have its origin in neighbouring high-burnup fuel, thus a significant fraction of the Cherenkov-light emission in the low-burnup fuel may be attributed to the near-neighbour effect. To be able to refer to the different neighbouring position in a storage rack, Figure 2 also labels the five neighbour positions considered in this work, where position N1 shares one side with the main assembly causing the near-neighbour effect in the studies, N2 shares a corner with the main assembly, and N3-N5 are one row/column further away. Other positions in a 5x5 grid may be referred to using these labels due to the symmetry of the storage situation.

2. Definition and characterization of the near-neighbour effect

In this work, the near-neighbour effect is studied in terms of the effect of one assembly emitting gamma radiation ("Main" in Figure 2) to its neighbours (N1-N5). The results will be presented as the ratio, NNR, of the Cherenkov light intensity in a neighbour ($I_{neighbour}$) produced by gamma radiation from the main assembly, as compared to the intensity in the main assembly itself (I_{main}), or

$$NNR = \frac{I_{neighbour}}{I_{main}} \qquad (1)$$

Note that by this definition, the intensity I_{main} is caused only by fission product decays in the main assembly. For real measurements, this value is not accessible due to the near-neighbour effect, though I_{main} can be predicted using one of the available prediction models [3] [6]. Furthermore, this study is limited to gamma-ray and bremsstrahlung emission, whereas it has been shown that beta particles may increase I_{main} by 1-10%, depending on fuel assembly type, irradiation history and cooling time [4]. There are negligible beta particle contributions to $I_{neighbour}$ because of their short travel range in water.

2.1 Simulations

To characterize the near-neighbour effect, simulations were run for two different fuel assembly configurations, BWR 8x8 and PWR 17x17, and for two different fuel storage situations. The simulations were performed using a toolkit based on Geant4 [7], which is a further development of a previously used toolkit for simulating the Cherenkov light production in irradiated nuclear fuel [8].

The fuel assemblies were modelled including fuel rods and control-rod guide tubes for PWR, respectively a water channel and a fuel channel surrounding the rod configuration for BWR. The dimensions of the simulated fuel assemblies are given in Table 1. In addition, walls of a square steel storage rack were also included in the simulations. Vertically directed Cherenkov light was analysed in the simulations, since the DCVD will measure the vertical light component given the measurement situation with the DCVD situated above the fuel. Cherenkov light at an angle smaller than 3 degrees to the vertical axis was considered representative of the vertical light component in the simulations, and this value also allows for comparisons with earlier simulation results [8]. This angle is wide enough that sufficient statistics can be obtained in the simulations in reasonable time, while being narrow enough to represent the vertical component.

	BWR 8x8	PWR 17x17
Number of fuel rods:	63	264
Fuel pellet diameter [mm]:	10.44	8.18
Cladding thickness [mm]:	0.91	0.57
Rod centre to centre distance [mm]:	16.3	12.6

Table 1: Dimensions of the simulated fuel assemblies.

The fuel depletion code ORIGEN [5] was used to assess the gamma spectrum for fuel assemblies with burnups of 10, 20, 30 and 40 MWd/kgU, and cooling times ranging from 0.25 to 60 years. The initial enrichment was set to 2% in all cases. These fuel parameter sets were chosen to be comparable to earlier studies [3] [8] . Fuels with 10, 20 and 30 MWd/kgU burnup were simulated as irradiated for four cycles, where each cycle consisted of 312.5 days of irradiation and 46 days of cooling, for a total of 1250 irradiation days. The power levels for the three lower burnups were 8, 16 and 24 kW/kgU, respectively. For the 40 MWd/kgU case, the power level remained at 24 kW/kgU, and the fuel was irradiated for 5 cycles. Note also that the gamma spectrum provided by ORIGEN includes both gamma-rays from fission product decays as well as bremsstrahlung produced when beta-particles are stopped in the fuel material.

2.2 Effects of burnup and cooling time for BWR assemblies

Figure 3 shows results of the simulations of the near-neighbour effect for BWR 8x8 fuels with a burnup of 40 MWd/kgU, for the fuel storage situation shown in Figure 2, with a 4 mm steel wall separating the assemblies. As can be seen, the N1 position is most strongly affected by the near-neighbour effect, with an NNR up to 2.9% of the main assembly intensity. For the N2 position, the near-neighbour effect is weaker, however; the NNR value is affected by attenuation in the assembly at N1, and will differ if the N1 position is occupied or vacant (called "N2" respectively "N2 only" in Figure 3). Accordingly, it is not only important to consider the properties of the emitting fuel assembly when estimating the near-neighbour effect; it is also important to consider which nearby positions that do not contain fuel to estimate the effect correctly. With N1 occupied, the intensity in N2 is up to 0.4% of the main assembly intensity, and with the N1 absent, it is up to 0.9% For the N3 position, if N1 and N2 are occupied the near-neighbour effect is at most 0.05%, and could be neglected. However, if N1 and N2 are absent, the near-neighbour intensity in N3 can be up to 0.5% (called "N3 only" in Figure 3), comparable to the intensity found at N2. The intensities in the N4 and N5 positions were found to be negligible in all cases simulated.

As can also be seen in Figure 2, the near-neighbour intensity ratios NNR, (Eq. (1)), reach maxima at a cooling time of around 1 year. As an example, the N1 position has a maximum NNR value at 1 year of 2.9%, which decreases to 1.9% after 40 years cooling. This is due to the changing gamma spectrum of the fuel assembly with cooling time [9]. For short-cooled fuel, several high-energy gamma-emitting isotopes are still present, which have relatively long range and thus contribute more to the near-neighbour intensity. As the fuel cools, the gamma emissions become dominated by the 662 keV emissions of Cs-137, which are of lower energy and has a relatively shorter range. As a consequence of the changing gamma spectrum with time, compensating for the near-neighbour effect will require assessing the gamma spectrum of all assemblies

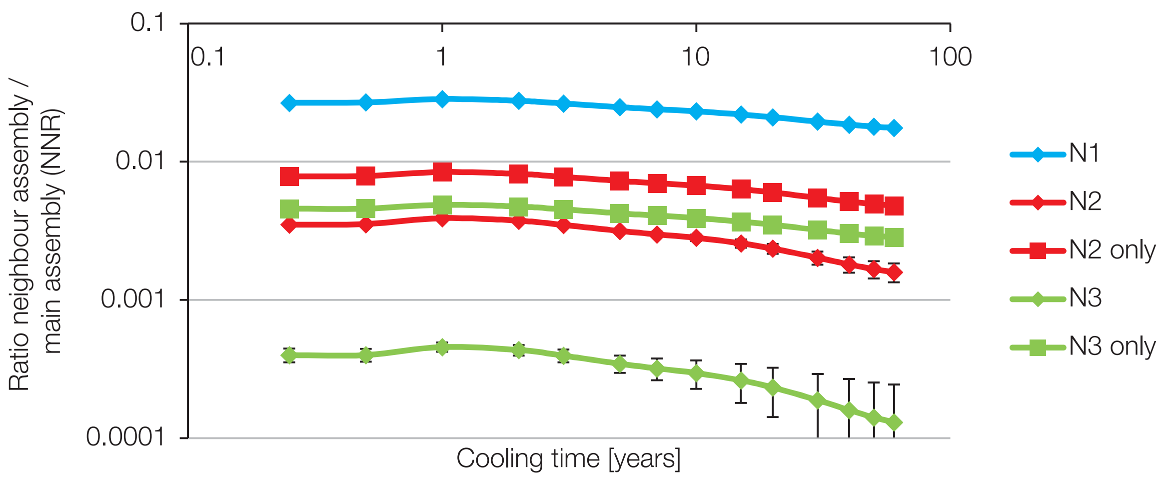


Figure 3: The magnitude of the near-neighbour effect as a function of cooling time, for BWR 8x8 assemblies. The N2 and N3 positions were simulated both for the situation that all neighbouring positions contained fuel (denoted N2 and N3, respectively), and for the situation that only two fuel assemblies were present, one at "main" and one at one neighbour position, (denoted "N2 only" and "N3 only", respectively). Error bars denote 1 σ uncertainty, and may be smaller than the point symbol for some data points.

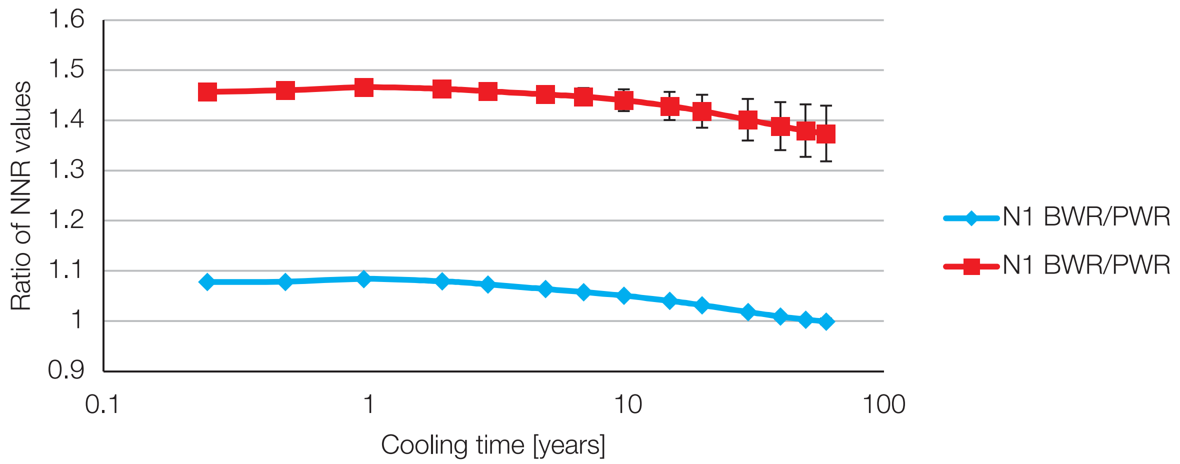


Figure 4: Ratio of NNR (see Eq. (1)) between BWR and PWR fuels as a function of cooling time, for the N1 and N2 neighbour storage positions. In the N2 case, all the N1 positions were occupied. Error bars denote 1 σ uncertainty, and may be smaller than the point symbol for some data points.

contributing to the measurable intensity at the event of measurement.

While the near-neighbour effect is noticeably affected by the cooling time, dependence on burnup is small, although a slight decrease with burnup is seen in the relative near-neighbour intensity at short cooling times. Note that while the near-neighbour intensity ratio changes little with burnup, the dependence of absolute Cherenkov light intensity on burnup is strong; high burnup implies high Cherenkov light intensity in both the fuel assembly emitting the radiation as well as in its neighbours.

2.3 Differences in the near-neighbour effect for BWR and PWR fuels

To investigate the differences in the near-neighbour effect for different fuel assembly configurations, the simulations in section 2.2 were complemented by simulations for a PWR case, where each PWR fuel was separated by a 5 mm steel wall, corresponding to closely stored fuel assemblies. In Figure 4, the ratios of NNR (see Eq. (1)) between BWR and PWR for the N1 and N2 positions are plotted, as a function of cooling time. The ratios are fairly flat at short cooling times, whereas for cooling times longer than 5 years, the near-neighbour intensity ratio, NNR, decreases more rapidly with cooling time for BWR as compared to PWR. This is likely due to a combination of the changing gamma spectrum and thus the mean free path of the gamma rays with time, and the differences in distance between an active rod and the water in the neighbouring assembly for the two configurations Accordingly, the near-neighbour effect depends on the fuel assembly configuration, and thus a compensation procedure may have to take the fuel type into account. Furthermore, one may note that NNR is higher for BWR fuel than for PWR fuel (the ratios between the fuel types is >1), given that both assembly types are stored closely.

2.4 Effects of fuel assembly spacing

To investigate the dependence of the near-neighbour effect on the storage distance between fuel assemblies, the simulations in sections 2.2 and 2.3 were complemented with an additional more spacious storage geometry, which corresponds to the storage situation for BWR fuels at the Forsmark Nuclear Power plant, for comparison with the experimental results reported in [10]. In these simulations, each fuel assembly was surrounded by a 2.5 mm-walled square steel channel, similar to the storage rack found at Forsmark. For the PWR simulation, the same relative fuel distance, as compared to fuel size, was simulated as in the BWR case, and the same wall thickness (2.5 mm steel) was used. The results for each simulated configuration are presented in Table 2 for 1-year cooled 40 MWd/kgU burnup fuel.

Storage configuration	Fuel size [mm]	Wall thickness [mm]	Fuel assembly centre-to-centre distance, [mm]	N1 intensity ratio (NNR)	N2 intensity ratio (NNR)
BWR close	130	4.0	135	2.84 ± 0.03%	0.39 ± 0.02%
PWR close	215	5.0	220	2.60 ± 0.01%	0.26 ± 0.01%
BWR spacious	130	2.5 + 2.5	195	1.43 ± 0.02%	0.54 ± 0.02%
PWR spacious	215	2.5 + 2.5	322	0.63 ± 0.01%	0.18 ± 0.01%

Table 2: The near-neighbour intensity ratio (NNR in Eq. (1)) for two fuel types and two different fuel centre-to-centre distances. In the spacious simulations, each fuel was surrounded by a separate steel wall. In the simulations for the N2 intensities, the N1 positions were occupied. The uncertainties are due to statistics in the Monte-Carlo simulations, and are presented for the 1 σ level. The simulated fuel assemblies had a cooling time of 1 year and a burnup of 40 MWd/kgU.

As can be seen in Table 2, the near-neighobur intensity in N1 is smaller for the more spacious storage geometry. For position N2 in the BWR case, the intensity becomes higher. The reason is that the assemblies in N1 positions strongly attenuate radiation travelling between the Main and N2 position in the close storage configuration. In the more spacious storage configuration, the N1 fuels interfere less with the radiation from the Main assembly, leading to a net increase in the N2 intensity ratio, despite the increased distance between them. For the PWR case, the result in larger absolute distances in the spacious storage geometry lower the N2 intensity ratio. Had the simulated distance been smaller, it may have been possible to observe the same effects as for the BWR case.

3. Comparison of simulations with experimental results

In 2012, a series of measurements were conducted at the Forsmark nuclear power plant, where the near-neighbour effect was quantified for the N1, N2 and N3 fuel positions [10], when all other positions were vacant. This was done by; (i) moving one active fuel assembly (defined as “Main” in Figure 2) to an isolated location and measure it to record the I_{main} intensity of Eq. (1), and; (ii) place it relative to a fresh fuel assembly in the N1, N2 and N3 positions and measure the subsequent intensity increase in the fresh fuel assembly, corresponding to $I_{neighbour}$ in Eq. (1). For details on these measurements, we refer to [9]. Here, the measured configurations have been simulated to provide an experimental benchmark of the simulation procedure, as further described below.

3.1 Measured and simulated geometries

In the measurements, the active assembly was of one BWR 10x10 type, while the fresh fuel was a different BWR 10x10 design. The properties of the storage racks at the Forsmark plant are accounted for in Table 2, denoted “BWR spacious”. The irradiation histories of the fuel assemblies were made available to the authors, courtesy of the operator, Vattenfall.

In the simulations, the fuel irradiation histories were used to calculate the assemblies’ gamma emission spectra by means of the ORIGEN code [5]. Using these spectra, simulations were run for the Forsmark storage configuration. However, the fuels simulated were BWR 8x8, while the irradiated fuels measured at Forsmark were all 10x10, including several part-length rods. The reason for not simulating the 10x10 fuel type was that information regarding the assembly manufacturer, and consequently the exact geometry data for the fuel types were unavailable for confidentiality reasons. However, the outer dimensions are similar for BWR 8x8 and 10x10 fuels, and both assembly types have a similar fuel to water ratio. Furthermore, differences in fuel pellet diameter results in different self-shielding by the rods, but an increase in absorption will lower both the Cherenkov light intensity in the assembly and in the neighbours, which partially compensates for the changes to the NNR. Consequently, the BWR 8x8 simulations may be considered to be representable also for 10x10 fuels in this context.

3.2 Results

In Table 3, the simulated near-neighbour intensities are compared to the intensities measured at Forsmark [10]. The overall agreement is good, especially for the N1 position where the near-neighbour effect is the strongest. One may note that the N1 position is slightly underestimated, while the N2 and N3 positions are overestimated. The deviations may be explained by differences between the simulated and measured fuel assembly configurations, or by measurement uncertainties. Further investigations would be required to draw more solid conclusions on the deviations.

Another result of these simulations is that for fuel assemblies in this storage geometry, the N2 intensity is not much affected by the presence or absence of a fuel in the N1 positions. In the case of both N1 positions occupied, the simulated N2 NNR is 0.41 ± 0.01 %, and with the N1 positions vacant it increases to 0.43 ± 0.01 %.

4. Detection limits in presence of the near-neighbour effect

As mentioned in section 1.1, partial defect verification using the DCVD relies on the fact that a 50% substitution of rods with non-radioactive content will reduce the Cherenkov light intensity by at least 30%, which, accordingly, is taken as the limit for partial defect. Fuel assemblies where measured intensities are more than 30% lower than predicted are detected as being subject to partial defect, whereas other assemblies pass the inspection. This

Neighbour position	Measured neighbour intensity	Simulated neighbour intensity
N1	1.25%	1.16 ± 0.02%
N2	0.36%	0.43 ± 0.01%
N3	0.12%	0.18 ± 0.01%

Table 3: Comparison of the measured near-neighbour effect (data from [10]), to a simulated near-neighbour intensity for a similar configuration, which was obtained using a gamma spectrum calculated with ORIGEN, taking into account the operator-declared fuel irradiation history. Simulation uncertainties are due to the Monte-Carlo nature of the simulations. Uncertainties in the measurements were not provided in [10].

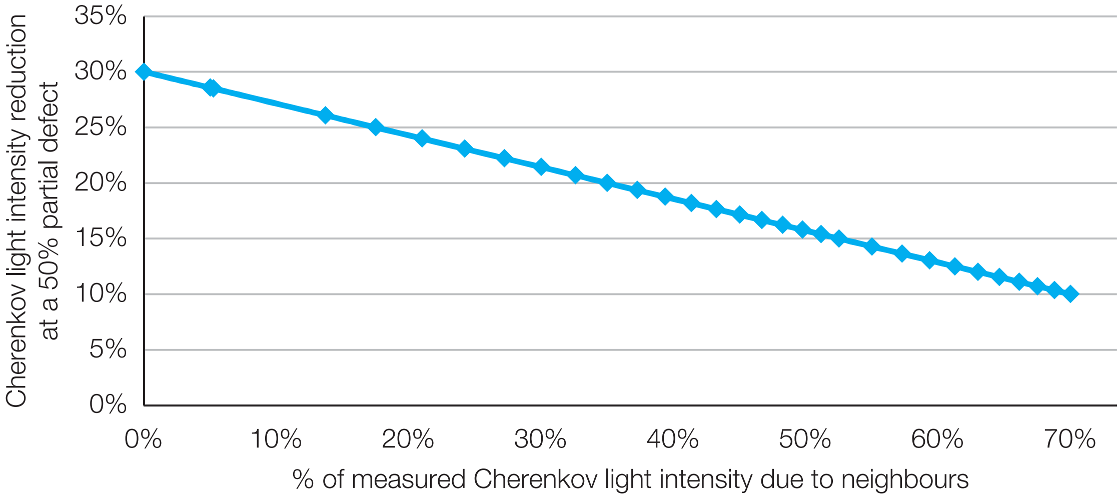


Figure 5: Calculated reduction of the 30% light intensity limit as a function of the near-neighbour intensity.

situation becomes slightly more complicated in presence of the near-neighbour effect, since the light being measured is partly caused by the fuel under study, and partly by the neighbouring fuels. As a consequence, the limit of 30% will be reduced when near-neighbour intensities influence the analysed data. This is shown in Figure 5, where the 30% intensity reduction is adjusted to also take into account the near-neighbour effect.

As a consequence of the data in Figure 5, the detection limit for partial defect at 30% lower intensity than expected would have to be lowered, unless the near-neighbour effect is corrected for. Lowering the detection limit would bring more stringent requirements on the accuracy of the models used for predicting the Cherenkov intensities as well as on the experimental precision that govern the accuracy of measured data in order to maintain the partial-defect detection capability. If the methods cannot meet these higher requirements, one must either allow a larger number of false alarms (using a lower threshold for partial defect according to Figure 5), or endanger the partial-defect detection rate (keeping the 30% intensity reduction threshold).

One situation where the near-neighbour effect would be particularly strong is when measuring a storage site with a population of neighbouring fuels with highly varying Cherenkov light intensities, due to largely varying burnups and cooling times. In such a situation, low-intensity fuel assemblies will be more strongly affected by the near-neighbour effect as compared to high-intensity ones. If the effect is not corrected for, the fuel intensity predictions will systematically underestimate the intensity of low-intensity fuels, while over-estimating the measured intensity of high-intensity fuels. Referring to Figure 5, considering a low-intensity assembly where as much as 50% of the intensity comes from neighbouring fuels, a diversion of 50% of its fuel rods may only cause a 15% decrease in measured intensity. It is doubtful that the current experimental and predictive methods may be further developed to offer the precision required for confident detection in such extreme cases, unless the near-neighbour effect is included in the analysis.

In conclusion, to avoid changing the detection threshold while maintaining the partial-defect detection capability, methods for correcting for the near-neighbour effect should be considered. Such correction methods are further discussed below.

5. Methods for correcting for the near-neighbour effect

In this section, two methods for correcting for the near-neighbour effect are presented. The basics of both methods are that each measured intensity can be expressed as a sum of the intensity from the assembly under study, I_0, and the intensities from its nearest neighbours:

$$I_{measured} = I_0 + \sum_i \left(I_i \cdot \varepsilon_i \right) \tag{2}$$

Here, ε_i denotes the ratio of the intensity that neighbouring assembly i emits in the studied assembly to the intensity it emits in its own position, (I_i). One may note that ε_i goes in the opposite direction compared to NNR defined in Eq. (1), but for symmetry reasons their values should be identical. The two methods presented below differ in how the ε_i are determined, where section 5.1 describes a method based on experimental data and section 5.2 describes a simulation-based method.

5.1 Least-squares fitting of experimental data

In [10], an experimental method to assess the near-neighbour effect was tested on a set of BWR fuel assemblies measured at Clab, under the conditions shown in Figure 2. The proposed method uses Eq. (2), limited to neighbours in relative positions N1 and N2 (referring to Figure 2). The method suggests collecting experimental intensities for the complete set of fuels in one storage rack and determining the ε_{N1} and ε_{N2} factors by performing a least squares fit of Eq. (2) for the experimental data set, based on predicted intensities. These fitted ε_{N1} and ε_{N2} can then be used to predict the measured intensity of an assembly, given a prediction of the intensity of the assembly and its neighbours, or alternatively to subtract the intensity caused by the near-neighbour effect from the measurements.

Ref. [10] presents values of ε_{N1} and ε_{N2} obtained from fitting of the experimental data set. In this work, simulations of the storage conditions at Clab for the assemblies under study have been performed to provide an independent evaluation of the deduced values. A comparison of the simulated and the experimentally fitted intensities from [10] is shown in Table 4.

Neighbour	Simulated ε_i	Fitted ε_i
N1	1.82 ± 0.05%	16%
N2	0.17 ± 0.02%	9.5%

Table 4: Comparison of the simulated near-neighbour intensity for closely-stored BWR fuels (as shown in Figure 2) and the fitted values reported in [10].

Table 4 shows poor agreement between simulated and fitted values, and the simulations suggest that the fitted values overestimate the near-neighbour effect by almost an order of magnitude. Considering the relatively good agreement between simulations and measurements shown in Table 3, there is reason to suspect that the fitting procedure may not be adequate to accurately quantify the near-neighbour effect. Probable reasons for this deficit are that the fit is based upon a rather small set of fuels, and that the equation system may be ill-conditioned, making it sensitive to stochastic noise. One may assume better results if larger data sets are used or if constraints are introduced on the near-neighbour intensities, based on expected ratios.

5.2 Simulation-based corrections

As shown in section 3, simulations can provide relatively accurate estimates of the relative intensities from neighbouring fuel. However, since the near-neighbour simulations are time-consuming, a method is needed to take the near-neighbour effect into account in a quicker way, which can be used by inspectors in the field. Here, a solution is suggested, where the near-neighbour effect is parameterised as a function of fuel geometry, fuel centre-to-centre distance, and gamma-ray energy. The parameterisation would be based on large simulations done in advance, allowing for fast deployment for in-field inspections.

Based on the results presented here, primarily the N1 and N2 positions would need to be considered when assessing the near-neighbour intensity, and only rarely will the N3 position be significant. Given the irradiation history, or at minimum the burnup and cooling time, of an assembly and all its neighbours, ORIGEN can be used to assess the gamma-ray energy spectrum of each fuel assembly. By binning the spectrum, it is possible to run simulations with initial gamma rays from each bin, to assess the near-neighbour intensity of gamma-rays of each energy. These simulations will have to be done for a large number of energy bins, for each fuel assembly configuration, and for several fuel centre-to-centre distances. The results will be the magnitude of the near-neighbour effect $\epsilon_{i,j}$ for a fuel at neighbour position i and for gamma rays with energy in bin j. These simulations can be done in advance, and only have to be done once for each case.

To calculate the near-neighbour intensity at the event of measurement, the user selects the pre-calculated $\epsilon_{i,j}$ values applicable for the fuel type and storage situation applicable to the measurement situation. These values are combined with the calculated, binned gamma-ray emission spectra of the neighbouring fuels, based on the operator declared fuel declarations. If the binned spectrum of a fuel is given by S_j for bin j, the intensity caused by one neighbour at position i ($I_{neighbour,i}$) can then be calculated as:

$$I_{neighbour,i} = \sum_{j=1}^{\#bins} \epsilon_{i,j} \cdot S_j \quad (3)$$

The total near-neighbour intensity contribution in an assembly is then the sum of the intensity of all present neighbours, each calculated using Eq.3. This value can either be added to a predicted assembly intensity I_0 to give a prediction of the measured intensity; alternatively it can be subtracted from the measurement to obtain an experimental value of assembly intensity I_0 twithout neighbours.

6. Conclusions and outlook

Fuel assemblies in wet storage are often verified using the Digital Cherenkov Viewing Device, which enables inspection without requiring the fuel to be moved to an isolated measurement location. Since the fuel assemblies are stored closely, gamma rays from one assembly may enter a neighbouring assembly and create Cherenkov light, the so-called near neighbour effect. This paper describes how simulations can be used to estimate the magnitude of the Cherenkov light intensity that occurs in a neighbouring position due to the near-neighbour effect. The simulations have been validated using experimental data. The near-neighbour effect will be particularly influential in cases where long-cooled, low-burnup fuels containing relatively low activity levels are stored next to short-cooled, high-burnup fuels containing relatively high activity levels.

It has been shown that the partial-defect detection limits may need adjustment unless the near-neighbour effect is corrected for. Two possible methods for such corrections have been described; one method based on experimental data and one simulation-based method. Building on the fact that simulations have proven capable reproducing experimentally recorded near-neighbour intensities, the latter method is recommended, and a methodology allowing for quick in-field use has been presented. The methodology is based on extensive, time-consuming simulations, which are done in advance to create parameterisations specific for storage configurations, assembly types and gamma-ray energies. These parameterisations may then be used for fast assessment during inspection.

While some experimental data is available regarding the near-neighbour effect, more is required to verify the simulations performed, and to assess the performance of the suggested method for predicting the near-neighbour effect. Knowing the accuracy of the near-neighbour prediction model will allow for higher limits to be set regarding what magnitude of near-neighbour effect can be tolerated in the measurements, which increases the partial-defect detection performance of the DCVD. Additional experimental data will also be useful for further refining the near-neighbour prediction model, which can further enhance the DCVD partial defect detection capabilities.

The studies presented in section 3 suggest that it may be possible to e.g. treat all BWR fuel assemblies as being identical with respect to the near-neighbour effect. Thus, it may be possible to simulate only a few selected fuel geometries of widely varying configuration, and use those simulations to assess the near-neighbour effect for all fuel types. This would greatly reduce the amount of simulations necessary to perform to parameterize the near-neighbour effect, but further studies are required to assess what uncertainties are introduced by this simplification.

7. Acknowledgements

This work was supported by the Swedish Radiation Safety Authority (SSM), under contract SSM2012-2750. The computations were performed on resources provided by SNIC through Uppsala Multidisciplinary Centre for Advanced Computational Science (UPPMAX) under project p2007011.

8. References

[1] J. D. Chen et.al, “Spent fuel verification using a digital cerenkov viewing device,” 8:th International Conference on Facility Operations - Safeguards Interface, 2008.

[2] J.D. Chen et al., “Partial defect detection in LWR spent fuel using a Digital Cerenkov Viewing,” *Institute of Nuclear Materials Management 50th annual meeting,* 2009.

[3] S. Rolandson, “Determination of Cherenkov light intensities from irradiated BWR fuel,” IAEA task ID JNTA0704, SKI Report #: SE 1-94, 1994.

[4] E. Branger, S. Grape, S. Jacobsson Svärd, P. Jansson and E. Andersson Sundén, “Comparison of prediction models for Cherenkov light emissions from nuclear fuel assemblies (accepted manuscript),” *Journal of Instrumentation,* 2017.

[5] S. Bowman, L. Leal, O. Hermann and C. Parks, “ORIGEN-ARP, a fast and easy to use source term generation tool,” *Journal of Nuclear Science and Technology,* vol. 37, pp. 575-579, 2000.

[6] E. Branger, S. Grape, P. Jansson and S. Jacobsson Svärd, “Improving the prediction model for Cherenkov light generation by irradiated nuclear fuel assemblies in wet storage for enhanced partial-defect verification capability,” *ESARDA Bulletin,* vol. 53, 2016.

[7] Agostinelli, S, et al (the Geant4 collaboration), “Geant4 - a simulation toolkit,” Nuclear Inst. and Meth. in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, Volume 506, Issue 3, 1 July 2003, Pages 250-303.

[8] S. Grape, S. Jacobsson Svärd and B. Lindberg, “Verifying nuclear fuel assemblies in wet storage on a partial defect level: A software simulaiton tool for evaluating the capabilities of the Digital Cherenkov Viewing Device,” Nuclear inst. and Meth. A, Volume 698, 11 January 2013, Pages 66-71, ISSN 0168-9002, 10.1016/j.nima.2012.09.048.

[9] S. Vacarro, S. Tobin, A. Favalli, B. Grogan, P. Jansson, H. Liljenfeldt, V. Mozin, J. Hu, P. Schwalbach, A. Sjöland, H. Trellue and D. Vo, “PWR and BWR spent fuel assembly gamma spectra measurements,” *Nuclear Instruments and Methods in Physical Research Section A: Accelerators, Spectromenters, Detectors and Associated Equipment,* pp. 208-225, 2016.

[10] D. Parcey, E. Sundkvist, J. Dahlberg, K. Axell, R. M Kosierb, B. Lindberg and S. Grape, “Determining the effect of adjacent spent fuel on Cherenkov light measurements,” in *Institute of Nuclear Materials Management symposium*, 2012.

iQS-01 InCounter Quantification System for Non-Destructive Assay: Report on Testing Procedures and Results for Device Performance and Holdup Quantification Model

Greg Peacock and Brent McGinnis

Innovative Solutions Unlimited LLC (inSolves)
1862 Shyville Rd., Piketon, OH 45661, USA
E-mail: gregorypeacock@insolves.com

Abstract:

The InCounter Quantification System and the associated iQS-01 software package are designed to support characterization of nuclear materials retained in process equipment associated with nuclear deactivation and decommissioning (D&D) processes at gaseous diffusion plants. The InCounter is designed to perform automated measurements inside process piping, resulting in a more efficient and reproducible estimate of process holdup. Each InCounter features a thallium doped sodium iodide scintillation (NaI(Tl)) gamma detector, a LIDAR sensor for surface mapping, and a video camera for visual inspection. This combination of sensors allows the user to make a better-informed decision regarding the nature and disposition of holdup deposits. The output from multiple sensors also allows the user to better understand variables that may cause biases in measurement results.

Several testing procedures were developed to assess the performance of the InCounter to initial design specifications. These testing procedures included verifying device physical capabilities for performance limits and positional accuracy, measurement quality and subsequent analysis algorithms for mass quantification and Total Measurement Uncertainty (TMU), and software reliability in accordance with the Nuclear Quality Assurance (NQA-1) Software Quality Assurance (SQA) regime. The InCounter underwent testing at various points in development at multiple facilities in the United States. This included testing at the Oak Ridge National Laboratory (ORNL) Safeguards Lab to evaluate initial detector characteristics, at the Portsmouth Gaseous Diffusion Plant (PORTS) to assess physical characteristics, at the Paducah Gaseous Diffusion Plant (PGDP) using uranium surrogate sources for measurement quality and analysis accuracy, and in-house at the Innovative Solutions Unlimited, LLC corporate facility to assess full integration testing.

Keywords: Non-Destructive Assay; Robotics; Gamma Detection; Non-Proliferation; Holdup.

1. Introduction

The iQS project is a design effort to produce a customizable platform for measurement acquisition and data tracking in regulated environments. The core software suite allows for a variety of endpoint sensors to be connected to an autonomous device that is controlled by pre-defined and scripted software routines. The current offering for such devices is the InCounter, an internal-to-pipe traveling robot that features an NaI(Tl) gamma detector, a light detection, imaging and ranging (LIDAR) sensor, and a camera as its sensor platform. The software system tracks and manipulates the resultant signals from these sensors from acquisition through final reports, thereby streamlining the analysis process and avoiding transcription errors [1].

From the beginning of the InCounter design process, various testing procedures were performed to ensure the software and hardware components satisfy defined quality assurance requirements. The objective of the different tests varied depending on the current status of the design. Subsequent testing protocols were developed based on the results of previous tests modifications made to the InCounter design.

A system capabilities description and design narrative are provided first to describe the baseline test requirements. Then, as a precursor to formal testing, exploratory measurements made at Oak Ridge National Laboratory (ORNL) are described.

The first formal test of InCounter capabilities was performed with the assistance of the Portsmouth Gaseous Diffusion Plant (PORTS) during the beta stage of design. The physical characteristics of the InCounter were tested at PORTS. Subsequent verification and validation (V&V) testing of version 1.0 of the combined hardware and software system was performed at inSolves main office. After ensuring proper functioning to requirements, the U-235 quantification model accuracy was tested using uranium standards at the Paducah Gaseous Diffusion Plant (PGDP) using version v1.1-beta of the system. Each set of testing is described sequentially to explain how each builds upon the successes and lessons learned of the previous testing.

Figure 1: The InCounter cart resting on a stand. From right to left, the camera, lights, LIDAR, detector and surrounding detector housing, and cart chassis.

2. Design Inspiration

The iQS project has its roots in brainstorming improvements to the Ortec Holdup Measurement System 4 (HMS4) in 2014 after it was recognized that the design of small detection hardware would be sub-optimal for characterizing holdup deposits of low-enriched uranium. Initial InCounter designs focused mainly on hardware improvements with some exploration of a modified version of the Generalized-Geometry Holdup (GGH) method. The current form of the project began in 2016 after studying the D&D projects at gaseous diffusion plants in both Portsmouth, OH, USA and Paducah, KY, USA.

The initially chosen measurement targets for the hardware system were long, straight pipes of varying sizes. Although holdup is more common at valves, elbows, expansion joints, or other uneven surfaces, these targets were chosen as the simplest measurement geometry to traverse with robotics and as the most applicable measurement need for the InCounter design, due to the fact that thin-film deposits for long sections of straight pipe can dominate the process holdup inventory given the large amount of surface area relative to that of the valves, elbows and expansion joints. Furthermore, although the reduction in overall measurement uncertainty has not been determined, improvements are expected because the deposit distribution can be observed and a more direct measurement can be performed without having to make corrections for the wall of the pipe. Each pipe is to be measured for ^{235}U holdup, typically in the form of UO_2F_2 attached to walls or within components [2]. An initial prototype system was designed to travel on rails external to the pipe (see Figure 2). However, this design was modified into the InCounter to permit in-pipe travel to simplify the engineering design of the InCounter system and improve the quality of the measurement.

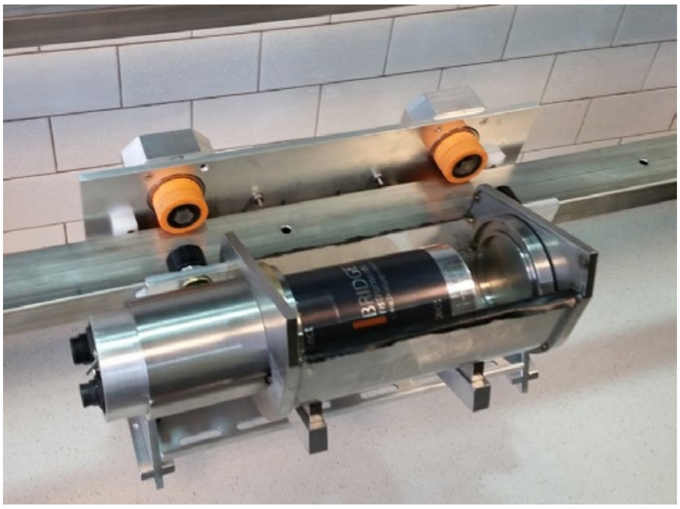


Figure 2: The GammaTrak: an early concept of external measurement system that became a precursor to the InCounter. It travelled on tracks and featured a rotating detector shield for field-of-view modifications.

Concurrent with the InCounter hardware development was the initial design of the software for controlling the data stream associated with acquisition, analysis, data package review, and management approval. Existing data processing methods included steps of handwritten transcription in the field and potentially weeks to months for generation and approval of the final data package. It was recognized that a networked, computer-centric data pathway would vastly speed up the process while also including automatic error checking. In addition, this allowed the custom developed ^{235}U mass quantification algorithms to be integrated directly into the data pipeline.

2.1 Generalized-Geometry Holdup

The procedure for mathematically determining the ^{235}U mass for a process deposit from the associated gamma spectrum has typically used the Generalized-Geometry Holdup (GGH) method. GGH models the deposit as a point, line, or area source. It relies on experimentally determined calibration data to provide a conversion coefficient from gamma peak counts to mass of the associated isotope [2].

It was noted that several aspects of the GGH method could be improved, including inverse-squared distance assumptions and source angulation during calibration [3]. In addition, correction factors for material attenuation do not necessarily account for thickness changes at an angle to the target [4].

When the focus of the iQS project became the internal measurement of pipes via the InCounter, a new model was necessary. The two-dimensional (2-D) geometries offered by GGH in the form of points, lines, and areas with finite width corrections, or a combination thereof with varying distances along the same detector axis, were not sufficient for three-dimensional (3-D) representation of material within the pipe. The new modelling method borrowed from GGH's basis of utilizing experimental responses to a standard to calibrate the instrument, but was expanded to applicable cylindrical geometries [4][5].

2.2 Quality System for Non-Destructive Assay

At the PORTS and PGDP sites, NDA measurement quality is ensured by the Quality System for Non-Destructive Assay Characterization (QSNDA) program document. QSNDA provides requirements for qualification of new instruments and the acquisition of new measurement data, such as requirements for duplicate measurements and control check frequency [6].

The iQS system is designed to support the stringency of calibration and measurement quality required under QSNDA. Several of the requirements of QSNDA resulted in the definitions found in the iQS system requirements document [7].

3. iQS-01 Project and InCounter Descriptions

3.1 InCounter Quantification System

The InCounter is designed directly to its intended measurement environment: the inside of long, straight pipes. This measurement environment was considered in the chassis design, sensor selection and placement, and the contamination control features.

Movement of the InCounter is accomplished via connected drive wheels at the front of the cart that have large, angled tires to improve traction by increasing contact against the potentially irregular contours of the pipe. Position tracking is handled by an interior encoder system attached to the free-spinning back wheels. The back wheels are designed with very thin wheels to provide a consistent contact point that allows for positional calibration [1].

Within the chassis, motors are located in the front compartment beneath the detector housing. The computing and power systems are located in the rear compartment. Power is provided by a hot-swappable lithium ion battery commonly used for hand-held power tools [1].

3.1.1 Sensors

The InCounter features three sensors: a NaI(Tl) gamma detector with a 5-cm. by 5-cm. crystal, a camera with associated lighting, and a LIDAR sensor. The sensor array is positioned along the axis of the pipe, allowing easier modeling of deposits using radial symmetry [1].

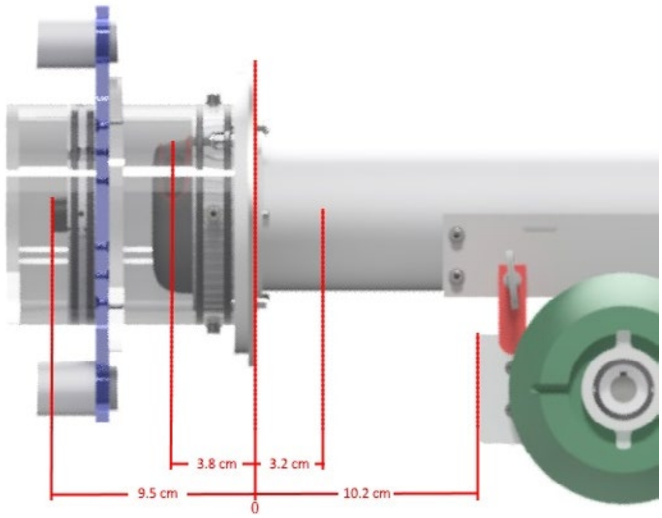


Figure 3: Sensor locations relative to a physical zero location. From left to right: camera, LIDAR, center of detector crystal, and the front of the cart.

Concurrent gamma detector is the key component of the InCounter system and features a field-of-view (FOV) of approximately 15 cm in front of and 15 cm behind the center of the NaI crystal. The detector is not collimated, but the chassis of the cart is designed to accommodate the addition of collimation if needed. Due to the absence of collimation, the detector response decreases with distance, which is accounted for in the detector calibration procedure. The detector exclusively provides input to the mass quantification algorithms discussed in section 3.2.1 [5].

The LIDAR sensor and camera provide the ability to perform visual inspection of the measured deposits and of internal pipe condition. The LIDAR sensor rotates along the pipe axis, providing a 2-D "slice" of the surface. Many of these slices can be combined into a point cloud representing the interior physical contour of the pipe. Due to limitations in the precision of LIDAR measurements, the point cloud data is not used for quantification analyses, but it does serve as a visual aid to operators and analysts for determination of future measurement plans and the application of the correct model selected for quantification [1].

3.1.2 Atmosphere

The chemical reactions that give rise to the typical UO_2F_2 deposit also release hydrofluoric acid (HF) as a byproduct, which can build up in the internal volume of the pipes and result in a corrosive atmosphere. In an effort to protect the more delicate sensor and computing hardware, the InCounter is designed to be resistant to HF.

The chassis and detector housing are made of aluminum to ensure low weight and resist corrosion without significantly attenuating the detector. The gamma ray transmission rate through the aluminum housing is approximately 95% at the 185 keV gamma ray of ^{235}U. Fasteners are made of stainless steel [1].

Seams between the chassis plates are sealed with a corrosion resistant caulk. The tires of the InCounter are a custom blend material and are 3-D printed [1]. The material was subject to hydrofluoric acid bath testing without losing integrity or corroding.

3.2 iQS-01 Software Backbone

The iQS software suite serves two purposes: to provide control for the InCounter or other hardware endpoint and to serve as the data pipeline. It is split into different programs: a control server within the InCounter that handles wireless networking, sensor interfacing, and low-level InCounter control functions; a control application allowing an operator to run the InCounter; and the Workbench, which allows an analyst to review data, run analyses, generate reports, and access administrative features of the suite [1].

Control of the InCounter can be done manually, but usually relies on a custom scripting language that grants the unit autonomy during data collection. This script can control movement and sensor acquisition and respond conditionally based on sensor input. Meanwhile, all data is sent to a central database in real-time, allowing operators and analysts to review data within seconds of collection [8].

3.2.1 Mass Quantification Methodology

Analyses are run within the Workbench, either on analyst command or automatically upon data collection. The mass quantification takes the results of the detector response calibration and the measured gamma spectra to return an estimate of the mass of ^{235}U in the pipe (units: g ^{235}U per foot (30.5 cm) pipe) [5].

The method begins with the detector response calibration. A small, mixed gamma source set consisting of exempt quantities of ^{133}Ba, ^{57}Co, ^{60}Co, and ^{137}Cs is moved around the internal surface of a test pipe. Spectra are collected for a sample of source locations within the pipe, spaced evenly around the circumference by angle and along the axis of the pipe by distance. The detector crystal is taken as the origin in these measurements, and on-axis measurements are taken over a few meters on each side of the detector to establish FOV boundaries and detector response. This calibration varies by the internal radius of the pipe, which depends on the construction standards. On each spectral peak, Gaussian fits are applied, and the counts in each peak are compared to the decay corrected source assays to determine a total efficiency of detection at that energy and location, including both detector and geometric efficiencies. High order (4th degree or higher) polynomial fits are applied to the total efficiencies as a function of energy. Polynomials were chosen after exploring several options including exponential and power functions. Although polynomials do not permit extrapolation, the energy range of interest is well bounded by the calibration peaks. These efficiency curve fits are used to interpolate the 186 keV peak efficiency needed to calculate the ^{235}U mass from the InCounter measurement [5].

On initial run of quantification for a set of measurements, default models of the physical deposit shape are created by the software. This default model assumes a thin film of holdup material over the entire interior surface of the pipe. Using a thin-film assumption, self-attenuation is not included in this first calculation, allowing for rapid calculation for a large number of measurements. After expert review of measurements of interest, custom models can be defined by an analyst to account for different distributions and thicknesses of deposit. Models are digitally represented as collections of voxels in cylindrical coordinates having a set thickness and spanning finite lengths of θ around the pipe circumference and z along the pipe axis. Each voxel corresponds to the location of a calibration response point with the associated efficiency curve. Self-attenuation of the deposit and attenuation of the cart chassis are included in the evaluation of each voxel's expected response to the detector, which are all combined to return a single quantification coefficient for the model [4].

Similar to GGH, the quantification coefficient is multiplied by the result of the peak analysis of ^{235}U - in this case, the Gaussian fitting of the 186 keV peak - to return total estimated mass.

4. Oak Ridge National Lab Testing – Detector performance

The Oak Ridge National Lab (ORNL) has several uranium standards in its Safeguards Lab that were key to establishing the basis of detector reliability on the InCounter (see Table 1 on next page). When testing at ORNL, the InCounter was in a v1.0-beta design, and the iQS software was still in active, pre-v1.0-alpha development. The purpose of this testing was to verify detector performance and compare uranium standards to mixed source sets, as well as run initial tests of the mass quantification algorithms that were updated in later versions of the system.

Source No.	^{235}U Weight Percent	Total Mass ^{235}U (g)
1	0.3166 ± 0.0002	0.52
2	0.7119 ± 0.0005	1.2
3	1.9420 ± 0.0014	3.28
4	2.9492 ± 0.0021	4.99
5	4.4623 ± 0.0032	7.54
6	20.107 ± 0.020	39.12 ± 0.04
7	52.488 ± 0.042	101.81 ± 0.10
8	93.1703 ± 0.0052	181.12 ± 0.12

Table 1: Sources used as standards during ORNL testing. All numbers provided by Safeguards Lab source certificates.

The sources listed in Table 1 come from two sets: SRM 969 for low assay sources, and CRM 146 for the 20% and greater assay sources. Specifically included are SRM 969-031 (1), SRM 969-071 (2), SRM 969-194 (3), SRM 969-295 (4), SRM 969-446 (5), CRM 146 – NBL 0021 (6), CRM 146 – NBL 0022 (7), and CRM 146 – NBL 0023 (8).

4.1 Mixed Source Qualification

The mixed source set used to calibrate the NaI(Tl) scintillation detector was chosen to allow easy licensing and location flexibility for calibration procedures. In addition, by utilizing a fitted efficiency curve, other isotopes besides ^{235}U can be measured. The challenge is in qualifying exempt quantity, non-uranium sources for use in a uranium measurement system. Measurements were taken of ^{235}U standards at varying enrichments, and efficiencies of each measurement were compared to the evaluated efficiency generated by fitting the efficiencies of the mixed gamma source: see Figure 4, in which measured efficiencies as red dots are compared to the expected efficiency for a specified energy represented by the blue line. All uranium sources from both SRM 969 and CRM 146 were used to conduct the comparison. Although the SRM 969 sources are similar to enrichment ranges expected for field measurements using the InCounter system, both sets provide a full enrichment range.

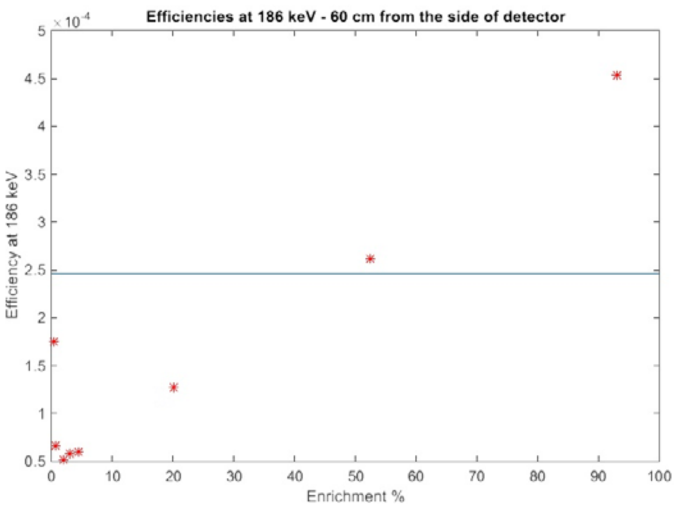


Figure 4: Total efficiencies of uranium standards as a function of enrichment. The blue line represents the value calculated using the mixed source set.

Figure 4 shows significant deviation from expected in the measurement efficiencies at distance. All efficiencies were calculated accounting for self-attenuation of the source material and attenuation of the canister window. The peak fits for the low enrichment sources had much smaller goodness-of-fit values than those for the high enrichment sources, likely due to the distance of measurement. Uncertainties for the measurements are comparatively larger at lower enrichment due to lower counts, further increased by the propagated fitting uncertainty that also increases with lower counts. The results from this test, however, did indicate that improvements were needed in the efficiency model across all ranges of enrichment. This issue was addressed in the next phase of development.

Preliminary tests of mass calculation were also run using two uranium oxy-fluoride sheet sources that contain 11.108 g ^{235}U and 11.113 g ^{235}U respectively. These sheet sources are made from approximately 93% enriched uranium. The physical dimensions of the sheet sources are 23 cm x 46 cm. The sheets were placed inside a pipe and the InCounter was run past them such that the extent of the sheet source was outside of the FOV of the InCounter's starting and ending positions. Calculated masses are found in Table 2.

Scenario	Calculated mass ^{235}U (g)
1 sheet (11.113 g ^{235}U)	8.99
2 sheets	23.875
2 sheets (2nd Pass)	24.060

Table 2: Results from early test of mass calculation algorithm with ORNL sheet sources.

For the first scenario in Table 2, the difference in the measured versus declared ^{235}U mass for the single sheet measurement indicates that the in-development quantification model calculated a mass value that was biased low. However, runs 2 and 3 show calculated results very close to the expected mass, which is inconsistent with the first run. In the first scenario, a single source sheet was in the bottom of the pipe, while in the two-sheet scenario, one source sheet is at the bottom of the pipe and one is at the top. Images show the top sheet hanging down from the pipe several centimeters, which suggests the measured response was artificially inflated. The conclusion that the quantification was under-estimating is applicable to all three scenarios.

The uranium efficiencies being lower than expected from the mixed source set naturally lead into the consistent underestimation in the quantification algorithm. As the iQS software was still in early development, changes were made to both systems in later software versions. The results of these changes can be seen in Section 7.

5. Portsmouth Testing – Physical Characteristics

With the assistance of quality assurance (QA) engineers from the PORTS site, the physical characteristics of the InCounter were tested. At the time of testing, the InCounter was a final v1.0-beta design and the iQS software was in early v1.0-beta, with some ongoing feature development.

The testing focused on the cart movement and positional tracking accuracy, battery functional limits at different feature loads, lighting limits, and visual inspections of sensor data quality [9].

Figure 5: Setup of the piping for testing of battery. The cart ran until the end of the pipe, then reversed to the point of origin, summing up total distance.

The first testing focused on battery lifetime as a function of operational features. Three scenarios were run: 1) continuous movement at 2.5 cm per second with no data collection, 2) continuous movement at 2.5 cm per second with lights on and data collection enabled on all sensors, and 3) movement punctuated with stops each 30.5 cm for 30 second spectrum acquisition and full data collection during movement. Results of these tests are presented in Table 3 [10].

Scenario	Battery life (hrs. & mins.)	Distance traveled (meters)
1	4:54	301
2	2:44	90
3	2:25	38

Table 3: Battery lifetime and maximum distance traveled as a function of features enabled.

Testing involved visual inspections of the real time data collection, available for viewing in the iQS InCounter Controller application. In addition, manual and scripted control methods were verified for positional accuracy. Results were consistently accurate to within 0.5 mm. of the arbitrarily determined target location, as verified by a traceable tape measure in the bottom of the pipe test stand [10].

Camera illumination was tested at 25%, 50%, 75%, and 100% lighting power at locations 30.5 cm., 61 cm., and 91.5 cm. in front of the InCounter cart. A maximum lighting of 685 lumens was achieved at 100% power at a distance of 30.5 cm. [10], and visual inspection confirmed the estimated maximum illumination distance of approximately 6 ft. [1].

6. In-house Testing – V&V and Integration

The successful testing of InCounter functionality and the completion of the v1.0 of the iQS software led to the administration of the verification and validation suite of the combined system and software quality assurance program. The SQA package was designed under NQA-1 and features the System Requirements Document, the collection of V&V tests, a traceability matrix to ensure V&V test coverage to the requirements, a QSNDA acknowledgement checklist, and a software interfaces chart for the interactions and interdependencies of development tools.

The V&V tests were run manually on each InCounter produced and for each version of the iQS software due to the nature of the integrated hardware-software design where neither functions fully without the other.. InCounter calibration activities are also included within the V&V tests. The iQS v1.0 V&V tests included:

- Test 0: Comprehensive,
- Test 1: Cart Calibration,
- Test 2: Cart Movement and Positioning,
- Test 3: Battery Endurance,
- Test 4: Sensors and Scripting,
- Test 5: Reports and Calibration Certificates,
- Test 6: Cart Inspection Test,
- Test 7: Uranium Data Analysis Test,

where Test 0 is included for compliance with internal SQA requirements [11]. These tests are categorized as end-to-end tests of various features of the system.

Test 7 is of special note due to the lack, at the time of the V&V testing, of comparison to experimentally simulated deposit measurements. As such, the data used for testing was generated from MCNP5 [12] simulations of ideal deposits.

6.1 Deviations from Requirements

One deviation from requirements was found in Test 2: Cart Movement and Positioning, in which the position of the cart as tracked by the iQS software remained constant after repeated back and forth runs but the real position varied irregularly by a few percent of the total distance of the run. This was found to be the consequence of physical modifications made to the position tracking to support movement over rough terrain in which the tracking wheels may not be on the ground at the same time. This issue was deemed to be less of a problem than potentially complete loss of tracking in uneven terrain. Designs to handle both scenarios are currently still in progress.

A further issue was found in the mass quantification algorithms when the input MCNP model was of a thick accumulation of UO_2F_2 in the bottom of the pipe. Test 7 includes three scenarios: 1) a fully uniform distribution of

30 g ^{235}U all the way around the internal surface of the pipe, 2) a uniform distribution of 20 g of ^{235}U in the bottom half of the pipe, and 3) an accumulated pile of 20 g of ^{235}U at the bottom of the pipe. The results of the quantification algorithm from the generated spectra were 33.47 g, 21.27 g, and 41.77 g respectively [11]. A study of the accumulated model revealed invalid assumptions, and a more complex model was devised to use prior knowledge inputs and iteration to remove the need for those assumptions.

Other aspects of the testing showed the v1.0 iQS software and InCounter to be sufficiently ready for version release.

7. Paducah Testing – U-235 Quantification

To make up for the lack of experimental measurements related to Test 7 of the V&V process, experiments to test the quantification algorithms were performed at PGDP using several uranium sources measured with different methods. The InCounter used was release v1.0, and the iQS software was v1.1-alpha, which was undergoing active development. This allowed quick turn-around in case adjustments needed to be made to the quantification algorithms.

At the initiation of the test it was initially determined that positioning issues similar to those identified during the V&V testing had not been resolved. The addition of a ballast to add weight to the rear wheels resolved this issue. The added weights were incorporated into future versions of the InCounter hardware.

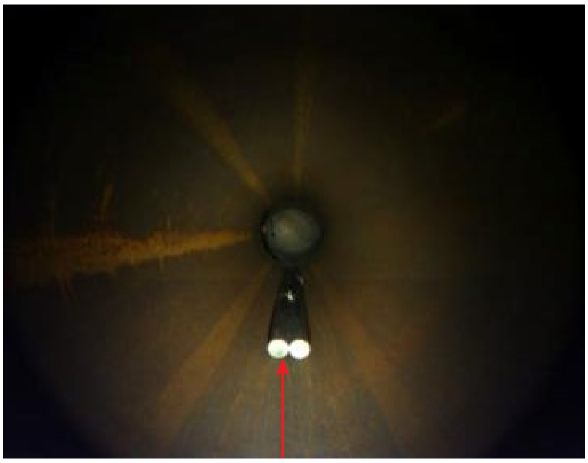


Figure 6: Image of four rod sources inside a capped steel pipe with an internal radius of 15 cm. These images were taken using the onboard camera and lighting.

Measurements using uranium sources began with various arrangements of rod sources used to simulate accumulated deposits. The rods were spread throughout 4.6 meters of pipe at lengths of either 20.3 cm or 40.6 cm (two rod-lengths), sometimes side-by-side (see Figure 6). Measurements were made using several predefined scripts:

1. The cart stopping every foot (30.5 cm) for spectrum collection while also recording scanning data during movement.
2. The cart stopping only when a threshold of countrate had been reached, otherwise always recording scanning data during movement.
3. The cart not stopping at all, instead traveling at various speeds while only recording scanning data.

Because the accumulated deposit model for quantification was identified to be non-functional during the v1.0 V&V testing (see Section 6.1), quantification on these sources was not performed. As expected, quantification of an accumulated source using a uniform model returned results significantly smaller than declared, confirming that an appropriate model is required to properly estimate the deposit mass.

The next step in testing was the usage of "mouse-pad" sources, which were wrapped around the pipe and held in place with a small piece of metal tubing (see Figure 6).

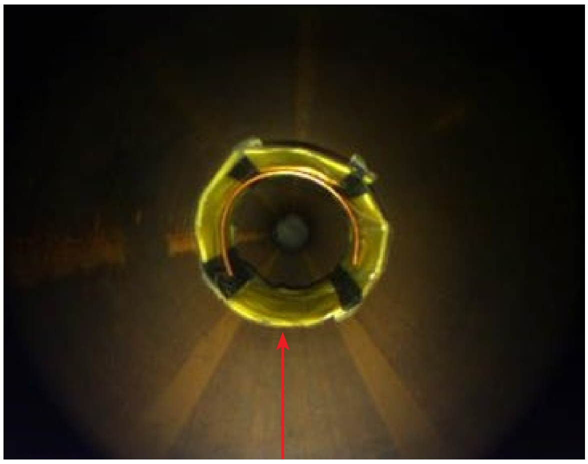


Figure 7: Image of four mouse-pad uranium sources taped to the inner surface of the test pipe with an internal radius of 15 cm.

Assay of mouse-pad sources was estimated at 1.5-2 g ^{235}U per source, with 4 sources being taped together inside the pipe. Multiple runs of the quantification identified masses found in Table 4, which are all within the expected range of the declared values for this source configuration. While this does provide basic confirmation of well-functioning mass quantification, proper testing requires sources of established traceability.

Run Number	Calculated ^{235}U Mass (g)
1	7.29
2	7.12
3	7.53
4	7.58
5	7.41
6	7.43
7	7.53
8	7.21

Table 4: Results of multiple runs for quantification algorithm testing using PGDP mousepad sources.

8. Ongoing and Future Efforts

The v1.1 update to the iQS software is recently completed, and V&V testing is imminent. The updated SQA package has been rewritten to include new features found in v1.1 and any bugs and associated fixes identified from v1.0 testing.

Proper sheet sources of uranium need to be researched further. In all detector testing with uranium sources, the sheet sources have not been traceable standards, while traceable canister or rod sources do not accurately represent deposit shapes.

9. Conclusions

The iQS project, including the InCounter quantification system and the iQS-01 software suite, is designed to improve quality and efficiency of piped holdup measurements. Likewise, its vantage within the pipe also provides a unique, unattenuated view of holdup deposits for gamma measurement and visual inspection. The tests performed show that the InCounter can produce measurements much faster and in a more reproducible fashion than a handheld detection system, and the integrated quantification analyses provide masses of holdup with improved accuracy.

The success of the latest rounds of testing indicates the InCounter and iQS software are ready for a more detailed, formalized qualification and testing program. Once complete, the InCounter should be qualified for incorporation into a facility NDA measurement program that complies with QSNDA.

10. Acknowledgements

The authors would like to thank the many scientists, engineers, and operators who assisted on-site during testing procedures. In particular, Jim Hurst at PORTS and Brent Montgomery at PGDP for coordinating testing efforts at their respective sites. In addition, to the late Graham Walford for his work in laying the basis for the iQS project.

The iQS project is an internally funded project of Innovative Solutions Unlimited, LLC (inSolves).

11. Legal

11.1 Privacy regulations and protection of personal data

The authors agree that ESARDA may print their names, contact information, photographs, and article in the ESARDA Bulletin or any other ESARDA publications and when necessary for any other purposes connected with ESARDA activities.

11.2 Copyright

12. References

[1] Innovative Solutions Unlimited, LLC; *iQS-01 InCounter System Description*; Piketon, OH; 2018.

[2] Russo, P. A; *Gamma-Ray Measurements of Holdup Plant-Wide: Application Guide for Portable, Generalized Approach*; Los Alamos National Laboratory report LA-14206; 2005.

[3] Walford, G., Dewji, S., DuBose, F., Roberts, D., Nicholson, A., Croft, S., Branney, S., Peacock, G., Miller, L., and Chapman, J.; *Quality Improvements in Calibration and Use of Portable Holdup Detectors for Uranium and Plutonium Fissile Deposits*; Presented at the INMM annual meeting; Indian Wells, CA; 2015.

[4] Peacock, G; *Development of Improved Methods for Modeling Piped Fissile Holdup Material* (Master's Thesis); Manuscript in preparation for the Georgia Institute of Technology; Atlanta, GA; 2018.

[5] Innovative Solutions Unlimited, LLC; *iQS-01 InCounter Technical Reference*; Piketon, OH; 2018.

[6] *Quality System for Nondestructive Assay Characterization*; U.S. Department of Energy Portsmouth/Paducah Project Office report DOE/PPPO/03-0235&D0; Rev. 0; 2011.

[7] Innovative Solutions Unlimited, LLC; *QSNDA Checklist for iQS and InCounter Applications v1.2*; Software quality assurance documentation; Piketon, OH; 2018.

[8] Innovative Solutions Unlimited, LLC; *iQS-01 InCounter User Manual*; Piketon, OH; 2018.

[9] *Test Plan for Pipe Crawler iQS-01 InCounter*; Internal test plan at Fluor-BWXT Portsmouth: TP-ND-2017-0001, Rev. 0; 2017.

[10] Tackett, T. and Taulbee, T; *FBP Surveillance Plan and Checklist: CM-SRV-FY17-5946*; Internal test result report at Fluor-BWXT Portsmouth; 2017.

[11] Innovative Solutions Unlimited, LLC; *Individual Test Plan & Report: InCounter Comprehensive Test for InCounter v1.0*; V&V result report ID #0.1-12130001-2; Piketon, OH; 2017.

[12] X-5 Monte Carlo Team; "MCNP - Version 5, Vol. I: Overview and Theory", LA-UR-03-1987; 2003.

KJ-BB-18-057-EN-N

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